



## Durham E-Theses

---

### *Personal identity and human animals: a new history and theory*

Southgate, Nicholas Charles James

#### How to cite:

---

Southgate, Nicholas Charles James (1999) *Personal identity and human animals: a new history and theory*, Durham theses, Durham University. Available at Durham E-Theses Online:  
<http://etheses.dur.ac.uk/4537/>

#### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

# PERSONAL IDENTITY AND HUMAN ANIMALS

## A NEW HISTORY AND THEORY

*Nicholas Charles James Southgate*

### ABSTRACT

The contemporary personal identity debate has divided into two entrenched positions. One supports the supposedly naïve and unpopular Bodily Criterion (the view that personal identity requires physical continuity). The other school is the Psychological Criterion (the view that personal identity requires psychological continuity). This has acquired the status of virtual orthodoxy.

The British Empiricists, John Locke and David Hume, are both supposed to give historical weight to this orthodoxy. This thesis argues this is a dramatic misrepresentation of history. Locke is supposed to found the personal identity debate in his *Essay Concerning Human Understanding*, arguing that personal identity is sameness of consciousness. It is argued that Locke in fact responds to a prevalent Cartesian View, called here the Compositional Account. The Compositional Account is the belief that a Human Being is composed of a Mind and a Body. Hume, in responding to Locke, is also responding to the Compositional Account. In opposition to widely established readings both philosophers are argued to be highly sympathetic to the Compositional Account.

Chapter 1 establishes Descartes' version of the Compositional Account and explains why Descartes needs no philosophical treatment of personal identity. These problems emerge only for the Empiricists, Locke and Hume. Locke's sympathies for the Compositional Account are established in Chapter 2, drawing on material prior to the *Essay* and normally uncited passages in the *Essay*. Chapter 3 argues that Hume presumed the Compositional Account in his *Treatise Concerning Human Nature*. This is argued to explain Hume's famous later recantation of his theory.

The thesis concludes by sketching a role for the Compositional Account in contemporary debate. The Compositional Account is argued to give strong support to a recently developed position known as Animalism. This provides the conceptual materials to move beyond the orthodox dichotomy between the Bodily Criterion and the Psychological Criterion.

University of Durham

Department of Philosophy

1999

Submitted for the degree of Ph.D.

# PERSONAL IDENTITY AND HUMAN ANIMALS

A NEW HISTORY AND THEORY

*Nicholas Charles James Southgate*

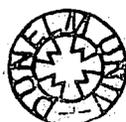
The copyright of this thesis rests with the author. No quotation from it should be published in any form, including Electronic and the Internet, without the author's prior written consent. All information derived from this thesis must be acknowledged appropriately.

University of Durham

Department of Philosophy

1999

Submitted for the degree of Ph.D.



19 JUN 2001

## TABLE OF CONTENTS

<i>Declaration</i> .....	iii
<i>Acknowledgements</i> .....	iv
INTRODUCTION.....	1
<i>Chapter 1</i>	
DESCARTES ON PERSONAL IDENTITY.....	7
I <i>Scholastic Influences on Descartes</i> .....	9
II <i>On Our Nature as Union of Mind and Body</i> .....	12
III <i>Substance Dualism</i> .....	14
IV <i>Two's Company, Three's a Crowd: The Place of Sensations</i> .....	17
V <i>Descartes' Substantial Principles</i> .....	23
VI <i>The Role of the Cogito Reconsidered</i> .....	31
VII <i>Souls, Human Beings &amp; Persons</i> .....	34
<i>Chapter 2</i>	
LOCKE ON PERSONAL IDENTITY.....	37
I <i>Are Lockean Persons Substances?</i> .....	39
II <i>Locke on Substance &amp; Identity</i> .....	43
III <i>Locke's Notion of Man</i> .....	51
IV <i>Locke &amp; The Rational Parrot</i> .....	62
V <i>The Famous Definition &amp; A New Interpretation</i> .....	65
VI <i>The Compositional Reconstruction of the Famous Definition</i> .....	74
VII <i>Locke &amp; Animalism &amp; The Man again</i> .....	84
<i>Chapter 3</i>	
HUME ON PERSONAL IDENTITY.....	86
I <i>Hume's Complaints &amp; The Self Of The Metaphysicians</i> .....	91
II <i>Hume on Identity</i> .....	97
III <i>The Contrasting Roles of Reason and the Imagination</i> .....	100
IV <i>Hume's Analysis of the Identity of Body</i> .....	104
V <i>Hume's Account of The Principium Individuationis</i> .....	112
VI <i>Identity in Hume's 'Of Personal Identity'</i> .....	123
VII <i>Hume's Explanation of Personal Identity</i> .....	132
VIII <i>The Failure Of Introspection</i> .....	136
IX <i>Three Relations</i> .....	140
X <i>The True Idea of The Human Mind</i> .....	143
XI <i>Hume's Missing Third Principle</i> .....	153
XII <i>Hume and the Compositional Account</i> .....	159
CONCLUDING POSTSCRIPT.....	164
I <i>Difficulties in Framing the Problem</i> .....	165
II <i>Dissolving The New Dualism</i> .....	169
III <i>Two Types of Animalism</i> .....	171
IV <i>Deflationary Animalism &amp; The Compositional Account</i> .....	177
BIBLIOGRAPHY.....	179

## DECLARATION

I confirm that this thesis conforms with the prescribed word length for the degree of Ph.D. for which it is submitted.

I confirm that no part of the material has been submitted by me for a degree in this or any other University. No material has been generated by joint work. In all cases material from the work of others has been acknowledged and quotations and paraphrases suitably indicated.

The copyright of this thesis rests with the author. No quotation from it should be published without prior consent and information derived from it should be acknowledged.



Nicholas Charles James Southgate

Date 12<sup>th</sup> May 2000

## ACKNOWLEDGEMENTS

As with any project of the longevity and magnitude of a thesis many people have contributed and should be thanked and acknowledged. That weaknesses remain in the final work is my responsibility alone. That there are not more is testimony to the efforts of those listed below.

First I should record the generosity of the University of Durham in awarding me the Studentship which made my studies possible at all.

I should acknowledge the support of Durham's Philosophy Department. I have been fortunate enough to study for two degrees at Durham. The Department's friendliness and warmth have made the experience both as undergraduate and postgraduate, student and teacher, an enriching and pleasurable one. My first tutor David Whewall cultivated my initial fascination with philosophy. The rest of the department ensured it took firm root and flourished. I know the department continues to make students feel as I do. I hope it will do so for many more years.

The department also encouraged me to tutor, and latterly lecture, during my studies. I gained much from the experience. For encouragement with tutoring I thank Chris Long. For giving me the opportunity to lecture I thank Andy Hamilton and Robin Hendry. I hope my own teaching matched the high standards I received from the department. I am greatly appreciative of having been given the opportunity to try.

Especial thanks are owed to my supervisor, Jonathan Lowe. Jonathan sets an example as a philosopher that is both something to aspire to and something to be inspired by. His comments on my work were always promptly delivered, unfailingly helpful, enviably well-informed and exemplary in their perception and insight. What is more, he always showed a faith in my abilities that meant that any doubts I entertained about completing the thesis successfully were always short-lived. My thanks to him cannot be warm enough.

My examiners Andy Hamilton and Harold Noonan deserve great thanks for their efforts. It is a great credit to them that I remember my viva as enjoyable and stimulating rather than nerve wracking. It is particularly fitting that Harold Noonan examined my thesis. It was his book *Personal Identity* that introduced me to the topic in my first term as an undergraduate. His writings have been a constant source of

illumination throughout my studies and I feel particularly honoured to receive his seal of approval. Andy Hamilton carried out his duties as internal examiner with notably efficiency. Further, his detailed and perceptive comments helped me to see where I might take the ideas in the thesis in the future.

Early versions of parts of the thesis were read at meetings of the Durham Department's Staff Seminar. I thank the audiences at those seminars for their attention and questions. Particular thanks in this regard are owed to Paul McDonald who read and commented extensively on the chapters on Descartes and Locke. Paul's comments improved immeasurably both the thesis as a piece of philosophy and myself as a philosopher. I also owe thanks to the EIDOS post-grad community. Having people to exchange thesis stories with made the process a lot easier.

I also presented versions of the arguments to Durham University's Student Philosophical Society, as well as in undergraduate lectures and tutorials. The challenge of presenting arguments clearly and persuasively to these audiences always served to sharpen my thoughts. There have been many students who have insisted that I explain why the philosophical problem of personal identity is an important one. Their demands were reasonable and have kept my studies grounded. I thank them for their interest, persistence and occasional indignation.

Lastly, I should thank my friends and family. They tolerated me giving over my life to the curious and wonderful pursuit of writing a philosophical thesis.

NCJS

*Durham, May 2000*

## INTRODUCTION

Historically, there have been two (allegedly) popular criteria of personal identity. The simpler of these takes it to be a necessary (conceptual) truth that people are to be reidentified in terms of their bodies. This criterion is much talked about, but I am unaware of any recent philosopher who has endorsed it, and I am inclined to regard it as a straw man. The only genuinely popular kind of criterion of personal identity is a mentalistic one.<sup>1</sup>

When Pollock wrote this ten years ago he was largely correct. Psychological accounts of personal identity were, just about, the only ones pursued by philosophers.<sup>2</sup> What is more this outcome was justified, indeed inevitable, given the history of the subject. Locke, regarded as the progenitor of the whole debate, held a theory identifying the person with consciousness.<sup>3</sup> Hume, Locke's extremist successor, though disagreeing with Locke on so much, held a theory in which each person was merely a bundle of perceptions.<sup>4</sup> History and Modern debate did indeed seem to concur. Pollock seemed correct. No-one took the physical criterion of personal identity seriously. What is more, no-one ever had.

In the last ten years a few philosophers have started to take the Physical Criterion of Personal Identity seriously. Bernard Williams had argued for a Bodily Criterion as long ago as 1959<sup>5</sup>, while Peter Strawson's *Individuals*<sup>6</sup> contained influential arguments that suggested our notion of person had an irreducible physical component. These arguments have been

---

<sup>1</sup> Pollock (1989) p. 30. Olson also cites this passage, Olson (1997a) p. 19. For a similar approach to the problem see Penelhum (1967) p. 95ff. Noonan also writes "the Bodily Criterion of personal identity has not proved popular with philosophers." Noonan (1989) p.3

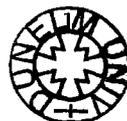
<sup>2</sup> Writers who pursued personal identity in terms of the soul are one possible exception (e.g. Swinburne) for it is not clear that the soul is a psychological entity in the same way persons are for subscribers to the Psychological Criterion of Personal Identity. The Brain Criterion is not an exception. The Brain Criterion is merely a (bad) mix between Physicalism in philosophy of mind and the Psychological Criterion of Personal Identity. The Brain Criterion is the assertion that a person is just a psychological entity that must be embodied as a certain sort of thing (e.g. a brain). Such an account remains a psychological one despite the embodiment restriction. For arguments I largely agree with see Olson (1997a), chap. 1. The situation is no better if embodiment must occur in a whole human being. This merely produces the New Dualism. See van Inwagen (1997), Burke (1997b), Olson (1997a).

<sup>3</sup> Writers who think this of Locke always cite *Essay* 335:11-14 (Book II, Chap XXVII, §9). There Locke defines a person as "a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it". For a full discussion of this passage see Chapter 2, §V

<sup>4</sup> E.g. Waxman (1994) n7p. 330 "Hume's account... is framed in exclusively psychological, subjective terms", Green (1999) p. 104 "Hume assumed that personal identity over time is equivalent to the identity of a person's mind over time. He paid no attention at all to the possibility that personal identity over time might consist in the identity of a human body over time.", Fang (1984) p. 59 "It is well-known that Hume has a quite unusual theory of personal identity. For him, personal identity is but the identity of mind."

<sup>5</sup> See Williams (1956). For further development of Williams' argument see Williams (1961b), (1966), (1970a), and (1973b).

<sup>6</sup> Strawson (1959). For a contemporary response see Williams (1961a). For a more recent, and lengthier, response see McCall (1990). Smart (1977) also raises some interesting points.



further advanced by Paul Snowdon<sup>7</sup>, Peter van Inwagen<sup>8</sup> and Eric Olson<sup>9</sup>. These writers, normally known as Animalists, replace the notion of *body* with that of *animal*. The difference between these writers and their Bodily Criterion predecessors rests in this emphasis on persons as *biological* entities as opposed to merely *physical* entities.

These writers find themselves ranged against two broad camps within those who subscribe to the Psychological Criterion. Those who follow Locke's insights are the Neo-Lockeans. Those who follow Hume's insights are the Neo-Humeans. The Neo-Lockeans prefer to attempt to offer a definition of person, and then attempt to provide non-circular identity conditions for this definition. The most well known Neo-Lockean is David Wiggins.<sup>10</sup> Although some feel Wiggins is some sort of Animalist, his writings are in the spirit of Locke. Wiggins suggests our notion of person contains a lengthy (and indeterminate) list of conditions, some of which are physical, but the core of which are psychological.<sup>11</sup>

The Neo-Humeans believe they follow Hume in denying that persons are independent entities of the sort Wiggins and the Neo-Lockeans investigate. Persons are bundles or constructions of some sort. The aim of philosophical investigation of personal identity is to elucidate what bundles together the more basic constituents of persons. The most influential Neo-Humean writer is without doubt Derek Parfit. Indeed the theories put forward in his *Reasons and Persons*<sup>12</sup> have the status of established orthodoxy in some quarters.

Neo-Lockeans and Neo-Humeans alike often proclaim that they only write in the spirit of their philosophical muses. However, it seems implicitly assumed that Locke and Hume would not be appalled by what has been propounded in their names. Adherents of the Psychological Criterion comfortably assume they have not only the arguments but also

---

<sup>7</sup> See Snowdon (1989), (1991) and (1998).

<sup>8</sup> See van Inwagen (1990)

<sup>9</sup> See Olson (1994), (1995a), (1995b), (1997a), (1997b)

<sup>10</sup> See Wiggins (1976) and (1980)

<sup>11</sup> Those who claim that Wiggins is an animalist regard him as a Neo-Aristotelian. On this interpretation Wiggins' theory of personal identity is an attempt to refine the Aristotelian claim that we are *rational animals*. Wiggins' project seems close to the Aristotelian one. Wiggins thinks we are animals with certain typical forms of mental behaviour. However, the two are importantly different because Wiggins regards the *core* qualities of personhood as being psychological. Wiggins' is persuaded of the importance of psychological factors for the same reasons as other Neo-Lockeans, i.e. causal continuity. Wiggins is therefore a Neo-Lockean who insists on embodiment (and further restricts this embodiment in typical cases to animal embodiment). A true Neo-Aristotelian holds we are animals of a certain kind first and foremost and is unimpressed by issues of psychological continuity. The Neo-Aristotelian view is best expressed by the strongest form of Animalism which claims that psychology is *irrelevant* to personal identity (though this is not incompatible with the claim that the animals that typically constitute persons are typically rational and exhibit a high-degree of psychological continuity). See Olson (1994), (1995a) and (1997a) chap. 1 & 3.

<sup>12</sup> Parfit (1984) esp. Part III

history on their side. In contrast Animalism looks to be nothing more than a cult among a few philosophical refuseniks.

It is the contention of this thesis that this picture is utterly mistaken. The Physical Criterion is not the straw man Pollock and his ilk suggest. In fact, the Physical Criterion has guided and influenced the entire history of the personal identity debate. In his guide to personal identity Harold Noonan is typical in his assertion that the Physical Criterion is a philosophically naïve position, quickly dismissed after a few moments reflection.<sup>13</sup> This slander of naïveté dramatically underplays the influence the Physical Criterion has had in philosophy's history. If our pre-philosophical assumption is that each of us is simply and primarily a human being then this view is worth taking seriously. Argument is needed to show that Locke and Hume did not take this position seriously. It seems unlikely that such careful and authoritative writers, both well aware of advances in the physical and medical sciences, both with a keen interest in Natural History, History and Anthropology, would dismiss such a persuasive and natural pre-philosophical view without argument.

The evidence, upon examination, shows overwhelmingly that they did not. Instead, they both attempted to *embrace* the Physical Criterion with their theories, not, as orthodox history would suggest, attack it. The traditional difficulties levelled in interpreting their theories arise only because modern commentators have not allowed for this trend in Locke and Hume's writings. The consequence of this error has been far reaching. The Physical Criterion has been reduced to nothing, and in its place two schools have been founded, which bear the names, but do not contain the theories, of Locke and Hume.

It is also frequently assumed that the personal identity debate only begins with the Empiricism of Locke and Hume. The contention of the first chapter is that the Empirical philosophy only gave rise to new difficulties about persons, not to the debate itself. Prior to Empiricism writers had assumed that person was a primitive notion. This position is explored by looking at Descartes' writings. Descartes has no explicit discussion of personal identity. However, this is because his philosophy avoids the problems Empiricism raises. Descartes assumed that persons were Human Animals. Each Human is *composed of a mind and a body*. By asserting that this union is primitive Descartes avoids the need to provide a theory of personal identity.<sup>14</sup>

---

<sup>13</sup> Noonan (1989) p.2ff

<sup>14</sup> By primitive Descartes means that experience of the union is available to us unmediated by higher cognition. Put differently, the experience of the union necessarily accompanies our thought. For further discussion see note 42 below.

Personal identity before Locke and Hume, therefore, presumed a *Compositional Account*. The second chapter re-examines Locke's theory as a response to the Compositional Account. Locke does not share Descartes' confidence about the knowledge we have of substance. However, if Locke could overcome the difficulties he feels afflict our knowledge of substance he would be able to accept the Compositional Account. Chapter 2 presents an interpretation of Locke that demonstrates how Locke both accepted and presumed the Compositional Account. His theory is not a psychological one. Throughout Locke gives a role to the physical, and does nothing to exclude physicality from his notion of person.

Hume clearly follows in Locke's footsteps. Yet, whereas Locke attempted to rehabilitate and utilise the rationalist notion of substance, Hume rejects it outright. Despite this rejection Hume's attempts to explain personal identity still presume that persons are composed of both a mind and a body. The third chapter reconstructs Hume's theory of identity and personal identity to demonstrate that not only does Hume not exclude the physical from his account of persons, he explicitly relies on it.

It might be asked what profit there is in establishing the Compositional Account. First, there is the simple fact that it seems to be the most historically accurate and insightful view. All three of the writers considered in this thesis have been interpreted so as to exclude an important element of their philosophies. Establishing the prevalence of the Compositional Account makes the wider theories of these great writers seem far more sympathetic to modern readers. These great writers deserve to be considered carefully, and do not deserve to carry the weight of convenient and crude caricature.

However, it is not the aim of this thesis to establish beyond doubt that the Compositional Account was held by Descartes, Locke and Hume (although in Descartes' case this seems almost incontrovertibly to be true). The Compositional Account should be adopted as a background assumption to better both understanding of these writers and the problem of personal identity. The Compositional Account serves not just to explain what is present in these texts. It also explains key absences; Descartes' neglect of the problem of personal identity, Locke's failure to explicitly proclaim a purely mnemonic or psychological theory, and the supposedly groundless extremity of Hume's theory.

The Compositional Account does, as will be demonstrated, fit the texts very well. Nonetheless, this thesis is a work of philosophy rather than of the history of ideas. Historians may find specifics of context make the Compositional Account seem implausible despite the support to be found in the texts. However, it is not the aim of this thesis to

establish that Descartes, Locke and Hume would have espoused these interpretations of their works as the correct ones. Instead the aim is to force a reconsideration of themes ignored in existing interpretations of their writings. The weight of evidence presented below demonstrates that orthodox views about each writer cannot be complacently accepted. The fundamental intention is to present a sympathetic but provocative interpretation of these much discussed texts. The Compositional Account is sympathetic because it brings all three writers closer to pre-philosophical common-sense than any existing interpretation. If the Compositional Account is accepted none of them can be dismissed as mere historical curiosities valued only as mistakes to be used as starting points. The Compositional Account is provocative because it obliges a re-evaluation of all three writers.

Secondly, the Compositional Account can add a great deal to current debate about personal identity. The division between the Bodily Criterion and Psychological Criterion had stultified the debate. The Bodily Criterion, reduced to a parody, left the Psychological Criterion effectively unopposed. Yet, it is both ludicrous and philosophically undesirable that the physical aspect of personal experience was so easily and commonly ignored.

Animalism is one possible route beyond this dichotomy. Animalism has the advantage of identifying persons with a common and metaphysically robust class of objects, i.e. animals.<sup>15</sup> This identity claim is interesting because it can be informative even if it cannot be expressed in a non-circular fashion. Even if the definition of a person must contain reference to an animal on both sides, nothing about this precludes the possibility of animals being investigated empirically. For this reason Animalism is sometimes also known as the Biological View because biology becomes the science of personal investigation.

The Compositional Account is of interest because it is the obvious precursor to Animalism. Persons and Humans are identified, as with Animalism. However, instead of understanding the Human Animal purely (and ultimately) biologically<sup>16</sup> a Human Being is understood as a composite of a mind and a body. Further, the existence of such Composites can be construed as a primitive or fundamental fact about the world.<sup>17</sup> This is interesting because

---

<sup>15</sup> Being metaphysically robust is a relative notion, i.e. some things are more metaphysically robust than others. Some metaphysicians talk of objects that are clearly the result of gerrymandering the borders of the universe (e.g. mereological sums of the Eiffel Tower and my foot). In comparison animals are more metaphysically straightforward and robust. They are also, in comparison, animals are metaphysically well-behaved, i.e. there are non-arbitrary conditions for their survival. The same cannot be said for gerrymandered objects and many artefacts. Some of these (e.g. Theseus' Ship) are notorious for being metaphysically badly behaved.

<sup>16</sup> As it must be to avoid the New Dualism. See van Inwagen (1997).

<sup>17</sup> Strawson (1959) argues along these lines.

persons are everyday objects in our ontology that we usually understand without complication. Persons should therefore be uncomplicated parts of any ontology to reflect the uncomplicated place they have in our folk practice. The Compositional Account captures this, and as will emerge, is adopted for just these reasons by Descartes, Locke and Hume. The implications for contemporary debate are sketched in a concluding postscript.

## CHAPTER 1

### DESCARTES ON PERSONAL IDENTITY

To many writers it would seem almost absurd to look to Descartes for a theory of personal identity.<sup>18</sup> There is no explicit theory of personal identity to be found in any of Descartes' texts. Indeed Descartes barely addressed the issue of identity,<sup>19</sup> let alone personal identity, and never uses the term 'person' in any sort of philosophical or technical sense.

This void of explicit theory has been filled by a widely accepted version of likely Cartesian theory. Personal identity is identity of the thinking thing that each of us essentially is. It is clear how this is straightforwardly derived from the *cogito* and the Cartesian doctrine that thinking is performed by a thinking *thing*. The Cartesian Mind (or Soul)<sup>20</sup> is indivisible and indestructible. Consequently its identity is a primitive fact. If personal identity reduces to identity of the Soul then this too is a primitive fact.

This view has been influential in commentaries. It has become particularly important in discussions of personal identity because the 'Cartesian Ego' has been isolated (and demonised) by Derek Parfit as the prime exemplar of one flawed conception of personal identity. Parfit writes, "besides assuming that every thought must have a thinker, Descartes assumed that a thinker must be a Pure Ego, or spiritual substance. A Cartesian Pure Ego is the clearest case of a separately existing entity, distinct from the brain and body."<sup>21</sup> Following a tradition owing to Ryle,<sup>22</sup> Parfit's accusation is that Cartesian Egos are wrongly postulated to explain personal identity.<sup>23</sup>

One part of this conception of Descartes can be upheld. Descartes does resolve the problem of personal identity into something other than questions about persons. Descartes has no need for a technical notion of a person. He does not introduce a specific philosophical

---

<sup>18</sup> For example Markie comments "Descartes seems to be insensitive to the distinction between a mind's continuing as the same substance through changes in its mental states and its continuing as the same person through such changes." Markie (1994) p. 84n24. This, of course, assumes that minds are persons.

<sup>19</sup> See Thiel (1997), especially pp. 356-361, and the references therein, for useful background discussion.

<sup>20</sup> Descartes does use the two terms interchangeable, but only on his own terms. He rejects the scholastic term *anima* in favour of *mens*. Descartes' conception of Soul (*Mens*), therefore, is interchangeable only with the Scholastic notion of the intellectual soul. In as much as Descartes has a use for the notion of an animal or vegetative soul, he explains these in mechanical terms.

<sup>21</sup> Parfit (1984) p. 224. Parfit currently believes there might have been Cartesian Egos, but in fact there are not. See Parfit 1995 pp. 16-19.

<sup>22</sup> See *The Concept of Mind*, Ryle (1949)

<sup>23</sup> Other examples include Eric Matthews comment "[I]t is an integral part of the Cartesian account that the self or person is to be identified with an immaterial substance." (Matthews 1977, p.11), and Hooker's attempt to "reconstruct Descartes' argument for person-body distinctness." (Hooker 1978b, p.183)

apparatus in response to the collection of problems that go up to make the personal identity debate. This distinguishes him significantly and interestingly from Locke and Hume. The rest of the Parfitian caricature of Descartes is quite wrong. It is the aim of this chapter to lay out an accurate account of Descartes' views and demonstrate why Descartes felt no need to explicitly deal with Lockean type problems.

This will involve four steps. Firstly, it is revealing to map out the Scholastic doctrines concerning the soul and persons to understand the influences Descartes inherited. Secondly Descartes' own position can be explored through his own account of substance. Third, Descartes' conception of substance can be viewed via his theory of nature. Finally, it will become clear that issues of personal identity can be resolved into questions about the soul, the body, and the union of the mind and body (a human animal) without use of the further notion of a person.

# I

## SCHOLASTIC INFLUENCES ON DESCARTES

The problem of personal identity has three elder and more distinguished theological brothers, the doctrines of the Resurrection, Immortality of the Soul and the Trinity. The first two problems can be seen as the problem of personal identity painted on a more expansive canvas. Whereas debate about personal identity primarily concerns itself with how one persists through an earthly career, such an account will have implications about what happens after death. Given that it is a central tenet of the Christian Church that one does survive bodily death Christians are likely to favour a theory of personal identity that is compatible with post-mortem existence. Survival of bodily death is a pre-requisite of the Resurrection, and Immortality of the Soul provides this most economically.

The Trinity raises the difficulty of explaining how three things, the Father, Son and Holy Spirit can be understood to all reside in one thing, God. The notion of person has been used to explicate this. Boethius, for example, claims that God is "one essence, three substances, and three persons of the Godhead".<sup>24</sup> The Trinity is a peculiar problem for Christian theology. It is revealing to an account of personal identity only in as much as the notion of person is stretched to explain the Trinity. The doctrines of the Immortality of the Soul and The Resurrection are more pertinent. For one, they are common to Christian and non-Christian eschatologies. Secondly, both are concerned with how a person (or whatever entity constitutes a person) can be said to persist through a change in mode of existence, i.e. from being alive to being dead. The Resurrection adds the problem of a person apparently existing intermittently and yet remaining identical. There is significance beyond theology to settling whether persons can be said to exist intermittently, and whether persons can undergo changes in modes of existence, or substance.

Scholastic definitions take a person to be the Aristotelian first substance whose essence consists in rationality. As substance has both form and matter it is obviously possible to claim that either the person is identical with the form, or the person is identical with the matter, or only with the combination of both. Aquinas hedges his bets; the soul is the form of man, but does not constitute a person on its own. However, unlike other forms the soul

---

<sup>24</sup> In *Liber contra Eutychen et Nestorium*. Cited in Thiel (1998b) p. 872ff.

can exist apart from the substance it constitutes.<sup>25</sup> The views of Boethius and Aquinas remained influential and well known via Suárez's writings.<sup>26</sup>

Although the Cartesians broke with the Scholastics on many issues they inherited an important principle from them about the nature of humans. The Scholastic view can be seen as offering a Compositional Account of a human person. A human person is regarded as being composed of a body and a soul, and identified with neither. The Aristotelian influence on the Scholastics meant that this composition was understood in terms of form and matter.

Descartes does at times adopt Scholastic terms when discussing the union of mind and body.<sup>27</sup> However, his official interpretation of the Aristotelian Scholastic account is clearly put in the *Fifth Set of Replies*:

[P]rimitive man probably did not distinguish between ... the principle by which we are nourished and grow ... [and] the principle in virtue of which we think. He therefore used the single term 'soul' to apply to both. [...] I, by contrast, realising that the principle by which we are nourished is wholly different – different in kind – from that in virtue of which we think, have said that the term 'soul', when it is used to refer to both these principles, is ambiguous. If we are to take 'soul' in its special sense, as meaning the 'first actuality' or 'principal form of man', then the term must be understood to apply only to the principle in virtue of which we think; and to avoid ambiguity I have as far as possible used the term 'mind' for this.<sup>28</sup>

Where Descartes differed from the Scholastics was that soul and body were related not as form to matter, but as two distinct substances brought together in union. It is this union that is central to understanding why Descartes believed he faced no difficulty about personal identity. This claim runs counter to the prevalent notion that it is the thinking thing of the *cogito* that accounts for personal identity. The following sections, therefore, will defend this interpretation of Descartes. §II will sketch the implications of regarding Human Beings as a Union of a mind and a body. This introduces Descartes' account of substance. §III examines this account, while §IV explains why a current trend in Cartesian exegesis,

---

<sup>25</sup> Thiel (1998b) p. 871

<sup>26</sup> Thiel (1998b) p. 871

<sup>27</sup> For example in *Rules For The Direction of The Mind*, Rule 12 (AT X: 411; CSM I: 40) All references are to *The Philosophical Writings of Descartes* (Descartes 1984, 1985 & 1991) edited and translated by Cottingham, Stoothoff, Murdoch & Kenny. Each reference takes the following form: Name of Work or Letter and any internal subdivision (e.g. *Meditation VI*) then reference to the standard Adam & Tannery Collected Works, by volume then page (e.g. AT X: 411) then reference to the standard English translation by volume and page (e.g. CSM I:40).

<sup>28</sup> *Fifth Set of Replies* (AT VII: 356; CSM II: 246). CSM add in a footnote that Descartes uses the standard Scholastic terms for 'first actuality' and 'principal form of man' derived from *De Anima* II:2.

Trialsm, should be dismissed. This done §V presents a preferred interpretation of Descartes. §VI then explains how the *Cogito* argument needs to be absorbed into this new framework. Finally, in §VII, Descartes' commitment to the Compositional Account can be asserted and Descartes' conclusions can then be contrasted with his Empiricist successors. The account that emerges is a robust and appealing one. It has intuitive appeal and also casts a very different light on post-Cartesian developments in thinking about personal identity. As Descartes' immediate critics did not believe Descartes to believe in 'Cartesian Egos' their comments will be best understood if we understand Descartes more as they did. This is all the more important because key among those critics was Locke.

## II

### ON OUR NATURE AS UNION OF MIND AND BODY

It does not seem to me that the human mind is capable of forming a very distinct conception of both the distinction between the soul and the body and their union; for to do this it is necessary to conceive them as a single thing and at the same time to conceive them as two things; and this is absurd.<sup>29</sup>

Although it is Descartes' avowed claim that our nature is as a union of mind and body it is also clear that he regarded this as an inherently obscure part of his philosophy. It would be improper to regard this as a knockdown criticism. Modern science in all its glory is no more clear about the nature of human beings; and Descartes has the distinct advantage of being clear about why our notions concerning our own natures must be mostly clouded and only occasionally penetrated by the light of reason.

The union of mind and body emerges as an important notion for Descartes in the Sixth Meditation and is discussed in replies and letters after that.<sup>30</sup> It is the closing and final claim of the *Meditations*. Despite Descartes choosing this as his end point it seems oddly neglected that this is his final conclusion. The *cogito* argument of the Second Meditation has generated far more interest as the *locus* of Descartes' claim as to our existence. This has occurred for two reasons. One is the interest of modern commentators in Descartes' epistemology<sup>31</sup>. The *cogito* is a far more fertile field for criticism than the claim that we are by nature human beings. The second is confusion about Descartes' subtle distinction between our natures and our essences. Nature and essence are taken to be synonymous by most modern writers. For Descartes, however, there is a strict distinction. Our essence is that which belongs to us by necessity of being a certain substance, i.e. thinking is essential to minds, and extension is essential to matter. Nature, on the other hand, is not understood by way of logical or metaphysical enquiry. Nature is the set of laws God has ordained that the Universe should run by.

---

<sup>29</sup> *Fifth Set of Objections* (AT III: 693; CSMK III: 227)

<sup>30</sup> *The Meditations* are primarily about human beings proper status as knowing beings. The union between mind and body is discussed in texts prior to *The Meditations*, but it is only in *The Meditations* that the union takes its fullest place in Descartes' philosophical scheme. Certainly it is *The Meditations* which is also cited as the source for alternative (mis)interpretations of Descartes' opinions. For an exhaustive survey of Descartes' mentions of mind-body union prior to the *Meditations* see the appendix to Voss (1994).

<sup>31</sup> See Cottingham's Introduction (1994b), Sorrell (1994), Markie (1994) and also Baker and Morris (1997) for criticisms to this effect.

At the level of *simple* substances the difference between nature and essence collapses. If there were only simple substances then nature would always conform to essence. Nature is significant in understanding *composite* substances. It will be discussed next how the union of mind and body is for Descartes a composite substance, while still being a single unit. However, it is clear that Descartes claims in Meditation II that our *essence* is most easily known as a thinking thing (thought after all being the medium of knowledge).<sup>32</sup> This is not in conflict with claiming that our *nature* is one of union of mind and body, and that this nature is less easily known than our essence. Descartes' Compositional Account makes it consistent that as human beings we have one nature, but two substantial essences.

---

<sup>32</sup> It seems to me that at the heart of Descartes' thought about the immateriality of mind is the belief that matter cannot *know*. Executive mental states can never be constituted by even the most complex material states, because they are of an essentially different order. This claim seems neither outrageous nor archaic.

### III

#### SUBSTANCE DUALISM

Famously Descartes is a dualist, if not *the* dualist. However, as several writers have demonstrated, most notably Cottingham, there are several distinct hues of dualism. Firstly, one can distinguish dualisms about *things* and dualisms about *stuffs*. To be a dualist about things would be to claim that all things could be classified as one of two types. It is therefore open for the dualist of things to be a monist (or pluralist) about the stuff that goes up to make these things.<sup>33</sup> A dualist of stuffs claims that all things are made up of two sorts of stuff. It is therefore possible for a stuff dualist to claim that there can be three types of thing: two pure types, and one mixed type.<sup>34</sup>

This distinction between things and stuffs brings out a significant difference in the aims of substance ontologies. The distinction is employed here as a heuristic device to explore Descartes' texts. Therefore, one wants to avoid embroiling the distinction in the pursuit of metaphysical completeness. What is useful about the distinction is it captures how substance ontologies deal with notions of constitution. Stuff ontologies regard what constitutes entities to be most revealing of their nature. Conversely, thing ontologies regard this as a lesser issue. It is the nature of the things themselves that is of real interest. Clearly, the two questions of *what* constitutes entities and *how* they are constituted do not have to be separate. If substances are homoeomerous there is no distinction between the nature of the whole and the nature of its parts. Conversely, if some entities are meant to have no parts (Cartesian minds being one example) it is not clear how they can be said to be composed at all.

It is no surprise to find, therefore, that these confusions are present in Descartes' ontology. On the one hand he seems to be a dualist about things. There are only minds and bodies and unions of minds and bodies. On the other hand he seems to be a stuff dualist and hold that there are minds, bodies, and human beings that are unions of mental and physical stuffs. The weight of explicit textual evidence favours Descartes being a thing dualist. However, he would have made things easier for himself if he had been an explicit stuff dualist. If the two substances are stuffs, then a combination of the two could still plausibly

---

<sup>33</sup> Russell's neutral monism is one such theory.

<sup>34</sup> If the degree of mix can be made to distinguish a type then a pluralistic ontology is possible. Leibniz might be an extreme case. Each substance (monad) is a unique mix (reflection) of mental and physical. Hence the infinite possibilities of mix provide an infinite number of monads.

be regarded as one thing. If Descartes insists on regarding minds and body as distinct things, then he is forced to regard human beings as unions of two distinct things. This appears antagonistic to the sort of intimate 'intermingling'<sup>35</sup> that Descartes suggests typifies the union of minds and bodies. Ultimately it would seem to leave no place for human beings in Descartes' ontology.<sup>36</sup> The apparent equivocation pivots around two things: Descartes' definition of substance, and the status of 'substantial union' in human beings.

Matters are further complicated because Descartes defines substances in terms of their leading attributes, i.e. thought and extension. Neither of these attributes necessarily implies that stuff or things make up the world. This is clearly true of thought. Thought is not a thing, although it is things that think. Thought also does not obviously make up a kind of stuff. Indeed the terms 'thing' and 'stuff' seem best suited to concrete objects, like tables, lumps of wax, and human beings. Descartes, of course, is clear that thought is essentially unextended. However, it is not clear that Descartes regards extension as being derived from concrete objects. Instead extension is the quality of taking up a portion of space. This is defined independently of any objects that might be in that space. Further, because Descartes is vague about providing identity conditions for material objects, it is not clear that he regards them as real in any strict sense.

Descartes could plausibly be seen as holding that there are merely two modes of apprehension available to us. We can either apprehend things as being extended or as being mental. This, of course, is the direction in which Spinoza took his metaphysics. Like Spinoza, Descartes can be seen as holding that there are no substances other than God. Descartes does go part way to accepting this with his doctrine that all substances are dependent ultimately on God.

However, to interpret Descartes as a witless Spinozist would be wrong. While Descartes may not have pursued all his principles to exhaustion, it is clear that he maintained a common-sense view about concrete objects. His lack of attention to these issues is itself evidence that he felt no need to justify the obvious claim that lumps of wax and tables are metaphysically reasonably well-behaved objects.<sup>37</sup> Descartes does not regard his division between the material and the immaterial as two ways of regarding some intangible

---

<sup>35</sup> Descartes suggests the mind and body 'intermingle' in Meditation VI (AT VII: 81; CSM II: 56). For more on this see §IV in this chapter.

<sup>36</sup> This is Voss's conclusion. See Voss (1994)

<sup>37</sup> For example he talks about mice being *entia per se* in his letter to Regius, December 1641 (CSMKIII: 200). Also in the *Meditations* Descartes calls a stone a substance (AT VII: 48; CSM II: 30). For a discussion of the notion of being metaphysically well-behaved see note 15 above.

unknowable substratum. Descartes is clear that the mind is more *distinctly* known than the body.<sup>38</sup> However, both body and mind can both come to be known fully and properly once their natures are properly conceived. One can have clear and distinct knowledge of both minds and bodies. Lack of clarity or distinctness is caused not by any inherently obfuscating factors between substance and ourselves, nor any intangibility about substance.<sup>39</sup>

It is this insistence on the possibility of clear and distinct knowledge of minds and bodies that means Descartes is best understood via the distinction between stuff and thing dualisms. Our knowledge of substances should allow us to know whether they are things or stuffs. As a consequence one would be better able to understand how they can be said to be in union.

It might be objected that while this investigation is plausible for material objects it might not be so obviously applicable to minds. Minds are not obviously things, at least not in the way hatboxes are. Likewise, thought is not obviously a kind of stuff, at least not in the way treacle is. However, one simply has to accept the idea of unextended, immaterial things or stuff. This is no more vague than Descartes' frustrating definition of thinking substance in negative contrast to material substance. This presumption seems to be most in tune with Descartes' intuitions. The other alternatives involve mystifying the objects of our knowledge, and that is counter to Descartes' aims.

The next section uses this distinction to explore a recent development in Cartesian scholarship. This is the debate between interpretations that accept Descartes' official dualism and those that suggest this official position is fundamentally strained or even flawed.

---

<sup>38</sup> For discussion see §VI in this chapter.

<sup>39</sup> Although, Descartes does express reservations about the *union* of the substances however clearly they are known. See the quote at the beginning of §II above.

## IV

### TWO'S COMPANY, THREE'S A CROWD: THE PLACE OF SENSATIONS.

Cottingham<sup>40</sup> has suggested that Descartes was tempted towards, and should have embraced, an explicitly Trialistic as opposed to Dualistic ontology. Cottingham therefore concludes that Descartes' dualism is fundamentally in unresolved tension as Descartes has left it. Baker and Morris<sup>41</sup> suggest that properly understood there is no tension in Descartes' substance ontology and that if one recognises Descartes' equivocal use of 'necessity' and 'independence' in defining substances Descartes' Dualism is coherent. Clearly these writers cannot all be correct. Yet it is paramount to settle this issue for on it rests the coherence of Descartes' plausible claim that human nature consists in being a union of mind and body. Descartes adopts a number of principles to define substance, and these are often apparently at odds with each other. Once these principles are brought into line with each other it will be possible to outline how Descartes understood the union of mind and body.

The key to Descartes' difficulties is his adherence to the Scholastic principle that a property can only inhere in one substance. Despite their differences the Scholastic and Cartesian conceptions of substance have enough in common to both embrace this rule. It would seem nonsensical that a single instance of a quality could be shared by two substances (at least, if qualities are not universals). Two substances may instantiate exactly similar qualities, but this would involve two instances of the quality. The problem arises because human beings are substance-like and we are therefore prone to ascribe qualities to human beings as if they were substances. Descartes' claim that human beings are a union of two substances obliges him *either* to allow that we can ascribe qualities to a union of substances *or* to explain how this mistaken practice arises and offer a systematic account of how these ascriptions belong to two substances. Although there are other conceptions of substance available Descartes does not consider them. Given the Cartesian conception of substance (and its Scholastic heritage) abandoning the rule would be tantamount to rejecting the notion of substance itself. It is not, therefore, an option. This creates particular difficulties in Descartes' treatment of sensations. Unlike pure actions of the intellect (such as willing to love God) or purely reflex physical actions (such as pulling one's hand away from the fire) sensations are seemingly both mental and physical events. Having a pain in one's foot is both an episode that is about physical damage to the foot and the mental episode of judging oneself to have

---

<sup>40</sup> In his article 'Cartesian Trialism' Cottingham (1985) and also Cottingham (1986), pp. 127-32

<sup>41</sup> In Baker & Morris (1997)

a pain. Whereas willing to love God can be predicated only to the mind, and reflex actions only to the body, sensations seem to need to be ascribed to both mind and body.

Cottingham suggests that Descartes should have solved the problem by introducing a third substance. Each Cartesian substance has a defining attribute. For minds this is thought, and for material things extension. Sensation should be the defining attribute of animals. This would preserve the need to have each predicate terminate in only one substance.

Cottingham feels that sensations are an irreducible part of our experience. Any attempt to parse sensations into either mental or physical episodes (or some combination of the two) will lose some part of what it is for an episode to be a sensation. Yet, Cottingham's interpretation is unsatisfactory. Sensations, though central, are not fundamental in the same manner as thought and extended matter. The category of sensation and the notion of a human being although tightly interrelated are too loosely connected to provide a definition of substance in the way thought and extension can. Indeed Descartes derives his notion of sensation from the division of mind and body, and has no metaphysical notion of sensation in itself.<sup>42</sup> Whereas bodies could not exist without extension, and thoughts could not exist without thinkers, this does not seem to necessarily hold for human beings and sensations. Sensations could belong to creatures other than human beings, and human beings might exist without having sensations.<sup>43</sup> Descartes *does* link human beings and sensations together but not by necessity of substance. Instead sensations arise because of the natural necessity of the union of minds and bodies. Appealing to the natural order makes a state of affairs dependent on God's grace.<sup>44</sup> Although Descartes is happy to use this device (most notoriously in the alleged Cartesian Circle) he prefers to offer proofs that are independent of God's direct concurrence. The two definitions of substance are of this order (hence they can be introduced in the *Meditations* *before* God's existence is proved). The union of mind

---

<sup>42</sup> In the letter to Elizabeth in which Descartes describes the union of mind and body as a 'primitive notion', sensations also appear to be thus defined. See Letter to Princess Elizabeth, 28<sup>th</sup> June 1643 (AT III: 690/1; CSMK III: 226). However, here 'primitive notion' is an epistemic and not a metaphysical principle. We know we have sensations immediately and without inference, and in this sense our knowledge of them (and the union they are produced by) is a primitive notion. Thus, although Descartes believes we have an innate capacity to understand sensations, he does not believe that sensations are metaphysically primitive. That he reserves for general material extension and minds.

<sup>43</sup> If animals existed without sensations then Descartes would have a problem distinguishing animals from plants. However, Descartes does not have to consider this possible because God has not lent his concurrence to such a natural order (I owe this point to Jonathan Lowe). Equally it is not entirely clear how Descartes distinguishes plants from mere bodies. He seems obliged to do so by order or mechanical complexity. Clearly, though, artefacts can approach (and maybe) even surpass the complexity of simple plants. It is disappointing that Descartes did not further consider these matters, and surprising given his interest in automata and other complex machines.

<sup>44</sup> Hence for the human body to exist without a soul is possible but would require a miracle. See the Letter to Regius, December 1641 (CSMKIII: 200)

and body, though, and with it sensations, are not in Descartes' opinion capable of this independent proof. Therefore it seems highly unlikely that Descartes would accept sensations as the defining attribute for a third substance.

As Cottingham notes Descartes at all times explicitly espouses dualism. Trialism makes most sense if Descartes is a stuff dualist. It is also clear that Descartes is not thoroughly enough committed to a stuff dualism to make adopting Cottingham's trialism acceptable.<sup>45</sup> Descartes always speaks of a union between mind and body and typifies this union as close, intimate, natural and complete.<sup>46</sup> This makes sense for a union of two things, but less so for a union of three things. If Cottingham is right human beings consist of thinking substance, material substance and sensing substance. If Descartes is a stuff dualist some sense could be made of these substances being mixed together. However, if they are things, then it is not clear how three things can be united in the intimate manner that two things can be. Each substance if in complete union with one cannot be in union with the third.<sup>47</sup> Sensing substance might take the role of the vegetative soul or animal spirits and be the mode by which soul and body are meant to connect. However, it is quite clear this merely generates another quite unnecessary level of explanation that Descartes neither needs nor desires.<sup>48</sup>

---

<sup>45</sup> Indeed, part of the problem with Cartesian Trialism as Cottingham has it is that it blurs the virtues of stuff and thing dualisms. A stuff dualism does produce a useful trialism, and this Descartes might be claimed to have held in some form. Cottingham, however, takes Descartes explicit claim that sensations (in humans at least) *emerge* from the union of mind and body to generate a third *stuff*, sensing stuff. This then gives a nine-fold series of possible stuff combinations. Descartes does seem to be ultimately a thing dualist. It may be the case that Descartes should have been a thing trialist. However, by favouring a stuff ontology Cottingham produces not a trialism, but a nonalism!

<sup>46</sup> Not that Descartes is particularly consistent in this matter. In Meditation VI Descartes talks of "the union and, as it were, intermingling of the mind with body." (AT VII: 81; CSM II: 56). However, when discussing matters with Mersenne Descartes argues that "[S]ince our soul is not double, but single and indivisible, it seems to me that the part of the body to which it is most immediately joined should also be single and not divided into a pair of similar parts. I cannot find such a part in the whole brain except this gland [the conarium or pineal gland]." Letter to Mersenne, 30 July 1640 (CSMK III: 149). Insisting on a single location seems in stark contradiction with the 'intermingling' position of Meditation VI. The same inconsistency affects Descartes analogy between the union of mind and body and gravity (*Fifth Set of Objections* (AT III: 693; CSMK III: 227-8); for discussion see Mattern 1978). However, Descartes seems to prefer dispersed metaphors, suggesting intermingling is his preferred interpretation. The phrase "most immediately joined" suggests that Descartes does not regard the pineal gland as either the only point of connection, or if it is the only join between soul and body, that this doesn't exclude the possibility of 'intermingling' or some other sort of close and intimate mixing.

<sup>47</sup> The Trinity is, of course, a union of three things. However, even if this union is comprehensible, it is not clear how it might offer any analogy for the union of substances in a human being. It would seem likely that claiming so probably constitutes some sort of heresy.

<sup>48</sup> "There is only one *soul* in human beings, the *rational soul*; for no actions can be reckoned human unless they depend on reason. The *vegetative power* and the *power of moving the body*, which are called the *vegetative* and *sensory souls* in plants and animals, exist also in human beings; but in the case of human beings they should not be called *souls*, because they are not the first principle of their actions, and they belong to a totally different genus from the *rational soul*." Letter to Regius, May 1641 (CSMK III: 182). Although Descartes does use the notion of animal spirits in the usual sense of an animating force in living bodies, he never understands them as being anything other than physical.

Further there are exegetical problems concerning Descartes' treatment of animals.

Officially Descartes claims that animals do not in the strict sense think. However, he does allow that animals do have sensations.<sup>49</sup> Animals can therefore sense pain but cannot *judge* themselves to be in pain. If such intellection is to distinguish humans from animals, Descartes must maintain that humans have a thinking component animals do not have. As a consequence while animals are unions of two substances (extended substance and sensing substance) humans will be unions of *three* substances (extended, sensing and thinking substances). Trialism's attraction is that it *removes* the complications of unions between substances, particularly in the case of sentient beings. Instead it only adds the complication of adding five new possible substance combinations to the furniture of the world, two of which (sensing but unextended and sensing and thinking but unextended) have no putative examples.

Baker & Morris<sup>50</sup> outline a different interpretation of Descartes that preserves the need to have every predication terminate in one substance, and a strict dualist account. Under this interpretation every sensory predicate is systematically ambiguous. Thus to say one has a pain is in fact always bivocal. In one sense to say one is in pain is to make a claim about the state of one's body. This claim is fallible, as in the amputee's claim to feel a pain in their absent limb.<sup>51</sup> This part of pain is ascribed solely to the body. The second sense of a pain is the mental episode of judging that one has a pain. This second claim, that one seems to have a pain, is infallible. Although one cannot be wrong in judging oneself to have a pain this does not necessarily imply the pain has the normal or claimed cause. Amputees may judge themselves to have pain in a phantom limb. The claim to be in pain is infallible despite the fact that the normal mechanical cause of physical attachment to the limb is absent. This remains the case despite the fact that the amputee might verbally assert that the pain is in the missing limb (if, for example, at that point in time the amputee is unaware of the loss of the limb).

This account is obviously not without difficulty. Certainly it does not seem to have a place for qualia. Judging one has a pain is not itself painful. And surely one cannot make a pain

---

<sup>49</sup> Much is made of Descartes apparently cruel attitude towards non-human animals. Descartes does maintain that animals do not think, however, this does not mean that he believes they do not suffer pain. Non-human animals do sense pain, but they do not ever *judge* themselves to be in pain. It is this distinction that separates humans from other animals. That Descartes regards pain in non-human animals as purely mechanical and a result of the disposition of their organs and the flow of animal spirits does not necessarily demean the pain non-human animals suffer. Pain is a primarily a bodily experience, and is painful whether or not contemplated. That non-human animals cannot contemplate their pain does not make it any less painful.

<sup>50</sup> Baker & Morris (1997)

<sup>51</sup> Meditation VI (AT VII: 77; CSM II: 53)

painless by refusing to judge that one has a pain.<sup>32</sup> This suggests that pains are genuinely mental events, in as much as we experience them, but that they are not *conscious* events. This seems to force Descartes into a dilemma. Either he refuses to admit we have pains, which is surely to be rejected. Otherwise he admits that pains are singular episodes which belong to the mind and body equally, instead of there being two singular but related events that belong one to the mind and one to the body.

However, it is not clear that this criticism can be levelled at Descartes. The range of events that Descartes regards as purely mental is far more restricted than the range currently assumed by philosophers. For Descartes pains are confused thoughts that the mind has as a consequence of the actions of the inner senses.<sup>33</sup> The inner senses and sensations arise because of the union of the mind to a body. The natural union of mind and body means that as long as the body is working properly the unclear thought of a pain will always arise in the mind. Inevitably the thought will be unclear as a consequence of the limited and finite nature of our minds.<sup>34</sup> However, the union is best served and preserved if such constant conjunctions operate. Therefore to be aware of a pain as a pain in one's foot is to make a confused thought a clear thought (or at least a clearer thought). The choice then is not between pains being conscious and painful or unconscious and painless. All pains originate as physical events that the mind is made aware of via the inner senses, at first in a confused manner, and then possibly in a clear manner (though never distinct). All pains therefore have a mental correlate that may be confused or clear.

As to whether it is the foot that is painful or the thought of the pain Descartes can answer quite clearly. It is the foot that is painful, the thought merely informs the mind that this is so. So while it is true that all pains are related confusedly to the mind via our nature as union of mind and body, it is also true that it is the body that is in pain. Descartes' claim that the link between certain physical events and certain mental events is arbitrary is not as preposterous as it might seem. God has ordained that the 'universal machine' of the body will best function if certain links are made. However, there is nothing inherent about pain

---

<sup>32</sup> The exertions of fakirs offer only a partial counter to this. We are all familiar with the possibility of *ignoring* pain. However, ignoring pain, even when elevated to an art form does not amount to the *dissolution* of pain.

<sup>33</sup> The notion of Inner Senses is, of course, itself a difficult one. However, these difficulties will not be pursued here.

<sup>34</sup> There is some conflict here with Descartes' claim that angels would know their bodies perfectly, given that angels are surely also finite minds. Presumably it is not finiteness that limits knowledge, but being in union with a body. Angels would instead *observe* a body. However, this somewhat undermines the value of the comparison. If observing a body is comparable to being in union with a body it is unclear why God did not create humans along these lines. If, however, there is not comparison, then it is not clear what is gained by describing an angel as 'having' a body as a consequence of observing it. This does not approximate even

that makes it painful, of makes pleasure pleasant. It is enough for Descartes that these states succeed in preserving the union of mind and body to each other's mutual benefit. To ask further questions is nonsensical. Either there are no answers, or it is to second guess the mind of God.<sup>55</sup>

---

vaguely to what it is like to sense a body. Surely it is sensing a body that really denotes its 'possession' by a finite mind.

<sup>55</sup> "God could have made the Nature of man such that this particular motion in the brain indicated something else to the mind; it might, for example, have made the mind aware of the actual motion occurring in the brain, or in the foot, or in any of the intermediate regions . . . But there is nothing else which would have been so conducive to the continued well-being of the body." (AT VII: 88; CSM II: 60-1)

## V

### DESCARTES' SUBSTANTIAL PRINCIPLES

Hopefully the discussion above shows that sensations do not create a complication for Descartes' official dualism. Nonetheless this strict thing dualism surely creates greater difficulties for Descartes, for the union of mind and body is now surely either a fiction or as contrary and mysterious as the traditional problems of Cartesian Interactionism suggest.

To bridge this gap Descartes claims a substance can be complete or incomplete

insofar as it is referred to some other substance in conjunction with which it forms something which is a unity in its own right . . . the mind and the body are incomplete substances when they are referred to a human being which together they make up.<sup>56</sup>

This statement is normally taken to be highly paradoxical.<sup>57</sup> Descartes claimed that by a substance we understand something that needs no other thing for its existence. Although this independence claim has to undergo some qualification (i.e. that all substances are dependent on God) it seems that a substance cannot be both independent and incomplete.

The following passage from *The Principles* states the general Cartesian conception of substance:

By substance we can understand nothing other than a thing which exists in such a way as to depend on no other things for its existence. And there is only one substance which can be understood to depend on no other thing whatsoever, namely God. In the case of all other substances, we perceive that they can exist only with the help of God's concurrence. Hence, the term 'substance' does not apply *univocally*, as they say in the Schools, to God and to other things; that is, there is no distinctly intelligible meaning of the term which is common to God and his creatures . . . {In the case of created things, some are of such a nature that they cannot exist without other things, while some need only the concurrence of God in order to exist. We make this distinction by calling the latter 'substances' and the former 'qualities' or 'attributes' of those substances. }

---

<sup>56</sup> *Fourth Set of Replies* (AT VII: 222; CSM II: 157)

<sup>57</sup> In, for example, Markie (1994)

But as for corporeal substance and mind (or created thinking substance) these can be understood to fall under this common concept: things that need only the concurrence of God in order to exist.<sup>58</sup>

The Independence rule of *The Principles* is a weakened form of qualification for substancehood. Writing about Scholastic debates in a letter to Regius<sup>59</sup> Descartes says that “It is inconceivable that a substance should come into existence without being created *de novo* by God”.<sup>60</sup> This would seem to exclude *all* material objects from being substances. The soul, on the other hand, “is thought to be immediately created by God”.<sup>61</sup> These two principles, as they stand, would utterly undermine any claim human beings have to be substances themselves. Even if human beings are not substances, but substantial unions, these principles would undermine the union being understood as Descartes intends. Although the soul is obviously still an individual substance, being created by God, the body is no longer an individual substance. A human being is, instead, a union between a soul (a substance) and an amount of matter (a non-substance). This would seem to be a surrender to Scholasticism by making the soul a substance governing a non-substantial body. Descartes could be seen to make this very capitulation when he writes, “This is confirmed by the example of the soul, which is the true substantial form of man.”<sup>62</sup>

The effects of this passage can be ameliorated if it is read not as an adoption of Scholasticism, but an attack on it. Descartes wishes to attack the Scholastic doctrine that all things have substantial form. Indeed, he wishes to illustrate the notion that form cannot isolate any true substance. The strictness of the principle of divine genesis excludes all material objects from being substances. If divine genesis is the mark of substancehood only a material object that has existed perpetually since the first moment God created the universe could qualify as a substance. Presumably there is no such material object.<sup>63</sup> This favours Descartes’ version of substance where a prime attribute rather than form defines each substance. Extension is the only necessary and perpetual quality of the material. Therefore, when Descartes says that man’s proper substantial form is the soul he should be understood to be claiming that when Aristotelians speak of the substantial form of man they mean no more than the soul as he, Descartes, defines it. Descartes is attempting a *reductio*

---

<sup>58</sup> *Principles*, pt. I, arts. 51-2 (AT VIII A: 24; CSM I: 210, section in curved brackets added in French edition AT IX B: 47)

<sup>59</sup> Letter to Regius, January 1642 (AT III: 503-9; CSMK III: 205-209)

<sup>60</sup> Letter to Regius, January 1642 (AT III: 505; CSMK III: 208)

<sup>61</sup> Letter to Regius, January 1642 (AT III: 505; CSMK III: 208)

<sup>62</sup> Letter to Regius, January 1642 (AT III: 505; CSMK III: 208)

<sup>63</sup> Other than the mereological sum of all material objects that have existed since the universe began. However, Descartes does not consider this possibility.

*ad absurdum*. He claims “All the arguments to prove substantial forms could be applied to the form of a clock, which nobody says is a substantial form.”<sup>64</sup> The result of the Scholastic approach is a difficulty in saying what is unique about man, as having substantial form will not differentiate man from machine. This problem is solved, again, by the adoption of a Cartesian analysis of substance. What is unique about man is the union of extended matter with a mind. The interrelating of mind to body as form to matter Descartes demonstrates cannot work. Instead each must be independently conceived of as substances in their own right.

Both the Principle of Divine Genesis and the Independence Principle seemingly exclude animal bodies from being substances. In a Letter to Mesland<sup>65</sup> Descartes offers an account of Transubstantiation developing from the one offered in his replies to Arnauld.<sup>66</sup> The account is obscure (maybe unsurprisingly given that it explains a miracle). However, Descartes explains that we can understand that the bread becomes the body of Christ because here numerical identity consists in having the same dimensions, even if the substance is itself changed.<sup>67</sup> It is by this same principle that we can call a river the same.<sup>68</sup> Descartes claims difficulty arises because the term ‘body’ is “very ambiguous”.<sup>69</sup>

Descartes adduces the standard principle that when we designate something as body this means that removal of even the smallest part of that body would prevent it from being the same.<sup>70</sup> This is not the case with the body of a human. This remains the same “so long as it remains joined and substantially united with the same soul”.<sup>71</sup> However, Descartes is careful to distinguish this from the notion of substantial form. The body itself must retain “the dispositions required to preserve that union.”<sup>72</sup> This requirement does not demand that the body is itself a substance in the sense that would contravene the Principle of Divine Genesis. Descartes notes that our bodies never do remain numerically identical from one moment to the next given the constant movement of particles within them. Considered as a portion of space the body has no strict identity conditions. It is only individuated as a

---

<sup>64</sup> *Fifth Set of Objections* (AT III: 505; CSMK III: 208.) See also Descartes attack on the attempt to define man as a ‘rational animal’ in Meditation II (AT VII: 25-6; CSM II: 17). Descartes is always scathing of any attempt to define individual substances. Substance can only be understood as substance-in-general.

<sup>65</sup> Letter to Mesland, 9 February 1645 (AT IV: 162-75; CSMK III: 241-6)

<sup>66</sup> *Fourth Set of Replies* (AT VII: 248-56; CSM II: 173-8)

<sup>67</sup> Hence the consecration of the sacrament should not alter the shape of the bread and wine. As Descartes remarks in the *Fourth Set of Replies* (AT VII: 255; CSM II: 177) if the bread is replaced with “some flesh, or a tiny child” (things of obviously variant shape) then this is a different miracle.

<sup>68</sup> Letter to Mesland, 9 February 1645 (AT IV: 165; CSMK III: 242)

<sup>69</sup> Letter to Mesland, 9 February 1645 (AT IV: 166; CSMK III: 242)

<sup>70</sup> Letter to Mesland, 9 February 1645 (AT IV: 166; CSMK III: 243)

<sup>71</sup> Letter to Mesland, 9 February 1645 (AT IV: 166; CSMK III: 243)

<sup>72</sup> Letter to Mesland, 9 February 1645 (AT IV: 166; CSMK III: 243)

portion of space like all material objects. However “*qua* human body”<sup>73</sup> it can remain the same because the preservation of this union, being fit for the union with a soul, is enough to identify a human body.

Even though only the soul itself can be regarded as a substance, the body is not ‘informed’ by the soul. Descartes instead emphasises that certain portions of matter can be suited to union with a soul. This is possible because of the way God has predetermined matter to operate. Hence Descartes can regard a human body as being “indivisible; because if an arm or a leg of a man is amputated, we think that it is only in the first sense of ‘body’ that his body is divided.”<sup>74</sup> Descartes regards the power to preserve the union as also resting with matter and not exclusively with the informing substantial power of the soul. This is caught in his conclusion:

Altogether then, provided that a body is united with the same rational soul, we always take it as the body of the same man, whatever matter it may be and whatever quantity or shape it may have; and we count it as the whole and entire body, *provided that it needs no additional matter in order to remain joined to this soul.*<sup>75</sup>

Descartes takes himself to have demonstrated the separateness of material and spiritual substances. However, his attempt to explain how two separate substances nonetheless form a true substantial union is not complete. Descartes contrasts a substance’s being independent with its being complete (or incomplete). As noted above this is normally regarded as paradoxical. However, Descartes can respond robustly to these accusations. Firstly it is not clear that independence and completeness should be said to go hand in hand. Descartes often treats artefactual objects such as cups and saucers as substances.<sup>76</sup> Now it is clear that there is nothing about the cup that requires the saucer to exist and vice versa.

---

<sup>73</sup> Letter to Mesland, 9 February 1645 (AT IV: 167; CSMK III: 243)

<sup>74</sup> Letter to Mesland, 9 February 1645 (AT IV: 167; CSMK III: 243) Locke echoes these sentiments at *Essay* 336:33-337:9 (Book II, Chap. 27, §11) and *Essay* 341:14-33 (Book II, Chap. 27, §17).

<sup>75</sup> Letter to Mesland, 9 February 1645 (AT IV: 167; CSMK III: 243) Emphasis added. This has clear implications for change in the human body if Descartes literally means no matter can be added. However, he surely means that body must be ‘whole and entire’ in the sense of being able to function as a body without needing any additional matter. If this is the case addition (or subtraction) of matter is consistent with the body being ‘whole and entire’. Descartes cannot, of course, appeal to the notion of completeness here, because the body is by definition an incomplete substance. Hence he uses the phrase ‘whole and entire’ to avoid having to speak of the completeness of an incomplete substance. This linguistic dodge may not ultimately be entirely satisfactory if Descartes cannot explain how the bodies ‘wholeness’ differs from completeness. If there is no difference a human body is both complete and incomplete, and this is clearly absurd. Parallel difficulties arise for the mind, though conversely because the mind lacks parts and therefore is hard to conceive of as ever being incomplete.

<sup>76</sup> It is not, of course, clear that artefacts should count as substances. However, it is clear that living bodies, minds, and human beings *are* good candidates for substances. As these are the crux of the matter and the

Both are therefore independent as good Cartesian substances should be. However, it is also clear that cups and saucers belong together, and in another sense, which in no way compromises the claim to be substances, one is incomplete without the other.” Therefore Descartes can claim that “the body and the soul, in relation to the whole human being, are incomplete substances; and it follows from their being incomplete that what they constitute is an *ens per se*.”<sup>78</sup>

Further, it is not clear that completeness or incompleteness can apply to substances when assessed under *The Principles* rule. For a substance to count as a substance we have to be able to understand that a world could exist in which God had chosen only to create that one substance. In such a world only two things exist, God and the substance. Now it seems likely that *all* things being dependent on God’s concurrence are incomplete in relation to God. This makes the relation trivial. The only meaningful use of the term, therefore, is in comparisons between non-divine substances. However, the test for substancehood is to be able to exist in a world where the only other substance is God. As such the issue of completeness is excluded.

Secondly Descartes is not claiming that the distinction between minds and bodies is of the same order as distinctions between various bodies or various minds. Correctly understood a mind and body can be truly distinct under one criterion, and a true unity on another. Hence Descartes’ claim

[E]ven if we suppose that God has joined some corporeal substance to such a thinking substance so closely that they cannot be more closely conjoined, thus compounding them into a unity, they nonetheless remain distinct.”

This needs to be squared with the claim that

The body and the soul, in relation to the whole human being, are incomplete substances; and it follows from their being incomplete that what they constitute is an *ens per se*.”<sup>80</sup>

---

analogy holds between natural and non-natural kinds here fairly uncontroversially, this matter can be put aside for the moment.

<sup>77</sup> It is in this manner Descartes claims that we can have the idea of a mountain or a valley, though each is incomplete without the other, in a Letter to Gibieuf, 19 January 1642 (AT III: 477; CSMK III: 202)

<sup>78</sup> Letter to Regius, December 1641 (AT III: 460; CSMK III: 200)

<sup>79</sup> *Principles*, pt. I, art 60 (AT VIII: 29; CSM I: 213)

<sup>80</sup> Letter to Regius, December 1641 (AT III: 460; CSMK III: 200) Compare Descartes discussion of adequate and complete knowledge in *The Fourth Set of Replies* (AT VII: 220ff; CSM II: 155ff).

The independence condition for substances means that minds and bodies are substances. There could be a world that contained only a body, or only a mind. Difficulties occur in extending this to the union of a mind and body. Clearly a world could exist that only contained a union of mind and body. However, surely this union is *not* independent of its constituents, a mind and a body. Hence unions of minds and bodies fail the independence test. The parts of the unions, however, pass the test, and hence the test properly reveals the distinct nature of mind and body.

Nonetheless, we have just seen that completeness offers a second definition of substance distinct from independence. Descartes can be claimed to be making metaphysical distinctions of different orders. The independence test relies on postulating the existence of a world that God has not created. In the world God *has* created there are many things, and as such none of them are independent. However, they could be complete or incomplete. One might suggest that the two tests are to establish different things. The independence test shows why God must create two substances, minds and bodies that can be said to be really distinct. The completeness test shows how these created substances exist together. This lines up with Descartes' appeals to nature to understand the world. The independence test gives a metaphysical test for substance, the completeness test a natural test. A proper natural substance is one that is complete, even if it could not itself be independent. Thus its unity is not a fiction, despite having truly distinct parts.

It is not, however, necessary for something that has parts that its parts be practically or obviously separate. Descartes claims the concurrence of God was necessary to create a world in which we are unions of minds and bodies. Equally, only by the concurrence of God could the two actually separate or combine. However, by use of the natural light it is possible to realise one's nature is that of a union. Although one realises that one is a union by use of the mind, this no more means one is solely a mind anymore than one would believe one was just one's ears in a world which only contained audio experiences. The method of discovery surely does not determine the nature of the discoverer.

This surely is Descartes' target with his famous comment that "I am not merely present in my body as a pilot is present in a ship, but . . . am very closely joined and as it were, intermingled with it, so that I and the body form a unit."<sup>81</sup> If the mind were united to the body in a simple abutment, then the mind should perceive its body in a manner that

---

<sup>81</sup> Meditation VI (AT VII: 81; CSM II: 56). This has echoes for those writers on personal identity who claim that we are essentially our brains. For just as surely as we do not experience ourselves as pilots in ships, we do not experience ourselves as merely brains in bodies.

reflected this.<sup>82</sup> Descartes does here use 'I' to refer to the thinking part of the union rather than the union itself. However, this is only to be expected, and does not entail that 'I' must refer to a Cartesian Ego. Human beings are necessarily unions of minds and bodies. This union is not typified by the mind having perfect knowledge of the body. The mind knows itself with *immediate* distinctness. Although the body can be known distinctly this is not achieved immediately. The mind's natural light of reason illuminates outwards and therefore always reveals itself first and most distinctly. When discussing the phenomenology of being such a union it is inevitable that Descartes uses 'I' to refer to the thinking part of the union. To use 'I' to refer to the union would preclude discussing the experience of being a union of mind and body in the manner of a human being.

As quoted earlier Descartes claims the human mind must "conceive them [mind and body] as a single thing and at the same time to conceive them as two things; and this is absurd."<sup>83</sup> The intermingling between mind and body means we experience ourselves as a unit. Yet this intermingling implies we are not one substance (one substance would hardly need to intermingle with itself) but two substances. Therefore, when Descartes comes to discuss the experience of human first-person thought he finds himself at something of a loss. It would be wrong to condemn Descartes for finding the expression of this phenomenon difficult. Equally, given this systematic difficulty (arguably the systematic difficulty that underwrites Phenomenology as a distinct philosophical movement) one should avoid building interpretations on Descartes' use of 'I' in this passage. The broader picture at this pivotal moment in the *Meditations* is more revealing. These issues are further examined in the next section.

The distinction between independent and complete substances also goes some way to settling vexations about the nature of Cartesian *Res Extensa*. Considered as an independent substance there is only extension. This is a notion of substance-in-general<sup>84</sup> and excludes individual material objects. However, in the natural world there are many things. Being part of a plurality these things must fail the independence test. However, they can pass the completeness test. Hence Descartes allows that they can be treated as substances and indeed this is his practice. There is a difficulty in explaining why the natural world contains many objects rather than one object, but Descartes would most likely have regarded this as a Scholastic *sophisma* given that the world manifestly does contain many objects.

---

<sup>82</sup> This would be the case with Angels if they had bodies. See discussion later in this section.

<sup>83</sup> Letter to Princess Elizabeth, 28 June 1643 (AT III: 693; CSMK III: 227)

<sup>84</sup> What is identified as 'pure-substance' in Markie (1994) pp. 81-86.

Descartes' reliance on our *nature* as union of mind and body is bolstered by his distinction between *natural reason* and *natural belief*. We learn that our nature is one of union of mind and body only by careful direction of the natural light. Reason can, therefore, reveal the essences that underlie our nature. However, we also know our nature by way of natural belief. Natural belief is an "impulse-to-believe [and] does not directly concern itself with speculatively warranted truth, however, or furnish its own guarantee of certainty". This is opposed to "the natural light [a] dynamic inclination toward improving our grasp of the truth about reality, an inclination which may indeed be weakened and diverted by sense prejudices and the pressures of utility and convention, but which cannot be entirely quenched without destroying the self in its central act".<sup>85</sup>

In a letter to Princess Elizabeth Descartes attempts to explicate his theory by describing the union of mind and body as being a 'primitive notion'.

First I consider that there are in us certain primitive notions which are as it were patterns on the basis of which we form all our other conceptions. There are very few such notions. First, there are the most general – those of being, number, duration, etc. – which apply to everything we can conceive. Then, as regards body in particular, we have only the notion of extension, which entails the notions of shape and motion; and as regards the soul on its own, we have only the notion of thought, which includes the perceptions of the intellect and the inclinations of the will. Lastly, as regards the soul and the body together, we have only the notion of their union, on which depends our notion of the soul's power to move the body, and the body's power to act on the soul and cause its sensations and passions.<sup>86</sup>

Descartes uses primitive notions in a manner closely allied with his notion of natural order. Primitive notions are the basis by which we can come to know the natural order correctly. Descartes defines a primitive notion as that which "can be understood only through itself."<sup>87</sup> Hence "human knowledge consists solely in clearly distinguishing these notions and attaching each of them only to the things to which it pertains."<sup>88</sup> Descartes is clear that all explanation must be in terms of these notions, and that they are themselves incapable of being analysed in terms of each other. Instead it is "in our own soul that we must look for these simple notions."<sup>89</sup> Indeed, Descartes continues that it is attempting to understand these primitive notions in terms of each other that is "the main cause of our errors".<sup>90</sup>

---

<sup>85</sup> James Collins (1971) pp. 86-7

<sup>86</sup> Letter to Princess Elizabeth, 21 May 1643 (AT III: 665; CSMK III: 218)

<sup>87</sup> Letter to Princess Elizabeth, 21 May 1643 (AT III: 666; CSMK III: 218)

<sup>88</sup> Letter to Princess Elizabeth, 21 May 1643 (AT III: 666; CSMK III: 218)

<sup>89</sup> Letter to Princess Elizabeth, 21 May 1643 (AT III: 666/7; CSMK III: 219)

<sup>90</sup> Letter to Princess Elizabeth, 21 May 1643 (AT III: 666; CSMK III: 218)

## VI

### THE ROLE OF THE COGITO RECONSIDERED

The previous sections have demonstrated that the role of the union of mind and body is central to Descartes' philosophical account. Further, it has been shown that this account is reasonably robust. However, it still remains the case that the *cogito* argument of Meditation II does claim that our essence is that of a thinking thing. Clearly this needs to be looked at in some detail, if only to dispel widespread myths about Cartesian Ghosts in the Machine.

The *cogito* argument appears not only in Meditation II, but also in the *Discourse* and *The Principles*. In each case Descartes claims that thinking must confirm to each thinker that they themselves exist. However, the role of the argument is quite different in *The Meditations*. Briefly the difference is this: in *The Principles* the argument is intended as an *independent* logical and metaphysical proof. In *The Meditations* the *cogito* is presented as an experiential datum by which one can better know one's own nature, for as Descartes remarks:

many more people make the mistake of thinking that the soul is not really distinct from the body than make the mistake of admitting their distinction and denying their substantial union, and in order to refute those who believe souls to be mortal it is more important to teach the distinctness of parts in a human being than to teach their union.<sup>91</sup>

The effect of the different presentations of the *cogito* indicates that the Meditation II version cannot stand alone from its place in *The Meditations* as a whole.

Recent writings, particularly those of Amelie Rorty<sup>92</sup>, have stressed and demonstrated that Descartes fully intended *The Meditations* to be taken seriously as meditative exercises. (It would otherwise be puzzling why Descartes chose to put so much effort into presenting an argument in this form, when he represents most of the arguments in more conventional form elsewhere). As such *The Meditations* fall into a long tradition of philosophical writings. Thus the *cogito* has to be seen as arriving in the context of Meditation I's doubts, and as a first step to the conclusions of Meditation VI.

Notably Descartes names all the things he will aim to re-establish in his Meditation I attack. *The Meditations* are meant to leave the reader where he started, but with a better

---

<sup>91</sup> Letter to Regius, January 1642 (AT III: 508; CSMK III:209)

understanding of that starting point. Thus just as Descartes ends with the claim that we are a union of a mind and a body, so he begins by assuming that we are human beings. This is assumed to be known by the senses, but Descartes wishes to show that it is best known, and only truly known, by reason of the natural light.

It is important to note the title Descartes gives Meditation II. Although Meditation II distinguishes between the mind and the body it does so *only on the grounds that the mind is more distinctly known than the body*. The *cogito* is not meant to spring the thinker onto a blank ontological canvas. Instead Descartes is removing items from our pre-meditative beliefs. The *cogito* arrives at the point when Descartes discovers that there is one thing that cannot be removed: his own thinking being.

The *cogito* is contrasted with our pre-meditative belief about our nature as human beings. That was known only by the senses, and as such is to be discredited: we must come to know this through the natural light. Hence immediately after the *cogito* passage Descartes returns himself to old thoughts “What then did I formerly think I was? A man. But what is a man?”<sup>92</sup>

At this point Descartes also makes clear that although he can know his own mind and guarantee his existence through thought alone it would not then be possible to discover his nature through consideration of abstract thoughts and concepts. To this end he dismisses the Scholastics who might attempt this:

But what is a man? Shall I say ‘a rational animal’? No; for then I should have to inquire what an animal is, what rationality is, and in this way one question would lead me down the slope to other harder ones, and I do not now have the time to waste on subtleties of this kind. Instead I propose to concentrate on what came into my thoughts spontaneously and quite naturally whenever I used to consider what I was.<sup>94</sup>

It might be thought that Descartes is rejecting the possibility of an animalist or biological explanation. What Descartes rejects, though, is that we can understand ourselves by providing a definition of what we are.<sup>93</sup> Merely providing a definition would not provide the enlightenment for which Descartes is searching. Descartes is interested in how we acquire self-knowledge. This does not preclude the possibility that the mode of self-knowledge does not itself directly reveal the precise constitution of a human being. Instead,

---

<sup>92</sup> See Rorty (1986b) and also Kosman (1986)

<sup>93</sup> Meditation II (AT VII: 25; CSM II: 17)

<sup>94</sup> Meditation II (AT VII: 25-6; CSM II: 17)

that must be achieved by meditating on the nature of substances and their unions. The entire structure of *The Meditations* reveals this. It is the mind we can know immediately, the body only relatively imperfectly via the senses. That we acquire knowledge in this manner is a consequence of being a mind and body in union. If we knew our bodies perfectly, as an angel might, we would not be in union with a body, merely an observer of one.<sup>95</sup> However, while angels are endowed with the capacity to perfectly perceive everything before them, we are not. It is this unevenness of human self-knowledge and the gap between what we experience ourselves to be, and what we actually are, that fascinates Descartes and drives *The Meditations*.

Immediately following this Descartes re-examines the pre-meditative claim that he is a man, and continues

Well, the first thought came to mind was that I had a face, hands, arms and the whole mechanical structure of limbs which can be seen in a corpse, and which I called the body.<sup>97</sup>

Even in the passage that establishes the *cogito* Descartes does not neglect the intuition that we are embodied creatures. This embodiment, however, is to be investigated by the new science. The body itself should be examined empirically. It is the aim of *The Meditations* as a whole is to show how correct direction of reason can give us knowledge of our nature as a union of mind and body. In this the *cogito* is only a stepping stone, and not the central claim.

---

<sup>95</sup> What I have defined as the Neo-Lockean approach. See Introduction to this thesis.

<sup>96</sup> For Descartes opinion on angels see his Letter to More, August 1649 (CSMK III: 380) and also the Letter to Regius, January 1642 (CSMK III: 206)

<sup>97</sup> Meditation II (AT VII: 26; CSM II: 17)

## VII

### SOULS, HUMAN BEINGS & PERSONS

There are two facts about the human soul on which depend all the knowledge we can have of its nature. The first is that it thinks, the second is that, being united to the body, it can act and be acted upon along with it.<sup>98</sup>

Descartes wishes the mind to take the role of the soul in a metaphysical system that will 'please the theologians'.<sup>99</sup> Yet his rational inclination was to provide an account of the soul that was accessible to human knowledge. If the mind is an indivisible substance we can have sure knowledge of, then it is a good candidate both for being the soul, and being easily known.

The range of thinking ascribed to the mind is restricted to those modes that might be classed as having a moral element: "A thing that doubts, understands, affirms, denies, is willing, is unwilling".<sup>100</sup> This is no accident. The soul is the bearer of guilt and sin that supposedly makes it possible to judge each individual on judgement day. Surely then, Descartes' avowed belief about personal identity should be that we are essentially, in part at least, a mind, and as this part is immortal, surely it is the soul that makes the identity of a person. Certainly the 'Cartesian Ego' is taken by many writers on personal identity to be the genesis of claims that identity of a person must lie in some unchanging, simple, thing. Yet it does not seem obvious Descartes' position has to add up to this.

This passage in Grosholz's *Cartesian Method and the Problem of Reduction*<sup>101</sup> captures why Descartes would not want to identify the person with a 'Cartesian Ego'.

The shadow of the real, substantial distinction between *res cogitans* and *res extensa* falls on the Cartesian self and threatens it with a disruption that the abstract unity of spirit may not redeem. What kind of self could experience its memory and desire, to say nothing of its perception and imagining, as alien and as merely contingent? For such a self, history and moral experience, its own projects, gratifications, regrets, and creaturely habits would play no role in the constitution of its true self. Even the pilgrim souls in Augustine and Dante,

---

<sup>98</sup> Letter to Princess Elizabeth, 21 May 1643 (AT III: 664-5; CSMK III: 217-218)

<sup>99</sup> It seems to me right to take Descartes comments on Religion as being sincere. Although he was a cautious man who disliked trouble, and at times it is clear he is not beyond a certain amount of dissembling to avoid controversy, the commonly made claim that Descartes *only* pleased the Theologians to stay out of trouble is too extreme a claim. In a freer age his faith might have found different expression, but it would still have been an expression of faith. Nor does it seem likely that Descartes is a crypto-materialist.

<sup>100</sup> Meditation II (AT VII: 28; CSM II: 19)

<sup>101</sup> Grosholz (1991) p. 142

merely passing through on their way to transcendence, are not so thoroughly stripped of their earthly particularity.

Although Descartes attaches the will to the mind, he is explicit in making memory a corporeal as well as intellectual process.<sup>102</sup> If Descartes did mean 'Cartesian Egos' to be persons, then he meant them to be memory-less and unimaginative persons. This metaphysically desiccated entity does not fulfil the role of person, and it is strange to suggest that Descartes might believe that it did.

One must consider that Descartes rarely uses the term person, and never uses it in a technical or philosophically significant sense. It is assumed, therefore, that personal identity must then be foisted onto either mind or body. Yet, could it not be the case that neither is equivalent to the notion of person that modern writers want? Instead Descartes employs the Compositional Account which does all the work he needs to solve what later writers would call the problem of personal identity. In the Sixth Set of Replies he clearly defines his notion of a human being:

In fact I have never seen or perceived that human bodies think; all I have seen is that there are human beings, who possess thought and a body. This happens as a result of a thinking thing's being combined with a corporeal thing: I perceived this from the fact that when I examined a thinking thing on its own, I discovered nothing in it which belonged to body, and similarly when I considered corporeal nature on its own I discovered no thought in it.<sup>103</sup>

Descartes uses this framework: during life, identity goes with the human being. Each human is a moral agent, because they have a mind that is a centre of conscious, which is to say culpable, behaviour. Moral acts are therefore willed and acted by humans, although it is their minds that make their acts moral. The same acts carried out by mindless brutes would not be blameworthy because they have no minds, and cannot judge right from wrong.

The mind, being indivisible, does not decay on the death of the body, and as such is immortal and survives. However, the mind although a substance is an *incomplete* substance. It can act as a store for the sins of an individual human, and God can re-unite this soul with a body so as to generate a new human being that would be *morally* identical with the previous human being. It is not clear, though, that they would be the same person. However, this issue simply does not arise for the Cartesian. Moral acts can be carried out

---

<sup>102</sup> Although latterly he does toy with the idea of a purely intellectual imagination.

<sup>103</sup> *Sixth Set of Replies* (AT VII: 444; CSM II: 299)

only by humans, and it is only these acts which can be judged. In as much as persons are moral actors, then for Descartes personal identity goes with the union of a mind and body.

This has the unexpected effect of making Descartes a kind of animalist. Descartes' belief in the immortality of the soul does not detract from this. The soul only makes the same person in the event of God's intervention to recreate actual human beings for the Day of Judgement. In the normal run of nature, however, God has determined that humans are identical from birth to death, and are identical as unions of minds and bodies. Descartes' faith in the natural order precludes the possibility of the problem cases that fixate theorists after Locke.<sup>104</sup> A single mind will always be united to one and the same human body for the natural life of that person. Descartes' claim for everyday identity would be answered by the question "Is this the same human animal?" and surely if the only thing that can gainsay this is divine intervention this is as strong a theory as any (as any account has divine counter-examples).

Descartes' account of human beings may not ultimately be coherent or defensible. However, for the purposes of this thesis this does not matter. What is clear is that Descartes took an uncomplicated view of the world, in which human animals are a natural and metaphysically well-behaved part. Descartes is not, then, the founder of a school of thought about personal identity that presumes that personhood is constituted by psychological continuity (whether guaranteed by a metaphysical simple or otherwise). The reverse is in fact true. Further, not only does Descartes focus his philosophy on human beings, he effectively excludes the traditional problems of personal identity. The fact of identity is a natural one, our knowledge of this identity innate. The traditional problems of personal identity are born with Empirical Philosophy. However, the Empirical Philosophy only makes sense in light of the prevalence of the Compositional Account.

---

<sup>104</sup> Descartes *does* face the set of problems that beset any claim that the universe is created with a certain number of souls.

## CHAPTER 2

### LOCKE ON PERSONAL IDENTITY

Noonan remarks that all philosophical enquiry concerning personal identity is a footnote to Locke.<sup>105</sup> Locke's discussion is the first explicit debate of personal identity as an independent and significant problem in its own right. As such the study of Locke's theory of personal identity remains as pertinent as ever.

The previous chapter has shown how Descartes adhered to a Compositional Account of personal identity. Locke's philosophy at large is a critique of Cartesianism and this is just as true for his writings on personal identity. Writers who take Locke to be the founder of the philosophical problem of personal identity are obliged to treat Locke as if he wrote his theory onto a philosophical *tabula rasa*. Instead, the following weaves Locke back into the fabric of philosophical history. This makes it clear which problems Locke felt himself to be responding to, and makes his philosophical motivations clearer.<sup>106</sup> Apart from a reaction against Cartesian doctrine Locke was also interested in challenging the intolerance and dogma of contemporary theologians. Locke's account of personal identity is also an account of moral and legal responsibility. It is constructed not only in opposition to prevailing metaphysics but also to prevailing theological disputations.<sup>107</sup> A tactic Locke uses throughout the *Essay* is one of under-cutting the value of traditional disputations. Where Locke sees an argument that tempts either scepticism or dogma he prefers to show that the debate is irrelevant to its putative subject. Thus he does not have to settle interminable issues and can provide an account that is free from the dogma and intolerance Locke spent his life opposing.<sup>108</sup> This tactic is in evidence throughout Locke's comments on personal identity.

---

<sup>105</sup> Noonan 1989, p. 30. For a similar verdict see also Thiel 1998b, p. 868

<sup>106</sup> It would also explain why Molyneux urged Locke to add a chapter about identity, and why Locke agreed to this (See Letters 1609 & 1620 in De Beer 1979). The problems were live ones Locke could not afford to ignore in a work claiming to be as comprehensive as the *Essay* does.

<sup>107</sup> It will not be possible to discuss here the theological implications of Locke's *Essay*. Harris notes that by placing moral worth with the individual consciousness sin is, in consequence, specific to each individual. Hence the sin of the Fall does not rest with each and everyone of us. See Harris (1994) pp. 235-6 & 301-305. For wider discussion see Tennant (1982).

<sup>108</sup> A good example of this is at *Essay* 438:28-30, 439:1-15 (Book III, Chap VI, §1). Here Locke states that we cannot know if there is only one sun, our Sun, and must allow that there might be as many suns as there are stars (as he notes some claim). As our sun is just another star, Locke's caution here has paid off! All references are to the Nidditch edited edition of the *Essay*, Locke (1975). References take the form of a page and line numbers marking the beginning and ends of the quoted passage (e.g. *Essay* 438:12-15) followed by Locke's own textual divisions (e.g. Book I, Chapter X, §15).

This chapter does not aim to defend a Lockean or Neo-Lockean account of personal identity. Locke is normally taken to present a psychological criterion of personal identity. This is naturally antagonistic to the animalist position. However, just as one can ponder if Descartes is a Cartesian one can equally wonder if Locke is actually a Lockean, at least concerning personal identity. Locke can be interpreted in a manner that shows he was highly sympathetic to the Compositional Account. When Locke's motives are properly understood this can be seen clearly. In this context it is possible to see that Locke could hold an animalist theory with very little adaptation to his explicit theory. Reading Locke in this way means there is no need to labour over the circularity objection, made famous by Bishop Butler, that supposedly dogs Locke's theory. Although Locke's memory theory has an intuitive appeal that has made it undeniably popular, and the circularity objection has likewise been as rigorously pursued (most notably giving us the contemporary debate concerning those mysterious things, Q-memories), this is not what makes Locke's writings of lasting importance and relevance.<sup>109</sup>

---

<sup>109</sup> Quasi-memory was introduced by Shoemaker (see Shoemaker 1959 & 1963) and made much use of by Parfit (see Parfit 1984). For interesting discussion see Wiggins (1980). All these writers believe themselves to be writing in a Lockean spirit. Schechtman provides arguments against q-memory which I regard as devastating. See Schechtman (1994a) and (1994b).

## I

### ARE LOCKEAN PERSONS SUBSTANCES?

Locke is taken to define persons as being distinct from souls and distinct from human beings. The person (or self) is equated with *consciousness*. The person (or self) is therefore equated with a psychological continuity, most typically of memories, and this continuity assumed to be causal. Locke's theory is therefore taken to imply that persons are non-substantial psychological entities. This foists an untenable position on Locke by way of a selective reading of the text and is at best a caricature of Locke's intentions in the *Essay*. Writers who want to adopt a Lockean theory for positive reasons assume Locke wished to do the same. This ignores the negative aspect of Locke's theory, and the manoeuvres that lead Locke to his final position.

Most writers accept that Locke explicitly denies that persons are substances.<sup>110</sup> What is strange is that many progress to treat persons as quasi-substantial entities, against both their own and Locke's claims. Also, having maintained that persons are not substantial, there is a general failure to place persons in a category that Locke would accept, i.e. a mode or a relation, as Locke takes this tripartite division to be exhaustive.<sup>111</sup> This interpretative malaise arises because of a lack of sympathy with Locke's aims. It is true that Locke sees personal identity as a specific problem that cannot be settled by the traditional apparatus of soul and body. However, it is not clear that Locke sees personal identity as the problem modern philosophers do. Locke did regard identity itself as a metaphysical principle worthy of explanation and investigation. However, the problem of personal identity does not represent for Locke a peculiar metaphysical problem. Personal identity is a problem of moral ascription, given our restricted state of knowledge about the identity of substances. This amounts to the claim that Locke sees *identity* as a *metaphysical* problem, but *personal identity* as an *empirical* problem.

The rejection of a substantial basis for personal identity is normally taken to commit Locke to a psychological theory of personal identity. The most famous and often propounded version of this is to interpret Locke as holding a memory theory of personal identity.<sup>112</sup> Yet Locke has not only denied that persons are substances he has also denied that persons are identical with any one spiritual substance. There is a confusion here, for Locke also defines

---

<sup>110</sup> For a contrary view see Alston & Bennett (1988) p. 25ff

<sup>111</sup> *Essay* 164:29-35 (Book II, Chap XII, §3)

a person (in part, at least) as “a thinking intelligent Being”.<sup>113</sup> It would seem to be obvious that a soul does constitute ‘a thinking intelligent Being’ and so qualifies as a person.<sup>114</sup> Yet Locke ignores this consequence. Otherwise if a person did have more than one soul in the course of a life, each soul would be a person in its own right. Yet Locke does not regard persons as possible constructs of other smaller person-parts. Locke regards the notion of personal as maximal, although he does not state this explicitly anywhere (it is fair to assume he took the point to be too obvious to need stating).

If Locke does not regard individual substances as persons (unless the substance and the maximal person coincide) Locke cannot define persons as collections of ‘personal substances’. This would suggest that the person would instead have to be either a mode or a relation. Locke asserts that all modes and relations are “ultimately terminated in Substances”.<sup>115</sup> If persons are modes or relations, then they could not exist without substances. If persons are not substances, but necessarily depend on substances they must be either collections of modes or a relation, or even a relation of relations or a relation of modes.

Locke does not think persons are relations. It would have been simpler if he had, for as he remarks “The *Ideas* then of *Relations* are capable at least of being more perfect and distinct in our Minds, than those of *Substances*.”<sup>116</sup> However, Locke does not once claim persons are relations in either of the two chapters on Relation or in the chapter on personal identity.<sup>117</sup> If Locke does not regard persons as substances or relations then he might regard them as modes. Yet he does not explicitly claim this either. To some extent Locke did think of persons as collections of modes. Certainly the extension of consciousness implies a collection of thoughts. Thoughts are modes of substances, and Locke even famously allows that thought, although mostly likely a mode of the soul, God might “superadd”<sup>118</sup> to body. Hence Locke could see a series of thoughts being supported by any series of substances, whether solely spiritual, solely material, or most likely a combination of the two. In this Locke accepts a Compositional Account of Man

It might be objected that to talk of a collection of modes is to talk of a way modes are *related*. Certainly the Neo-Humeans, and many Neo-Lockeans, believe that providing a

---

<sup>112</sup> See discussion in §V of this chapter.

<sup>113</sup> *Essay* 335:10-11 (Book II, Chap XXVII, §9)

<sup>114</sup> Although, as is argued in §III below the role of body may well preclude this.

<sup>115</sup> *Essay* 329:25-26 (Book II, Chap XXVII, §2)

<sup>116</sup> *Essay* 322 (Book II, Chap XXV, §8)

<sup>117</sup> *Essay* 319-348 (Book II, Chaps XXV-XXVII)

criterion of personal identity is the task of explaining what *relation* unites a person over time. Consciousness is the name Locke gives to the union of thoughts that make up a person “[f]or since consciousness always accompanies thinking... in this alone consists *personal Identity*”.<sup>119</sup> Locke never calls consciousness a relation. The reason for this can be found in Locke’s definition of relation; “there can be no Relation, but betwixt two Things, considered as two Things, either in themselves really separate, or considered as distinct, and then a ground or occasion for their comparison.”<sup>120</sup> Relation can only occur if two things are really distinct *or* we regard one thing in two ways such as to afford a comparison. Nothing truly identical, for example a person, could be united by relation, for this would mean that its temporally separated parts are really distinct, which is contrary to its identity. Of course, Locke allows that we can compare the same thing at two different times and suggest a relation between these observations. However, that is a different matter.<sup>121</sup> This is how Locke explains our knowledge of the identity of animals and vegetables through causal changes (cause and effect being a relation). We observe a relation between their parts and changes in size over time, and this reveals the ‘operation’ that constitutes animals’ and vegetables’ identical lives.<sup>122</sup>

Locke does call modes things on occasion. Consciousness could then be seen as genuinely a relation amongst distinct things. This is the very road Hume is to go down.<sup>123</sup> However, Locke upholds an account of identity that makes it very unlikely that he would hold such an account of personal identity. Consciousness is introduced because Locke needs a way to unite modes into persons that neither makes persons quasi-substantial entities nor is a relation. Locke emphasises the similarities between an animal *life* and a person’s *consciousness*. This comparison needs to be taken at face value. Locke conceived consciousness to be directly comparable to an animal life, and describes neither in causal terms. An animal life arises as a consequence of a certain aggregation of matter. Similarly a person is an aggregation of certain states in a thinking being. Locke is more interested in the fact of this aggregation than its cause. Further, as just discussed, Locke uses the notion of animal and vegetable identity to *ground* observations of cause and effect.

It would appear therefore that Locke is not, as regards personal identity at least, a Lockean at all. He does not regard persons as either modes or relations. This is clear enough from

---

<sup>118</sup> *Essay* 541:4 (Book IV, Chap III, §6)

<sup>119</sup> *Essay* 335 (Book II, Chap XXVII, §9)

<sup>120</sup> *Essay* 321 (Book II, Chap XXV, §6)

<sup>121</sup> *Essay* 328:4-6 (Book II, Chap XXVI, §1)

<sup>122</sup> *Essay* 324-5 (Book II, Chap XXVI, §1-2)

<sup>123</sup> See next chapter §§II, & IV-VI

the above discussion. Locke also doesn't think persons are substances. All he means to exclude by this is a pure Cartesian notion of substance, that is a mind or portion of matter identical in its own terms. To claim this Locke would have to know what sort of substance persons are. The very point of Locke's discussion is that the true nature of substances cannot be known well enough to equate any substance with a person. This, combined with Locke's desire to provide an account of personal identity where persons can be perfectly known, untainted by metaphysical vagary, obviously precludes claiming that persons are substances.

What then does Locke think persons are? He thinks they are quasi-substantial *compounds*, that is they contain substance, but cannot be known certainly to be pure substances. He makes this clear in his conclusions

[W]hatever Substance begins to exist, it must, during its Existence, necessarily be the same: Whatever Compositions of Substances begin to exist, during the union of those Substances, the concrete must be the same: Whatsoever Mode begins to exist, during its Existence, it is the same: And so if the Composition be of distinct Substances, and different Modes, the same Rule holds.<sup>124</sup>

Locke's account of personal identity is one of how these compound existences are held together. Having established the sort of relation Locke thinks persons have to substances his general theory of substance can be more closely evaluated. There is one more point that needs emphasising before continuing. All writers are clear that Locke distinguishes between an animal and the mass of matter that makes it up at any one time. Yet there is less consensus as to what constitutes the similar reciprocal term for Person. Person could be contrasted with substance, matter or spirit. None of these is obviously satisfactory. However, there is a further candidate that has been overlooked. Traditionally Locke is also taken to use Self and Person as synonyms. In fact Locke used these two terms to express the relation between a person and what constitutes that person at any one time. The argument for this claim will be made later, but for the moment the reader should be warned to be vigilant for this possibility.

---

<sup>124</sup> *Essay* 347:33-35, 348:1-4 (Book II, Chap XXVII, §28)

## II

### LOCKE ON SUBSTANCE & IDENTITY

Locke's account of substance is a complex one. This complexity arises because Locke combines apparently Cartesian and Aristotelian themes with his own empiricism despite the lack of easy compatibility and his opposition to both philosophies.<sup>125</sup> Locke adopts the notion that things have a 'shared organisation of parts' and that this can provide a criterion of identity. Yet this identity is not *substantial* identity as it would be for an Aristotelian. Identity of substance concerns only aggregates of Cartesian substances, i.e. thinking substance and physical matter. Into this inherent conflict is born Locke's account of personal identity. The problem, therefore, is to determine what led Locke to adopt such a mixed account of identity criteria and what mistakes it leads him into.

Locke outlines a strikingly Cartesian list of possible substances "We have the ideas of but three sorts of substances: 1. God. 2. Finite Intelligences. 3. Bodies."<sup>126</sup> Locke, like Descartes, regards the identity of God as being indubitable, and gives only two sentences to the issue.<sup>127</sup> Locke does not seem aware of the difficulties the rest of his account creates for him. To hold a corpuscular view in combination with there being three sorts of substance commits Locke to one of the following. First, there might be two types of corpuscular matter, spiritual and material. Or second, a substratum below all other substances forms both spiritual and material bodies, a sort of neutral monism. Locke is arguably sympathetic to both. He does seem to talk of substance as a substratum<sup>128</sup> and appears to talk of the possibility of particles of matter themselves gaining and losing parts.<sup>129</sup> He also talks as if

---

<sup>125</sup> For an insightful history of Locke's rejection of Aristotelian and Scholastic doctrines see McCann (1987).

<sup>126</sup> *Essay* 329:1-2 (Book II, Chap XXVII, §2)

<sup>127</sup> *Essay* 329:2-4 (Book II, Chap XXVII, §2)

<sup>128</sup> Alexander (1980) and (1981)

<sup>129</sup> *Essay* 329:8-9 (Book II, Chap XXVII, §2) If we take Locke to equate particle and atom here this result is strange. Locke could mean by particle a mass of matter, in which case his choice of words is strange, given the technical uses he puts particle and mass to elsewhere. If he does distinguish, so that particle should mean atom then he is committed to a substratum view, so that atoms can have different compositions. This would run against the obvious understanding and use Locke puts atoms to. Locke may mean that a particle of matter remains the same as long as it is not subsumed into a larger mass. However, it is not clear how an atom could lose its identity in this way, if it is a true simple. Again this can only be made sense of if there is a substratum and atoms can contain different amounts of matter. Mostly likely Locke uses 'particle of matter' to mean 'part of matter'. This Cartesian usage best captures Locke's intentions. However, it is not an atomistic use. So this clashes with Locke's later claim about the jumbling of atoms not changing the identity of a mass in §3 though this passage is ambiguous. Locke's phrase is "let the parts be never so differently jumbled". Locke is often taken to mean that arrangement is not important to a masses identity, but the 'never' could be read as meaning the opposite of this. This returns to the original difficulty. If Locke here doesn't mean atom when he writes part then maybe the jumbling is of the gross parts of a body, not its constituent atoms. Clearly there is either a delicacy in Locke's usage here that has not been widely commented on, or he is inconsistent in his use of language in these passages. McCann notes some of this difficulty (McCann 1987, p. 61) and comments that the

souls are constituted of parts in a manner analogous to material bodies, and therefore seemingly of spiritual as well as material atoms. Here, as throughout his account of substance, Locke is caught between the Cartesian attack on Scholasticism, to which he is sympathetic, and the more radical corpuscular mechanics of his scientific fellows. Locke's general position, though, is a pragmatic one. Matter is the substratum we presume to explain material objects, spirit the substratum we presume to explain thought.<sup>130</sup> However, all this aside, the real issue is to provide criteria of identity for finite intelligences and bodies.

The identity of both finite intelligences and material particles consists in the fact that each has a "determinate time and place of beginning to exist, the relation to that time and place will always determine to each of them Its identity, as long as it exists."<sup>131</sup> This solution is typical of Locke. It allows him to leave questions about the nature of substance untouched and still provide a robust theoretical framework. Locke believes that anything that can be said to have a beginning (although not necessarily an end) can have its identity fixed by that beginning in time. This principle can be extended to concretes and compounds which remain identical as long as they remain so composed.<sup>132</sup>

This has two implications for personal identity. First, persons can be accounted for without a full account of substance. Secondly, although having a beginning means an identical entity exists, it does not imply that informative non-circular criteria can be given for that object's constitution.

---

difficulty is unimportant to Locke's wider claim that coherent spatio-temporal bodies are identical. This fits well with the Lockean tactic of side-stepping more complex issues by emphasising a level of importance untouched by underlying debates. The final resolution of Locke's meaning in this passage can be left to others. It is possible that Locke wants to dress his theory in the colours of contemporary science but doesn't entirely manage. This is a not uncommon fate for philosophers.

<sup>130</sup> *Essay* 297:24-36, 298:1-10 (Book II, Chap XXIII, §5)

<sup>131</sup> *Essay* 329:5-7 (Book II, Chap XXVII, §2) If two things are formed by the splitting of one thing, they might both reasonably be said to have the same beginning. Amoebas do not provide such an example, because their splitting is clearly spatially if not temporally distinct. That is to say, it is clear which part of the previous whole has become each of the two new independent wholes. However, at a sub-atomic level two smaller atoms might be created from a larger atom in such a way that there was no question of spatial distinction prior to the split. The parts were mixed and until the point of the split could have belonged to *either* of the new atoms. This contrasts with the amoeba case where there is a clear *pre-figuring* of the division. Sub-atomic particles do not seem to exhibit this sort of ordered division. Until the division an electron could end up in *either* new atom. Although this may not be an accurate representation of chemical physics, there is a philosophical problem that arises when one considers the possibility of separating mixtures. Of course, this depends on how strong one feels the identity conditions of mixtures are, and whether or not they represent co-incident objects. A mixture of two types of powder can be regarded as either one mass, or two coinciding masses. If they are one mass and the two powders could be separated instantaneously the same problem would emerge. However, there is little reason to regard mixtures or masses of powders as anything other than aggregates, which surely have very weak identity conditions. For an alternative view, taking mixtures and aggregates to have more metaphysical weight see Simons (1987) §6.2, pp. 218-21 and §6 in general.

<sup>132</sup> *Essay* 347:33-348:4 (Book II, Chap XXVII, §28) quoted at the end of §I, above.

Nonetheless Locke also states that physical matter consists of atoms. In his atomism Locke follows the corpuscular mechanism of his day and assumes that atoms are finite and indivisible. Material substantial identity for Locke is derived from the identity of simples. The identity of thinking substance Locke says nothing about, feeling it is even more obscure in its nature than matter. As for modes, as each mode must terminate in a substance it will be the same as long as it persists in one substance. The exception to this is modes that are necessarily momentary. These, perishing as they begin, are necessarily diverse. Locke brings our attention to “things whose Existence is in succession, such as are the Actions of finite Beings, v.g. Motion and Thought, both which consist in a continued train of Succession.”<sup>133</sup> As memories are presumably examples of thoughts, the clear implication is that each memory must belong to some substance. This will be significant later. Locke’s statement that thoughts are necessarily diverse and exist as successions is under discussed in the literature. It has obvious ramifications for the kind of psychological account one can construct from Locke. It is also a fascinating pre-figuring of Hume’s advances on Locke’s theory.<sup>134</sup> Further it impinges on Locke’s views on whether the soul can be said to always think. This is discussed below in §III.

As discussed in the previous chapter Descartes’ account of individuation is somewhat jejune. Descartes believed that living bodies are machines. Yet material substance should minimally be described in terms of extension in space. The pure Cartesian project of material description in terms of extension is deeply inadequate for individuating objects. Broadly Descartes’ difficulty arises because mechanic description supplants the Scholastic notion of substance but cannot offer a substitute principle of individuation. This paucity of explanation combined with a desire to avoid a lapse into scholasticism may have been among Molyneux’s motivations in urging Locke that the *Essay* needed an account of the *Principium Individuationis* to be complete in its account of “metaphysick and logick”.<sup>135</sup>

Locke’s adoption of simples as paradigmatically identical leads him in an obvious direction. Aggregates and compound substances should be regarded as identical as long as they change none of their parts “the Mass, consisting of the same Atoms, must be the same

---

<sup>133</sup> *Essay* 329:28-30 (Book II, Chap XXVII, §2)

<sup>134</sup> Hume believes identity to be a relation *between* unity and diversity. If, like Locke, one assumes that thoughts are necessarily diverse, and, like Locke, one wishes to preserve the notion of the person being the same, one would seem to inevitably have to arrive at an account of mental identity akin to Hume’s. This will be fully discussed in the next chapter, esp. §§II & V.

<sup>135</sup> See Molyneux’s letter to Locke, 2nd March 1693 p. 650, Letter 1609, Vol. IV in Locke (1979)

Mass, or the same Body, let the parts be never so differently jumbled.”<sup>136</sup> However, Locke goes on for “‘Tis not ... Unity of Substance that comprehends all sorts of Identity, or will determine it in every Case”.<sup>137</sup>

Substantial identity, understood as unity of particles or the unity of mind, must be supplemented for “In the state of living Creatures, their Identity depends not on a Mass of the same Particles, but on something else.”<sup>138</sup> This ‘something else’ in plants is “such an Organisation of Parts in one coherent Body, partaking of one Common Life”.<sup>139</sup> Locke says further that “[t]he Case is not so different in Brutes”.<sup>140</sup> Indeed Locke sees no difference between the identity of plants and animals. In both cases identity consists in partaking in the same life. The section ‘Identity of Animals’<sup>141</sup> only adds a distinction between artefacts and living organisms. In a nod towards the mechanistic analysis of the Cartesians Locke likens the organisation of parts in an animal’s body to those of a watch. The only difference is that “in an Animal the fitness of the Organisation, and the Motion wherein Life consists, begin together, the Motion coming from within; but in Machines the force, coming sensibly from without, is often away, when the Organ is in order, and well fitted to receive it.”<sup>142</sup> Here Locke echoes the principle that Descartes had appealed to in his account of substance, that of independence. An animal could be conceived of separately, its motion coming from within. This is not true of an artificial machine, which would need a maker to provide motion from without.<sup>143</sup>

Locke is clear that neither animal nor vegetable identity is substantial identity. Yet this is not the entire picture. Locke suggests that there are only three sorts or kinds of substance;

---

<sup>136</sup> *Essay* 330:16-18 (Book II, Chap XXVII, §3). As discussed above in note 131 previously Locke’s exact meaning is not clear here. Ayers has made much of Locke’s failure to realise that a mass of matter can endure as long as its parts are changed in a regulated fashion (Ayers 1981). In defending Locke against this charge this passage is crucial. Clearly Locke allows that an animal or vegetable life is one type of regulated change. However, if he means by ‘never so jumbled’ a limited change, he appears to accept Ayers’ point. If he means any arrangement of the same simples is the same mass, then Ayers’ objection has weight. See note 131 above.

<sup>137</sup> *Essay* 332:23-24 (Book II, Chap XXVII, §7)

<sup>138</sup> *Essay* 330:20-22 (Book II, Chap XXVII, §3)

<sup>139</sup> *Essay* 331:4-5 (Book II, Chap XXVII, §4) Locke’s insistence on plants being ‘one coherent body’ might be too strong a condition, as some plant species arguably remain as coherent bodies while in fact containing different individual plants.

<sup>140</sup> *Essay* 331:19 (Book II, Chap XXVII, §5)

<sup>141</sup> *Essay* 331:19-33 (Book II, Chap XXVII, §5)

<sup>142</sup> *Essay* 331:29-33 (Book II, Chap XXVII, §5) Locke’s use of watches as an example of a substance is clearly based on an informed knowledge of the working of timepieces. See *Essay* 463:14-36, 464:1-22 (Book III, Chap VI, §39). Locke’s insistence on the watches parts being in fit order to receive motion from without tells against those who believe Locke believes in intermittent objects. If the watch was taken apart it would presumably not be in fit order, and would have gone out of existence. For a view that interprets Locke as holding that objects can have intermittent existence see Hoffman (1980)

<sup>143</sup> Locke’s account meets problems when trying to separate individual animals in a species. Each individual is born of a parent, and was at some point a *part* of that parent. The motion that the parent passes to a newly

finite intelligences, bodies and God. These are 'pure' substantial kinds. This does not preclude the possibility of either sub-sorts of substances or compound or mixed substances. If this is the case, objects like animals and vegetables can also be substances. Locke does make reference to the union of substances, commenting "during the union of those Substances, the concrete must be the same".<sup>144</sup> This decidedly Cartesian distinction between substances and their compounds ('the concrete') puts Locke in a similar position to Descartes. Animals and Vegetables both are and are not substances. Naturally, Locke does not develop as careful a defence of this position as Descartes does, because Locke has an entirely different apparatus for explaining knowledge of substance and their possible unions. Nonetheless it would be useful to settle whether Locke treats animals and vegetables as substances, or mere concretes, or, if the matter is obscure, whether it is safe to put it aside.

McCann persuasively demonstrates that Locke wished to avoid introducing substantial form in an explanatory role.<sup>145</sup> Scholastic doctrine<sup>146</sup> suggests that substantial form is not just the organisation of matter, but a real part of the thing it constitutes. Locke, in line with Boyle, wishes to claim that a plant is no more than the matter that is organised therein. Hence Locke's reluctance to call individual concretes substances. To do so would seem to concede that some further entity was present. Locke instead adheres to the Cartesian attribute version of substance. Noonan, therefore, is wrong to state that "something like Aristotle's *substantial form* holds a prominent place in [Locke's] thought, at least with respect to living creatures."<sup>147</sup> Although Locke is looking for a suitable *replacement* what he is looking for is quite unlike the Aristotelian or Scholastic conception. Of course, Locke might be mistaken in assuming that a difference of position is possible. A Lockean life unites parcels of matter, and Locke does not rise to the complexities such a position throws up. However, Locke is clear that "animal Identity is preserved in Identity of Life, and not of Substance."<sup>148</sup> The tension in Locke's position goes back to the very origins of the *Essay*. In Draft A he comments that it would be nonsensical "that a fortuitous concourse of atoms unguided by an understanding should frequently constitute the bodies of any species of

---

conceived embryo is essential to its coming into being, but this point when this stops being 'from without' and can be said to be 'from within' cannot always be easily pinpointed in a non-arbitrary fashion.

<sup>144</sup> *Essay* 347:35, 348:1 (Book II, Chap XXVII, §28)

<sup>145</sup> McCann (1987) pp. 55-57.

<sup>146</sup> Obvious different writers conceived the problem differently, but broadly speaking there is an emergent shared doctrine.

<sup>147</sup> Noonan (1978a) p. 344

<sup>148</sup> *Essay* 337:17-18 (Book II, Chap XVII, §12) Locke concedes that the term life can be prone to ambiguity at *Essay* 503:27-35 (Book III, Chap X, §22)

animals".<sup>149</sup> Here, though, Locke finds himself appealing to something beyond material constitution to justify mere masses not being animals. Locke is engaged in a struggle to avoid appealing to Aristotelian Form, while avoiding the problems of outright materialism. This is a difficult mid-ground to hold, and Locke's notion of Being may not succeed. Indeed the frequent empiricist reliance on existence to ground individuation is reminiscent of Scholastic haecceity.<sup>150</sup> Nonetheless, the broad aim and its impact on Locke's thinking can be appreciated without having to settle details.

For the above reasons some commentators<sup>151</sup> have suggested that Locke employs a thesis of relative identity.<sup>152</sup> Locke allows that something could be the *same animal* but a *different body*. Noonan insists that "what is noteworthy is the total absence in Locke of inferences made by way of the principle that if all Fs are Gs then the same F is always the same G".<sup>153</sup> Locke's continued treatment of both vegetables and animals as *things* promotes this view. Locke claims identity consists in sameness of life. This has to be made consistent with Locke's scepticism about substance. Although we can observe that the sameness of an oak does consist in its ability to take up new particles and organise them, this process cannot be completely known. Thus experience can only teach us that sameness of substance and sameness of vegetable or animal life are distinct. Locke does not believe that detailed knowledge about identity of life is available. Hence he cannot close the gap between material substance and lives. Locke's doubts about possible knowledge of substances leave him caught between two stools. On the one hand he cannot except an Aristotelian notion of substance as he holds we cannot know enough about natural kinds. On the other hand he cannot accept the Cartesian Rationalism that would appeal to the laws of mechanics to complete the explicatory power of material substance. Yet, experience clearly does teach us that material bodies are distinct in their identity from vegetable and animal lives. Locke, therefore, ends accepting *both* as simultaneously valid accounts of identity. The price of this is that (probably despite Locke's intentions) he does not class plants and animals as substantial things. They are, nonetheless, Beings, in the sense Locke gives Being to mean 'an existence'. Each has a beginning, and this is all that a being needs to be an identical existence.

---

<sup>149</sup> Draft A, §42 p. 74

<sup>150</sup> See Thiel (1998a)

<sup>151</sup> See Mackie (1976) and Noonan (1978a), Griffin (1977), Geach (1968). See Chappell (1989) for a survey of the debate.

<sup>152</sup> The modern source of this debate is to be found in Geach (1962) and (1968). Sympathetic arguments can be found in Quine (1948) and (1958b). For extensive (and convincing) argument against relative identity see Lowe (1989c).

<sup>153</sup> Noonan (1978a) pg. 346

However, Locke's insistence that masses of matter and Lockean lives are distinct does provide an answer to those who claim Locke is a relativist as regards identity. If something is an *F* and a *G* and the *F* and the *G* constitute different things at all times, then there is no question of relative identity, only of coincidence. Locke does maintain the thesis of sortal dependency, indeed the term sortal is Locke's. However, sortal dependency is not relative identity.<sup>154</sup> Locke's main thesis about identity is that existence itself accounts for identity, and each existence is individuated by its origin. Therefore a succession of things can be coincident with another kind of thing, but this is not identity. Masses of matter are unstable for Locke, incapable of surviving the loss or addition of any amount of matter. Masses are defined in almost complete opposition to Lockean lives, which are defined as things that *can* survive the addition of amounts of matter. Even for the time a mass of matter remains unchanged while in coincidence with an animal life it is not identical with that animal, because the two things have different histories, one having come into existence after the other. The only time the mass and the animal might be thought to be identical is when the animal first comes into existence. This, however, is only illusory.

Animals are not generated from nothing. Although the process is unclear, at some point a mass of matter becomes organised such as to constitute an animal. The animal and the mass appear to not only coincide but be identical if a change to the mass brings the animal into existence. Then, briefly, the mass' and the animal's constituent parts have the same histories. However, only the animal can survive the addition or subtraction of parts. Even at this initial moment of genesis this serves to immediately distinguish the animal from the mass. Should the animal come into existence with no addition or subtraction to the mass, then even this is not true, because the mass exists prior to the animal.<sup>155</sup>

---

<sup>154</sup> Sortal Dependency is the claim that an identity statement is incomplete without a sortal term. That is to assert ' $a=b$ ' is meaningless unless one can say  $a$  is the same *F* as  $b$ . Lowe (1989) claims Sortal Dependency is essential while defending a non-relative account of identity. For arguments about the worth of Sortal Dependency see Ayers (1974) and Alston & Bennett (1984).

<sup>155</sup> This defence of Locke is a development of that given in Chappell (1989). The possibility of one moment of coincidence collapsing into identity might be regarded by some as fatal to a coincidence account of objects. As I do not wish to defend such an account, and believe it can be shown to be undesirable without developing such an argument, I leave its development to others. If one wishes to deny coincidence there are other difficulties about how animals come into existence which have to be faced. In Draft A Locke writes that "the two Ideas are not exactly one & the same but in the understanding as distinct as the Ideas of one & two white & black or as of body & man" §27, p. 46. Locke seems to be suggesting that the distinction between man and body is only a distinction of reason. However, if he is adopting the scholastic term his other examples are curious, neither being obviously a mere distinction of reason.

Suffice to say Locke's apparent adherence to the doctrine of coinciding objects is not to be applauded.<sup>156</sup> Nonetheless, it is the position Locke adopted, and it is important to be clear that he held it in a manner incompatible with relative identity. Relative identity theorists who mistake Locke as one of their own do so for various reasons. Relative identity is associated with an attack on the coherence of the notion of objects as the most natural and primitive parts of any ontology. As such it is the last blow in a campaign started by Locke against the notion of substance. But Locke believed in substance, and he believed in objects. Hence Locke had no need for relative identity. Locke did adopt sortal dependency. However, Locke did this to explain the gap between our empirical knowledge and the nature of things. To claim we understand things only because we have an idea of them is a very different from claiming that things only exist as they do because of our ideas of them. Relative identity is a doctrine for the philosophically extreme, and Locke, curious as his philosophy might be at times, is not one of these extremists.<sup>157</sup>

It is normally assumed Locke hoped by settling the issues of vegetable and animal identity he could easily settle the question of Man's identity as merely a special case of animal identity. However, Locke's notion of Man is not nearly as straightforward as he intended. In the Second Edition of the *Essay* Locke discusses the definition of Man, introducing the famous example of the rational parrot. The place of this argument is discussed in §IV. However, Locke's philosophical journey to this mature position is a long, fascinating and revealing one. This is the subject of §III.

---

<sup>156</sup> For contemporary arguments in favour of coincidence see Lowe (1983a) & (1995). For replies, responses and alternatives see Burke (1992), (1994a), (1997a) & (1997b) as well as Doepke (1982), Simons (1985) and (1986), Noonan (1985) and finally Zimmerman (1995) & (1997).

<sup>157</sup> Locke's use of compound substances may mean he does *not* hold a theory of coincident objects.

### III

#### LOCKE'S NOTION OF MAN<sup>158</sup>

Having discussed the notion of identity (§1-3), and then vegetables (§4) and brutes (§5), Locke proceeds to discuss "wherein the Identity of the same *Man* consists; viz. in nothing but a participation of the same continued Life, by constantly fleeting Particles of Matter, in succession vitally united to the same organised Body."<sup>159</sup> Locke then makes a strong case for placing identity of the Man solely in the identity of the Animal for

He that shall place the *Identity* of Man in any thing else, but like that of other Animals in one fitly organised Body taken in any one instant, and from thence continued under one Organisation of Life in several successively fleeting Particles of Matter, united to it, will find it hard, to make an *Embryo*, one of Years, mad, and sober, the same Man, by any Supposition, that will not make it possible for *Seth, Ismael, Socrates, Pilate, St. Austin, and Cæsar Borgia* to be the same Man. For if the *Identity* of Soul alone makes the same Man, and there be nothing in the Nature of Matter, why the same individual Spirit may not be united to different Bodies, it will be possible, that those Men, living in distant Ages, and of different Tempers, may have been the same Man[.]<sup>160</sup>

Locke has two competing accounts of Man in his sights. One is the Aristotelian notion of Man that defines a man as 'a rational animal'.<sup>161</sup> The second is the Cartesian account that defines Man as a unit of mind and body, a fusion of thinking and extended substance, that is the Compositional Account.

Locke's queries with the notion of Man date back to the earliest draft of the *Essay*. Locke notes that a child might "call a Negro a devil rather than a man & at the same time call a dryl a man."<sup>162</sup> He explains:

---

<sup>158</sup> Unfortunate as Locke's choice of the word 'Man' is I have followed his usage to avoid confusion in exegesis and also because it is not possible to substitute any phrase without settling the very issue under investigation (obviously substituting 'person' would be disastrous, and 'human being' is confused by Locke's technical use of being. 'Human animal' might be acceptable, but this would prejudice which aspects of humans Locke thought were animal).

<sup>159</sup> *Essay* 331:34-5, 332:1-2 (Book II, Chap XXVII, §6)

<sup>160</sup> *Essay* 332:2-14 (Book II, Chap XXVII, §6) There being 'nothing in the Nature of Matter' preventing union to a different soul runs against Descartes' appeal to the body being the best fit for a soul, if this means a best fit to the individual soul the body is in union with.

<sup>161</sup> Locke attributes the definition *animal rationale* to Aristotle and the lengthier *animal implume bipes latis unguibus* to Plato.

<sup>162</sup> Draft A, §2 p. 9. 'Dryl' or 'Drill' is a term current in the 17<sup>th</sup> Century for a Baboon or any larger African monkey, a use maintained in the less obsolete term Mandrill. This example is repeated at §27, p. 48. Other relevant passages in Draft A not discussed here are: Man used as an example at §3, p. 12; Rationality is only part of the idea of man "not the knowledg of things existing in rerum natura" §13, p. 27/8; "body, sense & motion" alternative definition of man §29, p. 55; some men more different to each other than some men to beasts §41, p. 69.

for even that species which we may suppose to know best and call man is not so readily distinguished from beast or to speak more properly it is but very uncertainly or significantly yet determined what those number of qualities are which together are signified by the word man, whereof when any of them are wanting the thing wherein they are wanting is not to be called man.<sup>163</sup>

In this early version Locke does attempt to list the qualities that go together to make the idea of Man:

when instructed by the observation of my senses I have framed the Idea of a man i.e. have put together these following Ideas v.g. 1° A face usual to that species which though consisting of many particular Ideas for brevity's sake & the thing being obvious I take for one. 2° Two hands with five fingers on each. 3° Two legs. 4° Upright posture. 5° Living. 6° A power of laughing. 7° A power of speaking. 8° Of reasoning i.e. knowing the consequence of words or propositions one to another. 9° of judging i.e. guessing at the truth of words or being of things.<sup>164</sup>

Clearly this list has its problems. 1 is blatantly circular and 2 & 3 clearly not essential. 4 is still important in defining the species of *Homo Sapiens* but crawling men are still men. 6 is derived from the same school of Greek and scholastic definitions as 'rational animal' and 'featherless bipeds' and just as bad. 7 is favoured by the Cartesians. As already noted by the time of the *Essay* proper Locke has rounded on such definitions. Only 5, 8 and 9 need to be taken seriously. If Locke is not yet ready to condemn the former group he is aware of the difficulties that await. He continues:

I say when I have put all these 9 Ideas together & thereof in my mind framed the Idea which I call a man I cannot thereby certainly know that where any 8 of those 9 really exist together that there the 9<sup>th</sup> is necessarily also (at least if we except the 5<sup>th</sup> Life, which is not the specific Idea whereby I know man, or distinguish him from other things.) For when I say a man is rational or hath the power of reasoning, it hath one of these two meanings 1° either that the power of reasoning belongs to or is included in that Idea which I have framed & call man, & then the proposition is certainly true & is only verbal & reaches no farther than my own Ideas & words or names I apply to them 2° or that where the 7 foregoing Ideas meet together there this 8<sup>th</sup> [and 9<sup>th</sup>?] is certainly also, & then the proposition is real i.e. of things existing really without my mind but is not certainly true...<sup>165</sup>

---

<sup>163</sup> Draft A, §2 p. 9/10

<sup>164</sup> Draft A, §13, p. 28 Locke seems confused over his numbering and despite various crossings out and corrections does not seem to arrive at a consistent scheme. I have done my best to iron out the numbering consistent with the intended meaning.

<sup>165</sup> Draft A, §13, p. 28/9

Although Locke cites the various qualities uncritically, the problem of the shape of man guaranteeing reason is already present. Locke notes later that the shape of man does not imply the existence of a soul<sup>166</sup> and largely repeats the arguments above while becoming more critical. Locke urges that some “rational men” he knows “may demonstrat, that infants, changelings & maniacs are noe men”.<sup>167</sup> Some might allow that a dumb creature is a man, others insist on the power of speech. Some may even neglect all aspects of shape and allow “quadrupes” or animals or any shape as long as they “found speech and reason joynd”.<sup>168</sup> He concludes that all this confusion should be put to one side in favour of what is sure about substance. Definition “is noe very material question to our present purpose... [for] of such a complex Idea it is certain that **What is is.**”<sup>169</sup> Locke prefers to assert that the existence of individual substances is certain, however impoverished our definitions might turn out to be. This is the emergence of the role Being is to take in the mature *Essay*.

In Book III, Chapter XI, §16 ‘Morality Capable of Demonstration’<sup>170</sup> Locke considers the inadequacy of definitions of Man in discussions of morality.

Upon this ground it is, that I am bold to think, that *Morality is capable of Demonstration*, as well as Mathematicks: Since the precise real Essence of the Things moral Words stand for, may be perfectly known; and so the Congruity, or Incongruity of the Things themselves, be certainly discovered, in which consists perfect Knowledge. Nor let any one object, that the names of Substances are often to be made use of in Morality as well as those of Modes, from which will arise Obscurity. For as to Substances, when concerned in moral Discourses, their divers Natures are not so much inquired into as supposed; v.g. when we say that *Man is subject to Law*: We mean nothing by *Man* but a corporeal rational Creature: What the real Essence or other Qualities of that Creature are in this Case no way considered. And, therefore, whether a Child or Changeling be a *Man* in a physical Sense, may amongst the Naturalists be as disputable as it will, it concerns not at all the *moral Man*, as I may call him, which is this immovable, unchangeable *Idea, a corporeal rational Being*. For were there a Monkey or any other Creature to be found that has the use of Reason to such a degree, as to be able to understand general Signs, and to deduce consequences about general *Ideas*, he would no doubt be subject to Law, and, in that Sense, be a *Man*, how much soever he differ’d in Shape from others of that Name. The Names of Substances, if they be used in them as they should, can no more disturb Moral, than they do Mathematical Discourses: Where, if the Mathematician speaks of a *Cube* or *Globe of Gold*, or of any

---

<sup>166</sup> Draft A, §27, p. 48

<sup>167</sup> Draft A, §27, p. 49

<sup>168</sup> Draft A, §27, p. 49

<sup>169</sup> Draft A, §27, p. 49

<sup>170</sup> *Essay* 516:21-517:11

other Body, he has his clear settled *Idea*, which varies not, though it may, by mistake, be applied to a particular Body to which it belongs not.<sup>171</sup>

In this passage Locke tries to pull apart the same notions that he explores in the Second Edition discussion of the rational parrot. Locke insists that in some discourses not all aspects of a substance are relevant, we need not always know its 'real essence'. Hence the definition given to a Man varies with our concern. Locke notes that physical deformity or immaturity is claimed by some to exclude children and changelings from being men.<sup>172</sup> This, however, 'concerns not the moral man', which is 'this immovable, unchangeable idea' of 'a corporeal rational being'. That the moral man is a forerunner of Locke's notion of person is made clear by what follows. The Moral Man is defined by way of qualities exhibited. Hence anything, even a monkey, which exhibited these qualities, would 'in that sense be a man'. The qualities are those of reason. The moral man, however, does not need to be able to "consider it self as it self, the same thinking thing in different times and places"<sup>173</sup>

At this stage Locke has made a number of connections and distinctions that typify the account of persons, but some have yet to appear. The notion of a Moral Man is still tied to the notion of a Man, and to the notion of a 'corporeal thinking being'. In the second edition this has become "thinking intelligent Being".<sup>174</sup> Further Locke's claim that a monkey would be 'in that sense a man' is entirely dismissed when Locke writes in the Rational Parrot section "whoever should hear a *Cat* or a *Parrot* discourse, reason, and philosophize, would call or think it *nothing but a Cat or Parrot*".<sup>175</sup> This breaks the link with the Aristotelian rational animal that the Moral Man had. Locke does not exclude being corporeal, but does not insist upon it.

Locke has yet to make the connection between the Moral Man and the moral agent. The connection to agency and moral consciousness is crucial in Locke's introduction of 'Consciousness' and 'Person' as terms of art. At this stage moral agency is still connected with being a Man and seen as a defining aspect of humanity. Locke has to move to seeing agency as key to moral action. Agency connects to knowing one's own actions. This

---

<sup>171</sup> *Essay* 516:21-36, 517:1-11 (Book III, Chap XI, §16)

<sup>172</sup> Locke sees the failure to make an embryo and a mature man one and the same as a definite weakness, given his account of man. See *Essay* 332:7 (Book II, Chap XXVII, §6)

<sup>173</sup> *Essay* 335:11-13 (Book II, Chap XXVII, §9) For example the qualities of a unicorn can be imagined (a mixed mode) but never repeatedly experienced like the qualities of a horse (a nominal essence based on a real essence).

<sup>174</sup> *Essay* 335:10-11 (Book II, Chap XXVII, §9)

<sup>175</sup> *Essay* 333:9-10 (Book II, Chap XXVII, §8)

prompts Locke to add the self-knowledge found in Consciousness to his requirements for being a Moral Man. It is when this final connection is broken with Man that Locke starts to talk of personal identity instead.

Locke's analogy with mathematics is intriguing. In the early *Essays On The Law of Nature* Locke had entertained that a scientific proof of morals was possible.<sup>176</sup> This intention is fully clear in the opening part of §16 quoted above. Although he abandoned the attempt to fully expound such a moral system, Locke did not abandon the idea that somehow the reward and punishment for moral acts did have a sure and certain footing. Hence the analogy with geometry. A geometrical notion is a 'settled idea, which varies not'. Similarly, Locke believes, the Moral Man is a way of conceiving of a Man in a manner 'which varies not'. If we understand what the Moral Man is, then we will properly understand how to settle issues of personal identity and moral blameworthiness and 'substances ... can no more disturb moral ... discourses'.

The entirety of Locke's concerns with personal identity are formatively captured in this passage. All that is to change in the Second Edition is that Locke runs through all the ways that substances might upset our understanding of the Moral Man, and that Locke supplements the moral aspect with a legal and forensic gloss.

Locke's reasons for believing that an account of Man as a natural kind cannot be complete are also present in much greater length in the First Edition. Indeed the classification of Man as a species of animal could without little exaggeration be said to obsess Locke in these passages.<sup>177</sup> An extensive discussion of which qualities can be ascribed to Man occurs in Book III, Chapter VI, 'Of The Names of Substances'.<sup>178</sup> There is a complementary discussion in Book IV, Chapter XI, 'Reality of Knowledge'.<sup>179</sup> As these passages are rarely discussed it is worth surveying the different issues Locke raises.

---

<sup>176</sup> Locke (1988) p. 54ff and p. 199. Compare *Essay* 549:12-26 (Book IV, Chap III, §18)

<sup>177</sup> The references to the definition of Man can be found at *Essay* 440:4-15 (Book III, Chap VI, §3), *Essay* 440:27-30 (Book III, Chap VI, §4), *Essay* 441:5-10 (Book III, Chap VI, §4), *Essay* 450:19-22 (Book III, Chap VI, §21), *Essay* 450:23-35, 451:1-16 (Book III, Chap VI, §22), *Essay* 453:13-36, 454:1-21 (Book III, Chap VI, §26), *Essay* 456:25-34 (Book III, Chap VI, §29), *Essay* 460:30-36 (Book III, Chap VI, §33), *Essay* 462:24 (Book III, Chap VI, §36), *Essay* 464:15-22 (Book III, Chap VI, §39), *Essay* 465:30-34, 466:1-6 (Book III, Chap VI, §43), *Essay* 474:24-6 (Book III, Chap VIII, §1), *Essay* 500:5-18 (Book III, Chap X, §17), *Essay* 500:27 (Book III, Chap X, §18), *Essay* 502:19-24 (Book III, Chap X, §21), *Essay* 516:31-36, 517:1-6 (Book III, Chap XI, §16), *Essay* 518:33-4, 519:1-18 (Book III, Chap XI, §20), *Essay* 569:8-35, 570:1-36, 571:1-36, 572:1-37, 573:1-8 (Book IV, Chap XI, §§13-16). This list does not claim to be comprehensive. These passages are rarely cited although Mattern (1980) cites *Essay* 516:21-517:11 (Book III, Chap XI, §16). Two other examples prepossess Locke; Gold and the internal workings of a Watch. The uses of a watch as an example may follow Boyle's use in *The Origin of Forms and Qualities* (1666, Works III, 48-9).

<sup>178</sup> *Essay* 438:28-471:13

<sup>179</sup> *Essay* 630:12-639:8

Locke's wider aim in 'Of The Names of Substances' is to define and demonstrate the difference between *nominal* and *real* essences. This is summarised by Locke's comment that "Indeed, as to the *real Essences* of Substances, we only suppose their Being, without precisely knowing what they are: But that which annexes them still to the *Species*, is the nominal Essence, of which they [the real essences] are the supposed foundation and cause."<sup>180</sup> Nominal essences are the collections of qualities competent language users assign to substances that are recognised by their language. This is possible because real essences are premised to underwrite this practice. The nature of real essences is impossible to accurately discern. However, Locke insists that qualities consistently occurring together is good proof that there are real essences. This grounding distinguishes nominal essences of substances from the more arbitrary mixed modes.<sup>181</sup> Locke concludes "[t]his then, in short, is the case: *Nature makes many particular Things, which do agree one with another, in many sensible Qualities, and probably too, in their internal frames and Constitution...*"<sup>182</sup>

Locke also outlines his concerns with the definition of Man quite clearly:

And I imagine, none of the Descriptions of the word *Man*, which we yet have, nor Descriptions of that sort of *Animal*, are so perfect and exact, as to satisfie a considerate inquisitive Person; much less to obtain a general Consent, and to be that which Men would every where stick by, in the Decision of Cases, and determining of Life and Death, Baptism or no Baptism, in Productions that might happen.<sup>183</sup>

Locke furnishes us with examples of the defective definitions as well as the issues he cites as contentious. The issue of baptism is discussed using the case of the Abbot of St. Martin as an example.<sup>184</sup> The Abbot was born deformed, and "'Twas for some time under *Deliberation, whether he should be baptised or no.*"<sup>185</sup> The exclusion was based on the Abbot's failure to conform to the "sacred Definition of *Animal Rationale*"<sup>186</sup>. Locke is scathing of the worth of two Scholastic definitions of man, *Animal Rationale* and *Animal implume bipes latis unguibus*.<sup>187</sup> He argues there is no reason "why a Visage somewhat

---

<sup>180</sup> *Essay* 442:33-7 (Book III, Chap VI, §6)

<sup>181</sup> *Essay* 455:23ff (Book III, Chap VI, §28)

<sup>182</sup> *Essay* 462:14-16 (Book III, Chap 6, §36)

<sup>183</sup> *Essay* 455:16-22 (Book III, Chap VI, §27)

<sup>184</sup> *Essay* 454:1-21 (Book III, Chap VI, §26)

<sup>185</sup> *Essay* 454:7 (Book III, Chap VI, §26)

<sup>186</sup> *Essay* 454:2 (Book III, Chap VI, §26)

<sup>187</sup> *Essay* 453:15-16 (Book III, Chap VI, §26) These definitions are also discussed at *Essay* 500:5-18 (Book III, Chap X, §17) where Locke ascribes the former to Aristotle and the latter to Plato. At *Essay* 456:25-31 (Book III, Chap VI, §29) Locke attacks the *Animal Rationale* asking if "*Balaam's Ass* has, all his Life discoursed rationally" why it would not be considered a rational animal, thus prefiguring the rational parrot argument.

longer, or a Nose flatter, or a wider Mouth could not have consisted ... with ... a Soul”<sup>188</sup>

This argument is not complete, and Locke also argues for the complementary position, that external appearance does not imply a faculty of reason. This is a point familiar from the Rational Parrot argument in ‘Of Identity and Diversity.’<sup>189</sup> Here, however, Locke adds the further point that it is only our nominal essence of Man that is being challenged. Locke asks, “Shall the difference of Hair only on the Skin, be a mark of a different internal specifick Constitution between a Changeling and a Drill, when they agree in Shape, and want of Reason, and speech?”<sup>190</sup>

In ‘Of the Name of Substances’ Locke mainly uses Man as an example of a substance that is variously and contrarily defined. The moral and epistemological implications are discussed in ‘Reality of Knowledge’. The passages here are key in demonstrating why Locke would turn his back on the Moral Man in favour of the non-substantially and supposedly unequivocal status of a person.

Locke begins by arguing that Changelings are between Man and Beast. By Beast Locke means non-human animal. A Changeling is a baby supposed by folklore to be a child exchanged for another at birth by fairies. The replacement child has the form of a human baby, but lacks rational sense.<sup>191</sup> Loosely this matches with our notion of mental handicap. Locke reserves the term ‘Monster’ for the physically handicapped. Hence he comments when comparing the two “What will your drivling, unintelligent, intractable *Changeling* be? Shall a defect in the Body make a monster; a defect in the Mind (the far more Noble, and, in the common phrase, the far more Essential Part) not?”<sup>192</sup> Locke reminds us that Man and Beast represent only nominal essences, so a further nominal essence between them supposes no contradiction *unless* one supposes that the names reflect their real essences perfectly. If this were the case no species would exist between the two, but we cannot know this to be the case our knowledge being restricted to nominal essences.<sup>193</sup> This allows Locke to continue:

*Changelings*, which is as good a word to signify something different from the signification of *MAN* or *BEAST*, as the Names *Man* and *Beast* are to have significations different one from the other. This, well considered, would resolve this matter, and shew my meaning without any more ado. But I am not

<sup>188</sup> *Essay* 454:17-19 (Book III, Chap VI, §26)

<sup>189</sup> *Essay* 334:34-5, 335:1-8 (Book II, Chap XXVII, §8)

<sup>190</sup> *Essay* 451:8-11 (Book III, Chap VI, §22).

<sup>191</sup> Hence Locke’s comment about “some *Changelings*, who have lived forty years together without any appearance of Reason” *Essay* 569:14-15 (Book IV, Chap IV, §13)

<sup>192</sup> *Essay* 571:35-6, 572:1-2 (Book IV, Chap IV, §16)

<sup>193</sup> *Essay* 569:8-29 (Book IV, Chap IV, §13)

so unacquainted with the Zeal of some Men, which enables them to spin Consequences, and to see Religion threatened whenever any one ventures to quit their Forms of Speaking, as not to foresee what Names such a Proposition as this is like to be charged with: And without doubt it will be asked, If *Changelings* are something between Man and Beast, what will become of them in the other World? To which I answer, 1. It concerns me not to know or inquire. To their own Master they stand or fall.\* It will make their state neither better nor worse, whether we determine anything or no. They are in the hands of a faithful Creator and a bountiful Father, who disposes not of his Creatures according to our narrow Thoughts or Opinions, nor distinguishes them according to Names and Species of our Contrivance. And we that know so little of this present World we are in, may, I think, content ourselves without being peremptory in defining the different states, which Creatures shall come into when they go off this stage. It may suffice us, that he hath made known to all those who are capable of Instruction, Discoursing, and Reasoning, that they shall come to an account, and receive according to what they have done in this Body."<sup>194</sup>

The tactic is typical of Locke. He condemns pedants and puritans not because of any weakness in their argument (although it is clear Locke has little time for their arguments) but by undercutting their claims. The eternal fate of every creature will be decided by divine wisdom, and it is both impossible and unnecessary to attempt to discover what this decision might be, or second-guess the criteria it will be made upon. The attack, though, has to be regarded as limited. Locke himself will attempt to offer an account of correct moral blame in the personal identity chapter. His comments about a geometrical demonstration of morals and ethics show that he accepts the possibility. His account is not based, however, on 'Names and Species of our Contrivance.' The crucial difference is that although knowledge of compounded things like substances must necessarily be limited because of their complex nature, individual simple qualities can be known more certainly. Locke's later account is based on Consciousness, which he feels can be certainly known. Locke continues to distinguish between human justice and divine justice, and is clear their ways of knowledge are distinct.<sup>195</sup> Locke continues

"The force of these men's question (*viz.* will you deprive *Changelings* of a future state?) is founded on one of these two Suppositions, which are both false. The first is, That all Things that have the outward Shape and Appearance of a Man, must necessarily be designed to an immortal future Being after this Life. Or, secondly, that whatever is of humane Birth must be so."<sup>196</sup>

---

<sup>194</sup> *Essay* 569:30-35, 570:1-20 (Book IV, Chap IV, §14) \* Cf. Rom. 14:4, \*\* Cf. 2 Cor. 5:10. It is the Bible, not Locke, that urges an account for things 'done in this Body'. I take this to be among the passages in scripture to support insistence on bodily resurrection, something Locke nods towards at *Essay* 340:4-7 (Book II, Chap XXVII, §15)

<sup>195</sup> *Essay* 343:38, 344:1 (Book II, Chap XXVII, §22)

<sup>196</sup> *Essay* 570:22-27 (Book IV, Chap IV, §15)

Locke argues that if outward shape were necessary to possible immortality, then *anything* of human form would have to be immortal. As this would apply to statues<sup>197</sup> Locke suggests the consequence is absurd.<sup>198</sup> Further Locke finds it ridiculous to “place the Excellency of a Man more in the external Shape of his Body, than internal Perfections of his Soul”.<sup>199</sup> Finally, Locke rebuffs those who claim a certain shape is indicative of a soul being within. This would need proof, and Locke finds none. The only evidence of a soul being present is degree of reason, and reason is found sometimes in monsters, yet not in changelings, despite their shapes.<sup>200</sup>

Locke then moves to the question of parentage.<sup>201</sup> Locke does make considerable usage of the biological notion of a species. However, he is also willing to consider the possibilities represented by fantastical animal compounds, such as Mermaids, and winged fishes.<sup>202</sup>

Locke’s biological knowledge is often plainly wrong, for example he claims “I once saw a Creature, that was Issue of a Cat and a Rat, and had the plain Marks of both about it”<sup>203</sup> and no such creature could exist. The same goes for Locke’s claim that “Women have conceived by Drills”.<sup>204</sup> Locke has already pointed out the stupidity of depending upon parentage to determine knowledge of a thing’s species. If we depended on such information we would have to “go to the *Indies* ... to know whether this be a Tiger or that Tea”.<sup>205</sup>

Clearly no such journey is necessary. Locke argues that even with parentage known one still arrives at borderline cases.

For since there have been humane *Fætus*’s produced, half Beast and half Man ... it is possible they may be in all the variety of approaches to the one or the other Shape, and may have several degrees of mixtures of the likeness of a Man, or a brute, I would gladly know what are the precise Lineaments, which

---

<sup>197</sup> *Essay* 571:24-25 (Book IV, Chap IV, §15)

<sup>198</sup> *Essay* 570:27-36, 571:1-7 (Book IV, Chap IV, §15)

<sup>199</sup> *Essay* 571:8-10 (Book IV, Chap IV, §15)

<sup>200</sup> *Essay* 571:17-29 (Book IV, Chap IV, §15) This rebuttal would put Locke at odds with Descartes’ claim that the human body is a vessel fit to take a soul, and ideally fitted to do so by nature. This is an important difference between Locke and Descartes in their attitudes to the Compositional Account of human beings. Locke also comments at *Essay* 108:17-20 (Book II, Chap I, §10) “But whether the Soul be supposed to exist antecedent to, or coeval with, or some time after the first Rudiments of Organisation, or the beginnings of Life in the Body, I leave to be disputed by those, who have better thought of that matter.”

<sup>201</sup> See also *Essay* 451:17-32, 452:1-8 (Book III, Chap VI, §23) for a comparable discussion.

<sup>202</sup> *Essay* 447:1-29 (Book III, Chap VI, §12)

<sup>203</sup> *Essay* 451:28-9 (Book III, Chap VI, §23) Locke may have seen an African cat, some species of which have long rat-like snouts, and little or even no fur, giving them a very rodent like appearance. Confusions of this type are not uncommon in 17<sup>th</sup> Century travel literature. In 1696, De Vlamingh, for example, notes finding on an island off Western Australia “a kind of rat as big as a common cat, whose dung is found in abundance all over the island.” The Quokkas of Western Australia are, in fact, marsupials. (The example is borrowed from Jeff William’s *Western Australia*, Lonely Planet 1998, p. 148)

<sup>204</sup> *Essay* 451:23 (Book III, VI, §23)

<sup>205</sup> *Essay* 452:6-8 (Book III, Chap VI, §23)

according to this Hypothesis, are, or are not capable of a rational Soul to be joined to them.<sup>206</sup>

Locke feels this also means that his opponent must concede that Monsters and Changelings are a species between Man and Beast. Consequently guidance cannot be taken from the delineation of species.<sup>207</sup>

These examples demonstrate the centrality the notion of Man played in the composition and development of the *Essay*.<sup>208</sup> Locke's doubts are present from the beginning. However, it is only as the *Essay* evolves that Locke realises that the attack he has placed on the notion of man has left him without an account of the Moral Man. The Moral Man passage is an interim measure, before the emergence of Locke's theory of personal identity.

The chapter 'Of Identity & Diversity' attempts to provide a deflationary account of the notion of Man. The account of Man must answer to common sense, i.e. that we have an idea of Man we use successfully, but also not grant any inroads to the doubts Locke has identified. Locke is sympathetic to the notion that Man has a moral aspect and this aspect is what we morally judge. However, he is aware of the tensions of this account, given his own arguments about knowledge of substances. Further he is aware of the zeal with which some would use definitions of man to persecute others. Like his attack on innate ideas Locke finds himself obliged to attack the presupposition to avoid the ends the arguments would be put to. Locke presents himself with a dilemma. He might repair his account of the Moral Man so that uncertainties about substance both do not undermine his own account nor provide succour for the zealots. Alternatively he can develop an account of moral responsibility that does not rely on *any* notion of man. This is the prime motivation behind Locke's account of personal identity. He wishes to preserve the instincts that lead him to

---

<sup>206</sup> *Essay* 572:22-29 (Book IV, Chap IV, §16) The use of a slippery-slope paradox to demonstrate the notion of man not to depend on shape (by suggesting a small feature change, and then another, and so on) is echoed by Hume's suggestion that if one perception does not give us the idea of a person, then no number of perceptions ascribed to an animal can give one such an idea. See Hume's discussion of the oyster at *Treatise* 634:28-34 (App. 7)

<sup>207</sup> *Essay* 573:1-8 (Book IV, Chap IV, §16)

<sup>208</sup> Curiously references to Man are almost entirely absent from Draft B with the exception of a comment at §78 that "none of the definitions or rather descriptions of that sort of animall [Man] hath donne soe perfectly & exactly as to satisfie a considerable inquisitive person." (p. 185, this is extended to become *Essay* 455:16-22 (Book III, IV, §27) in the final *Essay*), and the repetition of the example of a child mistaking a negro for a devil and a drill for a man (§84, p. 192) and casual use as an example at §92, (p. 198) and §93b (p. 202). In Draft B Locke prefers to use swans, herons and cassowaries as examples. This disappearance might be indicative of both Locke's pre-occupation and confusion with the issue of Man as a substance. The difficulties with the issue in Draft A led him to remove the discussion from the later Draft B, but the importance of the issue led him to re-introduce the discussion to the complete *Essay*. The unpublished Draft C may reveal more of this development, as might an unpublished manuscript *Identity of Persons* dating from 1683 (Bodleian Library MS Locke f. 7, p. 107) alluded to by Thiel (Thiel 1998, p. 908, note 113, and bibliography).

believe in the Moral Man, while making his account immune to scepticism. This means Locke wishes to respect two strands. The Moral Man is a *corporeal* being regarded only in terms of *moral culpability*. Commentators have emphasised the moral strand to severe detriment of the corporeal strand. What is needed is a reconstruction of Locke's theory of personal identity that places embodiment of the person firmly in the picture as Locke intended. It is simply wrong to present Locke as disregarding non-psychological aspects of personal identity.

## IV

### LOCKE & THE RATIONAL PARROT<sup>209</sup>

Locke's discussion of Man reappears in §8 of 'Of Identity and Diversity'<sup>210</sup> in the discussion of the rational parrot. Normally commentators treat the rational parrot argument as if it is Locke's sole argument about the relation between Man and personal identity. This shows the dangers of concentrating on arguments in isolation from the wider text. In this case the effects would be minimised if the rational parrot argument only restated Locke's arguments elsewhere. It is therefore necessary to try to discern to what extent the rational parrot argument is a restatement of Locke's position prior to the Second Edition, and to what extent it is an advance on those views. If, as the previous section suggests, the personal identity discussion was added because of the difficulties Locke found in his discussion of Man one might expect the rational parrot argument to reflect this departure in Locke's thinking.

Before telling the story of the talking parrot, Locke repeats his definition of an animal secured in §§5 & 6 and then makes the major claim of §8:

And whatever is talked of other definitions, ingenuous observation puts it past doubt, that the *Idea* in our Minds, of which the Sound *Man* in our Mouths is the Sign, is nothing else but an Animal of such a certain Form: Since I think I may be confident, that whoever should see a Creature of his own Shape and Make, though it had no more reason all its Life, than a *Cat* or a *Parrot*, would call him still a *Man*[.]<sup>211</sup>

Clearly this statement repeats Locke's earlier contempt for scholastic attempts to define Man in terms of rationality (although we should not forget the role Locke himself gives to rationality in his definition of the Moral Man). This is picked up later when Locke asks

---

<sup>209</sup> Locke is often attacked for using thought-experiments. It would be reasonable to assert that there are no conversing parrots (although there are parrots that repeat the sounds of words parrot fashion!). Yet, Locke is *not* presenting a thought experiment. He takes himself to be presenting reliable evidence. Those who seek to remove thought experiments from the personal identity debate, such as Wilkes (1988), only replace the problem with one of interpretation, i.e. whether unusual cases really do challenge our concepts. There does not seem to be anything to choose between the alleged weaknesses of enquiry in thought alone, and the obscurities of empirical interpretation. Further, no one has ever suggested that Locke's argument fails because parrots do not in fact talk (although maybe someone should). In short, evidence that demands careful appraisal is no better a tool than a carefully conceived thought experiment. For this reason I have not excluded thought experiments, or indulged in a lengthy discussion of their worth. What is more, if Locke had asked us to *imagine* a race of rational animals who were not human it is hard to see what would be wrong with this.

<sup>210</sup> *Essay* 332:35-335:8

<sup>211</sup> *Essay* 333:2-8 (Book II, Chap XXVII, §8)

whether talking parrots “would not have passed for a race of *rational Animals*”.<sup>212</sup> However, Locke’s claim seems to run in direct contradiction to his statements elsewhere that shape *cannot* provide any guide to the distinction between humans, changelings, monsters, drills and so forth. In fact Locke can be extricated from contradiction, but only at cost to the strength of his position.

Although Locke has argued that Man, Changeling, and Monster represent nominal essences, it is also clear that he does regard Man and Beast as real essences. That is, he assumes that the distinction between Man and Beast is a real one, even if our nominal essences do not reveal this.<sup>213</sup> Borderline cases are not at a deep level borderline, but as nominal essences clearly cannot settle which real essence they fall under one should tolerate agnosticism on this issue. Locke’s belief in a real essence of man means he thinks there is a minimal possible definition, and this is shape. This goes back to the Draft A list cited above, with its preponderance of physical characteristics. It is expressed here by the phrase ‘own Shape and Make’. The use of the word ‘Make’ is revealing. Mere external shape would not be enough, however, ‘make’ suggests the inner constitution of the animal. Locke has urged that when the inner constitution of something is fully known, i.e. with a watch, we do go near to knowing something’s real essence. Locke’s own knowledge of medicine and Natural Sciences influence him to believe that biology might ultimately provide the real essence of Man *qua* animal. Ideally science will be able to reveal the internal constitution of Man so it will be as clear as the workings of a watch.

However, Locke does not exclude rationality from his final definition of Man. He finishes §8 with a clear embracing of the Compositional Account.

For I presume ‘tis not the *Idea* of a thinking or rational Being alone, that makes the *Idea* of a *Man* in most Peoples Sense; but of a Body so and so shaped joined to it, and if that be the *Idea* of a *Man*, the same successive Body not shifted all at once, must as well as the same immaterial Spirit go to the making of the same *Man*.<sup>214</sup>

Locke’s aim is not to define Man without reference to rationality. His objection is to those who “shall place the *Identity* of Man in any thing else, but like that of other Animals in one fitly organized Body”.<sup>215</sup> Locke is not claiming that the *definition* of Man should not contain

---

<sup>212</sup> *Essay* 335:1-2 (Book II, Chap XXVII, §8)

<sup>213</sup> He opens Draft B by remarking “Since it is the Understanding that sets man above the rest of sensible beings” (Draft B, §1, p. 101)

<sup>214</sup> *Essay* 335:3-8 (Book II, Chap XXVII, §8)

<sup>215</sup> *Essay* 332:2-4 (Book II, Chap XXVII, §6) emphasis added

reference to rationality. It would wrong just to define Man as 'a rational Being alone'.<sup>216</sup> To give a Man's identity conditions one need only refer to the biological continuity of a human being. Giving identity in terms of immaterial substance would also have to fail. As Man is like other animals there would be inevitable contradiction "Unless they will say, 'tis one immaterial Spirit, that makes the same Life in Brutes; as it is one immaterial Spirit that makes the same Person in Men, which the *Cartesians* at least will not admit, for fear of making Brutes thinking things too."<sup>217</sup>

---

<sup>216</sup> Consider also his comment that "Likewise, to say, that a rational Animal is capable of Conversation, is all one, as to say, a *Man*. But no one will say, That Rationality is capable of Conversation, because it makes not the whole Essence, to which we give the Name *Man*." *Essay* 450:19-22 (Book III, Chap VI, §21)

<sup>217</sup> *Essay* 337:24-7 (Book II, Chap XXVII, §12)

## THE FAMOUS DEFINITION & A NEW INTERPRETATION

Having dismissed the possibility of understanding personal identity by looking at what the idea of Man stands for Locke begins §9 with his famous definition:

This being premised to find wherein *personal Identity* consists, we must consider what *Person* stands for; which, I think, is a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it; It being impossible for any one to perceive, without perceiving, that he does perceive.<sup>218</sup>

This is not the totality of Locke's definition, although it is often taken to be.<sup>219</sup> The definition is qualified throughout Chapter XXVII. Uncovering these qualifications and development and influences elsewhere in the *Essay* is the purpose of this section.

Concentrating on this definition has also led commentators to treat Locke's use of self and person as being synonymous.<sup>220</sup> As already advertised the two should be read and understood apart. Before establishing this claim in detail the following loose distinctions can serve as place holders. By Person Locke means the enduring thing that one is. By Self he means the present substantial union of personal presence, immaterial and/or material substance. This he sometimes calls the *Personal Self*.<sup>221</sup> Locke also has two distinct uses of self. One is merely that of self-identity, the other the name of the Self. These are distinguished quite reliably in Locke's text by the former appearing in plain text and the latter in italics, or with a capital 'S'.<sup>222</sup> This distinction is not unique to Locke, but it is important not to conflate the uses (for example, in the above definition, Locke is not talking about the substantial *self*).

Locke's definition of *Self* is given at §17. Locke writes:

---

<sup>218</sup> *Essay* 335:9-16 (Book II, Chap XXVII, §9)

<sup>219</sup> Wiggins is an influential example of one writer who assumes this. See Wiggins (1980, p. 149 and Chapter 6 at large).

<sup>220</sup> For example Noonan (1998) p. 315. This seems to be a shift from the views of in Noonan (1978) and the more careful considerations of Noonan (1989).

<sup>221</sup> At *Essay* 336:24 & (Book II, Chap XXVII, §10) & *Essay* 337:6 (Book II, Chap XXVII, §11). Locke also thrice uses 'Personality', twice in the same sentence when discussing moral reward, *Essay* 344:5 (Book II, Chap XXVII, §22) and in the same context at *Essay* 346:28 (Book II, Chap XXVII, §26).

<sup>222</sup> I will mark the distinction by following Locke in this convention.

*Self* is that conscious thinking thing, (whatever Substance, made up of whether Spiritual, or Material, Simple, or Compounded, it matters not) which is sensible, or conscious of Pleasure and Pain, capable of Happiness or Misery, and so concern'd for it *self*, as far as that consciousness extends. Thus every one finds, that whilst comprehended under that consciousness, the little Finger is as much a part of it *self*, as what is most so.<sup>223</sup>

For commentators committed to presenting Locke as giving a purely psychological and causal related account of personal identity there is great temptation to ignore this subtlety in Locke. Yet it would be an insult to a great thinker to continue as if the second definition is entirely pleonastic. Although Locke is prone to repetition in his discussion he nearly always introduces nuances with each reworking of a theme. The onus falls to those who would read *Self* and *person* as synonyms to provide argument for this.

There is a decided Cartesian cast to Locke's definition. Descartes had claimed with the *cogito* that we could be certain of our own existence through the certain connection we have with our own thoughts. If thinking confirms our existence, then constant and uninterrupted thinking would confirm enduring existence. Locke subscribes to the *cogito* along with Descartes, but does not accept that the soul always thinks. Locke's approval for the *cogito* is longstanding. In Draft A Locke had written:

The Understanding knows undoubtedly that while it thinks reasons or imagins it is or hath existence or that there is something that knows & understands which according to Cartes & I thinke in truth is the most certein & undoubted proposition that can be in the minde of a man.<sup>224</sup>

Locke's opinion is similar in Draft B where the *cogito* is "a certainty as great as humane nature is capable of"<sup>225</sup>. By the time of the complete *Essay* Locke had separated the argument into the parts he found useful and those he found suspect. He asserts that

In every Act of Sensation, Reasoning, or Thinking, we are conscious to our selves of our own Being; and, in this Matter, come not short of the highest degree of *Certainty*<sup>226</sup>

and

'tis altogether as intelligible to say, that a body is extended without parts, as that any thing *thinks without being conscious of it*, or perceiving, that it does so.<sup>227</sup>

---

<sup>223</sup> *Essay* 341:14-20 (Book II, Chap XXVII, §17)

<sup>224</sup> Draft A, §27, p. 42

<sup>225</sup> Draft B, §35, p. 144

<sup>226</sup> *Essay* 619:1-3 (Book IV, Chap IX, §3)

Locke adopts the Transparency Thesis that follows from the *cogito*. However, he devotes a considerable amount of the first chapter of Book II<sup>228</sup> to attacking the notion that the Soul always thinks. By insisting that the mind always thought, Descartes was able to identify the soul and the mind. This Locke cannot do because he feels that sleep interrupts thought, indeed that “every drowsy Nod shakes their Doctrine, who teach, That the Soul is always thinking”<sup>229</sup>. Hence Locke is obliged to distinguish the soul and the mind. To this end he introduces a definition of consciousness “Consciousness is the perception of what passes in a Man’s own mind.”<sup>230</sup>

Locke’s phrasing ‘a Man’s own mind’ is no accident. It is the mind of a man, that is an animal body united with a soul, Locke refers to. Chapter I of Book II appears in the First Edition of the *Essay* and as such pre-dates the personal identity Chapter. It is also here that Locke makes his first references to personal identity. The example he introduces, that two persons could share one body, is used as a *reduction ad absurdum* of his opponent’s position. However, Locke comes to realise his own position is equally vulnerable to this conclusion. This stands as a further motivation to the Second Edition arguments.<sup>231</sup>

Therefore it is more than worthwhile to pick out the distinctions between Soul, Man, Mind and Consciousness that Locke is making in this chapter to see how they impinge on his later fuller statements about personal identity.

Locke says he must “confess my self, to have one of those dull Souls, that doth not perceive it self always to contemplate *Ideas*”.<sup>232</sup> From this Locke concludes against the Cartesians that “the perception of *Ideas* being ... not [the Souls] Essence, but one of its Operations”.<sup>233</sup> This is consistent with the Lockean view that immaterial substance is the substratum we presume to make thought possible. Locke feels that we can assert from the fact that we sometimes think that there is a thinking thing within us.<sup>234</sup> It is this limited version of the Cartesian argument that Locke is looking to defend.

---

<sup>227</sup> *Essay* 115:15-8 (Book II, Chap I, §19) At *Essay* 108:22-3 (Book II, Chap I, §10) this argument is put as “[I] can conceive it any more necessary for the *Soul always to think*, than for the *Body always to move*”.

<sup>228</sup> *Essay* 108:17-118:31 (Book II, Chap I, §§10-25)

<sup>229</sup> *Essay* 111: 22-23 (Book II, Chap I, §13) In this Locke prefigures Hintikka’s discussion of the *cogito* as performance. See Hintikka (1962).

<sup>230</sup> *Essay* 115:23-4 (Book II, Chap I, §19)

<sup>231</sup> Thiel (1998b), p. 888 is unusual in placing any weight on these passages. Otherwise they are infrequently alluded to and more rarely discussed.

<sup>232</sup> *Essay* 108:20-22 (Book II, Chap I, §10) Hume will later confess to having a soul so dull it cannot perceive itself at all.

<sup>233</sup> *Essay* 108: 23-25 (Book II, Chap I, §10)

<sup>234</sup> *Essay* 108:31, 109:1-2 (Book II, Chap I, §10)

Just as dreaming had caused Descartes so much perplexity, Locke finds sleeping and dreaming the nub of the difficulty. This is inevitable after he grants "that the Soul in a waking Man is never without thought, because it is the condition of being awake".<sup>235</sup> The following discussion is entirely, and explicitly, couched in terms of the Compositional Account of Man. If a Man is a unit of a soul and a body, then the soul could think while the corporeal body slept. Hence the Man, as a unit, might be unaware of the thoughts of the soul. The consequences of this would, Locke thinks, be absurd. Hence he argues:

It is certain, that *Socrates* asleep, and *Socrates* awake, is not the same Person; but his Soul when he sleeps, and *Socrates* the Man consisting of Body and Soul when he is waking are two Persons: Since waking *Socrates*, has no Knowledge of, or Concernment for that Happiness, or Misery of his Soul, which it enjoys alone by it self whilst he sleeps, without perceiving any thing of it; no more than he has for the Happiness, or Misery of a Man in the *Indies*, whom he knows not. For if we take wholly away all Consciousness of our Actions and Sensations, especially of Pleasure and Pain, and the concernment that accompanies it, it will be hard to know wherein to place personal Identity.<sup>236</sup>

The concerns of the Second Edition account of personal identity are in place here in the shape of the importance attached to Consciousness of Actions. What is sharply different is that Locke derives the absurdity of the claim of an always thinking soul by appealing to the coherence of the Compositional Account, i.e. that identity of person should be expected to run with the identity of a Man. Locke clearly has no difficulty in understanding the Man and the Person to be co-conscious. Indeed, Locke derives the absurdity from an appeal to co-consciousness. The periods of the soul's supposed thinking during sleep are *not* co-conscious with the Man, unlike waking states. If they exist at all they are only co-conscious with themselves. However, Locke uses the common-sense account that personal identity goes with the waking man to form the conclusion that the soul *does not* think during sleep. Clearly Locke finds coherent the consequence that if the soul always thinks then two persons might inhabit one Man. However, it is also clear that he *rejects* that this possibility occurs and this is the central plank in his argument against the soul always thinking. This demonstrates incontrovertibly Locke's sympathy for the Compositional Account of Man, and its role in personal identity.

---

<sup>235</sup> *Essay* 109:31-2 (Book II, Chap I, §11)

<sup>236</sup> *Essay* 110:12-22 (Book II, Chap I, §11) Locke shows his first concerns here over the possibility that consciousness might be transferred, although here he is quite damning of the possibility. Also note that consciousness of thought and consciousness of happiness and misery have not yet been separated out as they are in the 2<sup>nd</sup> Edition so the former concerns the person and the latter the self.

This sympathy is continued in Locke's tone in the next section. Locke supposes it possible "the Soul of *Castor*, whilst he is sleeping, retired from his Body"<sup>237</sup> but only because this "is no impossible Supposition for the Men I have here to do with, who so liberally allow Life, without a thinking Soul to all other Animals. These Men cannot then judge it impossible, or a contradiction, That the Body should live without the Soul; nor that the Soul should subsist and think, or have Perception, even Perception of Happiness of Misery, without the Body."<sup>238</sup> It is clear from Locke's tone that he disagrees with the truth of these Cartesian propositions. In this section Locke suggests a case where a soul migrates between two Men. The passage that concludes this section reads:

Just by the same Reason, they make the Soul and the Man two Persons, who make the Soul think apart, what the Man is not conscious of. For, I suppose, no body will make Identity of Persons, to consist in the Soul's being united to the very same numerical Particles of matter: For if that be necessary to Identity, 'twill be impossible, in that constant flux of the Particles of our Bodies, that any Man should be the same Person, two days, or two moments together.<sup>239</sup>

This could be read as an attack on the Compositional Account. The addition of this comment at this point seems dissonant. Locke is of course right that no-one would (or indeed could) make identity of persons consist in the union of a soul with a numerically identical collection of particles of matter. However, Locke is only attempting a clarification of the Compositional Account. He identifies the person and the Man here. If personal identity consists in sameness of Man, *qua* unit of soul and body, then this union must allow for the change of parts usual in a living body. This is consistent with Locke's aims in this section. Identity of person is assumed to be identity of Man, and Man is a unit of soul and living body. This has to be granted by those who would allow the sharing of a soul between two bodies, or their account would lapse into incoherence, for the Soul would constantly find itself in Union with different bodies. This would result, though trivially, in a succession of different Men being present. Although modern writers like Lewis<sup>240</sup> might accept this consequence, Locke finds it incoherent. This being granted, Locke forces the point to demonstrate the incoherence of the soul-sharing account and its premise that the Soul can be in union with a body forming a Man, and the Man yet remain unconscious of the Soul's perceptions.

---

<sup>237</sup> *Essay* 110:28 (Book II, Chap I, §12)

<sup>238</sup> *Essay* 110: 29-34 (Book II, Chap I, §12) Note Locke's emphasis that the soul might not be necessary for perception, but is surely necessary for the perception of happiness and misery.

<sup>239</sup> *Essay* 111:14-21 (Book II, Chap I, §12)

<sup>240</sup> See Lewis (1976) and (1983b)

Locke concludes that "If such a definition [that the soul always thinks] be of any Authority, I know not what it can serve for, but to make many Men suspect, That they have no Souls at all, since they find a good part of their Lives pass away without thinking".<sup>241</sup> He resolves "Not to mention again the absurdity of two distinct Persons, which follows from this Supposition".<sup>242</sup>

The supposition that follows from Locke's position is that the body plays an essential facilitating role in thinking. In fact, Locke regards the proper unit of rational thought as the Man and not the Soul, Mind or indeed Consciousness. If the Man is properly rational, then the Man is also, by Locke's own lights, a Rational Being, the very sort of thing that Persons are. The Moral Man discussion emphasised the corporeal. Here the role of the corporeal is explicitly linked with discussion of personal identity.<sup>243</sup> Even more tellingly Locke's discussion of the body's role in thought is carried out in terms of memory.

Locke considers a counter argument to his claim that the mind does not always think; "Twill perhaps be said, That the *Soul thinks*, even in the soundest *Sleep*, but the *Memory retains it not*."<sup>244</sup> Locke rounds on this counter-argument in various ways. First he argues that "Most Men, I think, pass a great part of their Sleep without dreaming."<sup>245</sup> Whether we all dream, and for how much of our sleeping existence we dream, is a question of perpetual dispute, and no more settled today than it was in Locke's time. Whatever one is inclined to think of Locke's assertion that common experience tells us that we do not dream all the time we sleep<sup>246</sup> this argument is not Locke's most interesting and can be put to one side.

---

<sup>241</sup> *Essay* 116:8-11 (Book II, Chap I, §19)

<sup>242</sup> *Essay* 112:22-23 (Book II, Chap I, §15)

<sup>243</sup> The discussion in this chapter attempts to follow the chronological development of Locke's views. Clearly this is difficult for those arguments that appear simultaneously for the first time in the First Edition. The earliest arguments are about Man as a substance, as clearly shown by the discussion in Drafts A & B. I would suggest that the Moral Man argument predates in composition the arguments about the always thinking soul which mark a mid-point between the former argument and the Second Edition personal identity passages. The unpublished Draft C may shed light on this (as might other manuscripts). Without this material (and of course this material may not shed any light) one can only rely on the circumstantial evidence of placing each argument in a likely progression of thought. Although it would be interesting to establish the order of composition in Locke's argument this is not essential here. His final position, in the sense of arguments of latest composition, is clearly that of the Second Edition additions. It suffices to demonstrate that Locke was attempting to deal with conflicting themes and that his final position emerges from them. It does not matter if he held them sequentially or simultaneously (and one must allow that the intellectual process is rarely purely linear, and that old notions leave their ghosts if not their corpses).

<sup>244</sup> *Essay* 111:29-30 (Book II, Chap I, §14)

<sup>245</sup> *Essay* 112:1-2 (Book II, Chap I, §14)

<sup>246</sup> *Essay* 112:2-8 (Book II, Chap I, §14)

More interestingly Locke claims that “*To think often, and never to retain it so much as one moment, is a very useless sort of thinking*”.<sup>247</sup> Locke then makes his claim about the body’s role in memory:

[I]t will be said, that in a waking Man, the materials of the Body are employ’d, and made use of, in thinking; and that the memory of Thoughts, is retained by the impressions made on the Brain, and the traces there left after such thinking; but that in the *thinking of the Soul*, which is not perceived *in a sleeping Man*, there the Soul thinks apart, and *making no use of the Organs of the Body*, leaves *no impression on it, and consequently no memory of such Thoughts*.<sup>248</sup>

The possibility of so-called ‘Intellectual Memory’, i.e. purely spiritual memory, intrigues Cartesians and Cartesian commentators. Descartes’ acceptance that the body is responsible for most memory led him into an awkward compromise that the Soul must be allowed some faculty of memory itself. Locke clearly finds the notion of Intellectual Memory unpromising, and provides arguments against it.<sup>249</sup>

Locke argues that to suggest that the Soul can acquire Ideas without the body should imply that it can remember these Ideas as well, otherwise “to what purpose does it think?”<sup>250</sup> This forces a dilemma. Either the Soul does not remember, which Locke has further arguments against, or it does remember, but in a manner which is not co-conscious with the Man. From this flows ‘the absurdity of two persons’. What is more, to claim the Soul can think without remembering is tantamount to materialism. Locke’s target is a certain kind of materialism that reduces memories to organisation of matter, presumably with no role for thinking matter:

Characters drawn on Dust, that the first breath of wind effaces; or Impressions made on a heap of Atoms, or animal Spirits, are altogether as useful, and render the Subject as noble, as the Thoughts of a Soul that perish in thinking; that once out of sight, are gone for ever, and leave no memory behind them.<sup>251</sup>

These separate attacks on Materialist and Intellectual theories of memory makes no sense other than against the background assumption of a Compositional Account of Human Beings and their rational capacities. It seems clear that Locke believes that thinking matter thinks, but lays up its storehouse of memory in the corporeal body. In short, Locke can only make sense of memory as a function of the Composite Man. With this presumption

---

<sup>247</sup> *Essay* 112:9-10 (Book II, Chap I, §15)

<sup>248</sup> *Essay* 112:15-22 (Book II, Chap I, §15)

<sup>249</sup> For lengthy and fascinating discussion of Locke’s theory of memory see Sutton (1998).

<sup>250</sup> *Essay* 112:31 (Book II, Chap I, §15)

Locke can conclude that thinking matter does not think apart from the body, or remember without it. This is justified because to continue to suppose we think in our dreamless sleep would imply

[T]hat our infinitely wise Creator, should make so admirable a Faculty, as the power of Thinking, that Faculty which comes nearest the Excellency of his own incomprehensible Being, to be so idly and uselessly employ'd, at least  $\frac{1}{4}$  part of its time here, as to think constantly, without remembering any of those Thoughts[.]<sup>251</sup>

Locke relies on the Compositional Man even more clearly when he asks us to consider the nature of the dreams we do remember;

How *extravagant* and incoherent... they are... This I would willingly be satisfied in, Whether the Soul, when it thinks thus apart... from the Body, acts less rationally... If its separate Thoughts be less rational, then these Men must say, That the Soul owes the perfection of rational thinking to the Body: If it does not, 'tis a wonder that our Dreams should be, for the most part, so frivolous and irrational; and that the Soul should retain none of its more rational Soliloquies and Meditations.<sup>252</sup>

Locke also insists that the Body is the source of all Ideas, otherwise

'Tis strange, the Soul should never once in a Man's life, recal over any of its pure, native Thoughts, and those *Ideas* it had before it borrowed any thing from the Body<sup>253</sup>

Locke also feels the ignorance and lack of reason in Children to be further proof that the Soul does not always think.

[F]ew Signs of a Soul accustomed to much thinking [are found] in a new born Child, and much fewer of any Reasoning at all. And yet it is hard to imagine, that the rational Soul should think so much, and not reason at all... [Instead] Infants,[...] spend the greatest part of their time in Sleep, and are seldom awake, but when... some... violent Impression of the Body, forces the mind to perceive[.]<sup>254</sup>

---

<sup>251</sup> *Essay* 112:34-6, 113:1-3 (Book II, Chap I, §15)

<sup>252</sup> *Essay* 113:5-9 (Book II, Chap I, §15)

<sup>253</sup> *Essay* 113:16-26 (Book II, Chap I, §16) Unless, it might be argued against Locke, it is deemed useful by God for us to have thoughts we *do not* remember. Indeed passages of fanciful and irrational thought might be useful if they are a gateway to enlightenment as some suggest.

<sup>254</sup> *Essay* 114:8-10 (Book II, Chap I, §17) This argument is used in an adapted form against the transmigration of the soul. See *Essay* 339:5-34 (Book II, Chap XXVII, §14)

<sup>255</sup> *Essay* 116:25-32 (Book II, Chap I, §21)

A foetus also thinks very little, and this is because it is denied any external stimulus.<sup>256</sup>

The weight of evidence from Book II, Chap I is overwhelming in its support for reading Locke as subscribing to the Compositional Account of Man. It is also clear that Locke uses this very assumption to attack alternative theories of personal identity. It is also clear that Locke placed memory and rational activity, that is rational being in a Man, not solely with that thinking part of us that we are so in the dark about. Locke's appeal to the Benevolence of Creation as guarantor of the parsimony of mental activity is akin to Descartes' appeal to Nature's design of Soul and Body so as to be a perfect fit for each other. This is compatible with an evolutionary account of human beings, as long as evolution can be made a suitable surrogate for Divine Benevolence. A naturalistic evolutionary account favours Animalism, for it bases understanding of human endeavours on humans being biological entities. Identifying persons and animals is a natural part of this. The question, then, is to establish to what extent Locke does deviate from this version of the Cartesian Compositional Model in the Second Edition personal identity discussion. As has already been suggested there is a decidedly Cartesian cast to Locke's Famous Definition. In light of Locke's explicit use of Cartesian themes elsewhere this can now be spelt out in detail.

---

<sup>256</sup> *Essay* 117:2-9 (Book II, Chap. I, §21)

## VI

### THE COMPOSITIONAL RECONSTRUCTION OF THE FAMOUS DEFINITION

There is nothing in the first part of §9 that is incompatible with the Book II, Chapter I Compositional Account. A person is “a thinking, intelligent Being”,<sup>257</sup> “the same thinking thing”,<sup>258</sup> “the sameness of rational Being”.<sup>259</sup> As Locke regards the Compositional Man as a rational being (i.e. a rational existence) there is no dissonance. Locke also repeats the Transparency Thesis<sup>260</sup> while emphasising “When we see, hear, smell, taste, meditate or will anything, we know that we do so.”<sup>261</sup> All of these perceptions are according to Locke dependent on the body.<sup>262</sup> Locke gives a role to substance in the sameness of *self* (not person), but emphasises that “It [is] not being considered in this case, whether the same *self* be continued in the same, or divers Substances.”<sup>263</sup>

The uniting principle is *consciousness*. Locke states “consciousness always accompanies thinking, and ‘tis that, that makes every one to be, what he calls *self*”.<sup>264</sup> Locke is creating a tight knot of concepts here. It is important to note that Locke talks of being the same *self* with oneself, and not the same person. This is because Locke is reserving *self* for (what will be called for the moment) the instantiated person. Sensation and Perception (both inherently bodily) create the sense of Self *whatever substances presently constitute the Self*. The Transparency Thesis provides an awareness of whatever belongs to the Self and “thereby distinguishes [it] from all other thinking things”.<sup>265</sup> In all of “this alone consists *personal Identity*”.<sup>266</sup> There is nothing in this that is necessarily antagonistic to Locke’s previous commitment to the Compositional Man being the basis of personal identity.

What differs from before is the additional statement:

---

<sup>257</sup> *Essay* 335:10-11 (Book II, Chap XXVII, §9)

<sup>258</sup> *Essay* 335:12 (Book II, Chap XXVII, §9)

<sup>259</sup> *Essay* 335:24 (Book II, Chap XXVII, §9)

<sup>260</sup> *Essay* 335:14-16 (Book II, Chap XXVII, §9)

<sup>261</sup> *Essay* 335:16-17 (Book II, Chap XXVII, §9)

<sup>262</sup> Meditation, it might be objected, is not. However, even Descartes gave some role to the body in imagination, e.g. when discussing our ability to understand in reason what a thousand sided shape is like, while being unable to picture the shape in our mind’s eye in Meditation VI (ATVII: 72-74; CSM II: 50-51). If Descartes allows a role for the body in such acts of contemplation, then surely Locke, with his suspicion for Intellectual Memory, would to. There seems no necessity to read meditate here to be a purely intellectual exercise, and every reason to imagine Locke was suspicious of such purely intellectual acts.

<sup>263</sup> *Essay* 335:19-21 (Book II, Chap XXVII, §9)

<sup>264</sup> *Essay* 335:21-2 (Book II, Chap XXVII, §9)

<sup>265</sup> *Essay* 335:22-23 (Book II, Chap XXVII, §9) Hence Locke provides an account of a person’s boundaries, something which is claimed to allude Hume’s bundle theory. See §IV of the next chapter for arguments that this claim is a spurious objection anyway

<sup>266</sup> *Essay* 335:23-24 (Book II, Chap XXVII, §9)

And as far as this consciousness can be extended backwards to any past Action or Thought, so far reaches the Identity of that *Person*; it is the same *self* now it was then; and 'tis by the same *self* with this present one that now reflects on it, that that Action was done.<sup>267</sup>

In the understanding of this one sentence the greater part of all the woes of Lockean scholarship concerning personal identity are to be found. Consciousness here should be understood as the unifying principle of substances that creates the *self*. The *self* now can come to stand in *some relation*<sup>268</sup> to a *self* that performed a previous 'Action or Thought'. This constitutes the identity of a person. Locke provides no better clue than that consciousness can be 'extended backwards'. Nonetheless, the following sections contain material enough to construct a reading compatible with Locke's Compositional sympathies, and his scepticism about knowledge of substance, especially Man. In short Locke substitutes his reliance on Man in Book II, Chapter I for a reliance on Rational Being. A Man is a Rational Being, but here the emphasis is on the Cartesian *cogito* and the sureness of knowledge provided by The Transparency Thesis. A Rational Being always knows the extent of its substantial *self*. The Rational Being of Chapter 27 is the heir to Locke's Moral Man, a corporeal rational being.

It has seemed beyond even the obvious to many that Locke is referring to the faculty of memory when he talks about consciousness being extended backwards.<sup>269</sup> The historical precedent for this is, of course, Reid's famous criticism of Locke. Reid insists:

Mr. Locke attributes to consciousness the conviction we have of our past actions, as if a man may now be conscious of what he did twenty years ago. It is impossible to understand the meaning of this, unless by consciousness be meant memory, the only faculty by which we have immediate knowledge of our past actions.<sup>270</sup>

It is of course true that memory is the faculty by which we know about the past. There is also a subset of memories that not only provide us with knowledge about the past, but also appear to inform us about our own pasts. This is personal or autobiographical memory. It

---

<sup>267</sup> *Essay* 335:24-28 (Book II, Chap XXVII, §9)

<sup>268</sup> Though for reasons discussed in §I above this is not a relation in Locke's technical sense.

<sup>269</sup> See, for example Carruthers' blithe assertion that Locke held a memory theory in Carruthers (1986) p. 78. Quinton presents a version of the memory theory which he takes to be completely in the spirit of Locke. See Quinton (1962). Shoemaker (1959) takes himself to be similarly motivated. Perry attacks the memory theory, but also takes Locke to be the source of this theory. See Perry (1975b).

<sup>270</sup> Reid (1785b), p.115 in Perry (1975a). Reid's criticism is dependent not only on a typically wilfully abrupt reading of Locke but his own insistence that memory provides direct or immediate access to our pasts. Such a theory of memory is arguably incoherent, and any objection based on it must therefore suffer.

is this latter form of memory that is usually supposed to be the basis of personal identity.<sup>271</sup> Loosely the claim is that a person is constituted by the collection of memories to which each individual has a unique and peculiar psychological relation.

Consciousness is a relation persons have presently with themselves, and a relation they have to their past selves. However, this does not imply that Consciousness is memory.<sup>272</sup> It would be trivial to insist that all our relations to the past are necessarily memories. Locke does indeed distinguish between memory and extended consciousness. Bringing this out will sideline memory interpretations of Locke.

First it should be noted that Locke has a developed theory of memory given in Book II, Chap X, 'Of Retention'.<sup>273</sup> Locke is perfectly aware of what memory is, and as his discussion of the soul's alleged perpetual thinking above demonstrates, is willing to make philosophical use of memory to prove arguments. If this were the case it would seem odd at the very least that if Locke intended a memory theory he didn't just use the term memory to explicate it. Instead he introduces the apparatus of extended consciousness. It seems unlikely that he went to this difficulty unnecessarily. There are two motivating factors. Locke felt it highly probable (although not proved) that memory relied on the body, and the Compositional Man. Therefore, while attempting to give an account of personal identity that is indifferent to substantial composition, Locke needs to use a neutral term. The second reason is simpler. Locke simply does not intend to provide a memory theory.

Nonetheless various passages give easily to a memory reading. Locke writes early on in his discussion:

For as far as any intelligent Being can repeat the *Idea* of any past Action with the same consciousness it had of it at first, and with the same consciousness it has of any present Action; so far it is the same *personal self*.<sup>274</sup>

---

<sup>271</sup> Hughes in his defence of Locke's theory emphasises, correctly I feel, the importance of non-personal memory in reinforcing our belief in the reliability of personal memory. See Hughes (1975).

<sup>272</sup> Many commentators simply assume that Locke identifies memory with extended consciousness. Laird, pithy as ever, manages an interesting interpretative near-miss when he writes "Locke's general conclusion, sometimes expressed with some diffidence but, in the main, very firmly, was that personal identity lasted only *so long as* the same continued consciousness lasted. [...] Such reflex appropriation, it would seem to many of us, need not be identified with explicit memory. Locke, however, interpreted it in no other sense, even forensically." Laird (1932) p. 169. If only Laird had pursued the doubts he attributes 'to many of us' (indeed, if this many had not themselves been so silent) Locke's subtle and intriguing opinion on this matter might have been correctly established sooner. Laird, despite asking the right question, leaves prevailing orthodoxy unscathed.

<sup>273</sup> *Essay* 149:19-155:21

<sup>274</sup> *Essay* 336:21-24 (Book II, Chap XXVII, §10)

Locke's theory of memory is that past Ideas are recreated in the image of the original. The problems of such an account aside, the phrase 'repeat the *Idea* of an past Action' is consistent with Locke's theory of memory. However, the account is qualified in important ways. The relationship between a repeated Idea and the 'Idea of a past Action' is not clear. Ideas caused by sensation can be repeated as representations. But this is mere repetition of an Idea. Yet Locke emphasises it is the Idea of a past Action. This suggests something more akin to personal memory. To repeat a past Action in the mind is to re-experience it, not just to represent it.

The evidence for this distinction is in the latter part of the quotation. The consciousness of a past act is the same as the consciousness of a present Action. It seems clear that the relation a person stands in to a present Action is not one of (mere) memory. This implies the relation consciousness has to a past Action cannot be (mere) memory either. The repetition of consciousness has something distinct.

The waters are muddied further when Locke introduces what some have taken to be a prototype of Q-memories. In the light of our ignorance about "what kind of Substances they are, that do think"<sup>275</sup> we cannot settle "whether the consciousness of past Actions can be transferr'd from one thinking Substance to another."<sup>276</sup> The issue *would* be settled according to Locke were it not for this difficulty:

I grant, were the same Consciousness [of a past Action] the same individual Action, it could not [be transferred]: But it being but a present representation of a past Action, why it may not be possible, that that may be represented to the Mind to have been, which really never was, will remain to be shewn. And therefore how far the consciousness of past Actions is annexed to any individual Agent, so that another cannot possibly have it, will be hard for us to determine, till we know what kind of Action it is that cannot be done without a reflex Act of Perception accompanying it, and how perform'd by thinking Substances, who cannot think without being conscious of it.<sup>277</sup>

Here Locke clearly embraces a representation account for the presentation of past Actions. This is placed in contrast with the ideal situation in which this is not possible. This would seem to destroy the privileged reading of consciousness suggested above. However, Locke's reasons for believing in the possible transfer of memory are *premised* on the privileged reading. Consciousness is the experience of being intimately related to an action, the experience that *I* did it. Locke can find no evidence of actual memory transfer occurring

---

<sup>275</sup> *Essay* 337:31-32 (Book II, Chap XXVII, §13)

<sup>276</sup> *Essay* 337:31-33 (Book II, Chap XXVII, §13)

so he appeals to what he regards as an analogous situation for possible evidence.

Sometimes we have "representations in Dreams, [...] which yet, whilst dreaming, we take for true."<sup>278</sup>

It has already been noted Locke remained impressed by the *cogito* even in the mature *Essay*. Here Locke makes clear use of a related Cartesian tactic, the dreamer hypothesis. In a dream we appear to have a relation to an Action which is *whilst dreaming* indistinguishable "with that consciousness, which remarks our waking Thoughts."<sup>279</sup> Locke feels this establishes that while the test of an Action being our own is that it is performed with the same consciousness that we would perform a present action, that this awareness can be confused. We can have the phenomenological experience of a veridical experience but be deluded. Dreaming acquaints us with this possibility because we experience false veracity in dreams, but upon awakening realise our error. Likewise Locke regards memory as a certain acquaintance with past events, accompanied by a peculiar phenomenological quality. It is therefore equally susceptible to delusion. For this reason Locke recognises and explicitly makes a case *against* a memory theory. Memory cannot provide the sort of extended consciousness that Locke is looking for. Our experience teaches us that memory is unreliable. Locke's prime motivation is to exclude unreliable factors from an account of personal identity. Hence he refers to memory as "*that which we call the same consciousness*"<sup>280</sup> rather than same consciousness proper.

Locke is clearly aware that he must explain

[W]hy one intellectual Substance may not have represented to it, as done by it self, what it never did, and was perhaps done by some other Agent... [and] may not possibly be without reality of Matter of Fact

Locke feels that in the absence of a refuge against this possibility our notions will

[T]ill we have clearer views of the Nature of thinking Substances, be best resolv'd into the Goodness of God, who as far as the Happiness or Misery of any of his sensible Creatures is concerned in it, will not by a fatal error of theirs transfer from one to another, that consciousness, which draws Reward or Punishment with it.<sup>281</sup>

---

<sup>277</sup> *Essay* 337:33-338:6 (Book II, Chap XXVII, §13)

<sup>278</sup> *Essay* 338:11-12 (Book II, Chap XXVII, §13)

<sup>279</sup> *Essay* 336:6-7 (Book II, Chap XXVII, §10)

<sup>280</sup> *Essay* 338:6-7 (Book II, Chap XXVII, §13) emphasis added

<sup>281</sup> *Essay* 338:14-18 (Book II, Chap XXVII, §13)

With God guaranteeing that our memory will only reflect consciousness as far as consciousness is actually extended Locke can safely conclude that what we represent to ourselves actually happened.<sup>282</sup> This is important, because Locke is quite clear that forgetfulness does occur. Equally Locke is quite sure that forgetfulness has no effect on personal identity. Even if God does not guarantee the faculty of memory we can trust that God's review of our consciousness will be perfect and fair, even if ours is flawed.<sup>283</sup>

The extension of consciousness is like a temporally extended version of the *cogito*. At any point in time this consciousness allows someone to be aware of what presently constitutes themselves. This is true of spatial extension, so our self is extended into that part of matter consciousness goes with. Locke gives the example that "Upon separation... should this consciousness go along with the little Finger,[...], 'tis evident the little Finger would be the *Person*, the same *Person*".<sup>284</sup> Then Locke asserts this is true temporally; "[C]onsciousness ... constitutes this inseparable *self*: so it is in reference to Substances *remote in time*."<sup>285</sup>

Locke's term 'inseparable *self*' is intriguing. If Locke is holding a memory theory he is obliged to claim that persons have gaps in their existences, and that parts of their pasts can be lost forever through (ordinary) forgetfulness. This would seem to conflict with the *self* being inseparable. How is this to be understood when considering the passage where Locke discusses forgetfulness? Locke begins the discussion while considering if the Person is one identical substance:

This few would think they had reason to doubt of, if these Perceptions, with their consciousness, always remain'd present in the Mind, whereby the same thinking thing would be always consciously present, and, as would be thought, evidently the same to it self.<sup>286</sup>

This clearly draws on the certainty of the Transparency Thesis. Locke agrees that if we were Cartesian Souls that thought all the time, then we should be sure of our identity. Locke also seems to assume that this would make us sure we were the same substance. This is a departure from modern writers who presume that changes of whatever substance supports thinking could take place without interruption of first-person experience. Locke

---

<sup>282</sup> The obvious correlate of this is that we should reliably be able to identify imaginings as imaginings.

<sup>283</sup> Locke discusses Pascal's prodigious memory, but remarks that only Angels regularly can have all of their pasts laid out in their minds at one time while "humane Minds are confin'd to... having ... *Ideas* only in succession, not all at once." *Essay* 154:20-21 (Book II, Chap X, §9). Locke's appeal to Man's finite powers to explain our fallibility recalls Descartes' arguments in Meditation IV.

<sup>284</sup> *Essay* 341:20-24 (Book II, Chap XXVII, §17)

<sup>285</sup> *Essay* 341:25-28 (Book II, Chap XXVII, §17)

provides no reason why constant first-person experience would provide a guarantee against change of substance. Presumably as he is examining a full-blooded Cartesian version he does not think the problem could arise because thought is the essence of thinking matter. It would be odd if a substance's essence could be interrupted without awareness. This, though, is all by-the-by as Locke is about to dismiss the relevance of the full Cartesian account to personal identity. Locke continues:

But that which seems to make the difficulty is this, that this consciousness, being interrupted always by forgetfulness, there being no moment of our Lives wherein we have the whole train of all our past Actions before our Eyes in one view: But even the best Memories losing the sight of one part whilst they are viewing another; and we sometimes, and that the greatest part of our Lives, not reflecting on our past selves, being intent on our present Thoughts, and in sound sleep, having no Thoughts at all, or at least none with that consciousness, which remarks our waking Thoughts. I say, in all these cases, our consciousness being interrupted, and we losing the sight of our past *selves*, doubts are raised whether we are the same thinking thing; *i.e.* the same substance or no. Which however reasonable, or unreasonable, concerns not *personal identity* at all.<sup>287</sup>

That Locke concludes that forgetfulness and interrupted consciousness 'concerns not *personal Identity* at all' would, one would have thought, been enough to end speculation about Locke's allegiances in this matter. Locke's point is merely a re-iteration of the earlier Book II, Chapter I attack on the Cartesian Soul. Locke rejects the view that the unifying feature of a thinking being can be uninterrupted continuity of thought. If the soul, as Cartesians maintain, did always think then this could be the case. However, as Locke demonstrated we often do not think. Locke therefore needs to find some other unifying principle. He finds this by analogy with Animal identity:

Different Substances, by the same consciousness (where they do partake in it) being united into one Person; as well as different Bodies, by the same Life are united into one thinking Animal, whose *Identity* is preserved, in that change of Substances, by the unity of one continued Life.<sup>288</sup>

---

<sup>286</sup> *Essay* 335:30-34 (Book II, Chap XXVII, §10) Compare this with both Descartes' and Locke's remarks about how an embodied angel would apprehend their corporeal abutment.

<sup>287</sup> *Essay* 335:34-336:11 (Book II, Chap XXVII, §10) Locke's misgivings about memory are long-standing. On Saturday 22<sup>nd</sup> January 1678 Locke wrote in his Journal "For the minde endeavouring to retaine only the traces of the patterne loosening by degrees a great part of them and not havinge the liberty to supply any new colours or touches of its owne, the picture in the memory every day fades and growes dimmer and is often times quite lost." (Aaron & Gibb 1936, p. 104). See also Draft B, §23, p. 134 for the precursor to 'Of Retention' (Book II, Chap X) in the published *Essay*. Locke also comments that "the memory is very weake" at Draft B, §27, p. 137. Memory is not discussed in Draft A.

<sup>288</sup> *Essay* 336:14-18 (Book II, Chap XXVII, §10).

In his discussion of Lockean Lives earlier Locke had avoided suggesting that the relation of one body to the next was a causal one. It is reasonable to assume that he does not see the union of substances in a person as a causal relation either. What is more, if the Lockean Life of a Person is consciousness this seems to preclude a memory theory. There is no analogy between memory and a life. There is, however, an analogy between union of substance into an animal life, and self-aware union of substances into a thinking being.

There is one further passage that needs dealing with to establish the non-memory reading of consciousness.<sup>289</sup> Locke's discussion of the Drunk Man and The Sober Man has produced a small industry in Lockean scholarship, not to mention the extensive correspondence with Molyneux on the topic.<sup>290</sup>

Locke maintains that there is something peculiar in drunkenness that removes consciousness. Putting aside whether this state is self-induced, and therefore still blameworthy, one needs to make further sense of this claim. Drunkenness can cause forgetfulness, that much is common-knowledge. What, though, is of more interest to Locke is the possibility that Drunkenness can cause *present* action to occur without consciousness. It is in this sense that Locke finds a parallel with the sleep-walker. While drunk (one would presume *extremely* drunk) Locke feels it is possible to act without knowing one is doing so.<sup>291</sup> Drunkenness therefore is a possible example of that "kind of Action... done without a reflex Act of Perception"<sup>292</sup> that Locke had mentioned earlier in his discussion of the transfer of consciousness. Drunkenness, in other words, seems capable of causing a failure of the Transparency Thesis. Locke feels if this failure is genuine, then someone could be excused.

---

<sup>289</sup> The example of the Day-Man and Night-Man given in §23 also appears to mention forgetfulness and that "the Mind many times recovers the memory of a past consciousness, which it had lost for twenty Years together." *Essay* 344:36-345:1 (Book II, Chap XXVII, §23). However, it is not forgetfulness that makes the example coherent but that we could "suppose two distinct incommunicable consciousnesses" *Essay* 344:18 (Book II, Chap XXVII, §23). Forgetfulness is not the same as being incapable of recalling something, and so this example I take not to rely on memory (although Locke *illustrates* it in this manner to make the sceptical possibility appear closer to actual experience).

<sup>290</sup> In De Beer (1979): Letter 1609, from William Molyneux, 2<sup>nd</sup> March 1693 (Vol. 4:650), Letter 1620 Locke to W.M., 28<sup>th</sup> March 1693 (Vol. 4:665), Letter 1655, Locke to W.M., 23<sup>rd</sup> August 1693 (Vol. 4:722), Letter 1661 from W.M., 16<sup>th</sup> September 1693 (Vol. 4:729), Letter 1685, from W.M., 23<sup>rd</sup> December 1693 (Vol. 4:767), Letter 1693, Locke to W.M., 19<sup>th</sup> January 1697 (Vol. 4:785), Letter 1712, from W.M., 17<sup>th</sup> February 1694 (Vol. 5:21), Letter 1744, Locke to W.M., 26<sup>th</sup> May 1694 (Vol. 5:58). For discussion see Helm (1981) and Allison and Jolley (1981).

<sup>291</sup> Locke had a dispute with Thomas Stringer about the ownership of portrait. Stringer claimed Locke had given him the painting. Locke claims to have no recall of this and remarks in a letter "I am sure I could never give it away in a letter to anyone, or say anything like it unless I were drunk when I writ it." This is maybe indicative that Locke's view of drunkenness was based on peculiar personal experience. For a discussion of the whole affair see Yolton (1985, pp. 10-14).

<sup>292</sup> *Essay* 338:3-4 (Book II, Chap XXVII, §13)

However, the extent of this possible excuse is not delimited by memory as those who have fulminated over this case assume. Locke is attempting to provide a theory of moral retribution as well as one of personal identity. Consciousness transforms here into Conscience.<sup>293</sup> This is the moral counterpart of consciousness. For two states to be personally co-conscious they have to be performed with the same consciousness even if done at different times. In this sense the co-consciousness relationship is the same for simultaneous as for temporally distinct actions. The Conscience is the totality of moral actions that fulfil this relationship. Memory does not necessarily uncover all these connections. Lockean bodily memory could not do so, and Locke denies the existence of Intellectual memory<sup>294</sup>. Thus Locke writes:

But in the great Day, *wherein the Secrets of all Hearts shall be laid open*, it may be reasonable to think, no one shall be made to answer for what he knows nothing of; but shall receive his Doom, his Conscience accusing or excusing him.<sup>295</sup>

Divine judgement will discern the true relations of consciousness that our finite intellectual capacities cannot always reveal to us. Conscience, being a repository of every act a person performed with full consciousness, is the marker of how each person will be judged.

The same approach explains §20<sup>296</sup> where Locke considers what would happen on the supposition “I wholly lose the memory of some parts of my Life, beyond a possibility of retrieving them, so that perhaps I shall never be conscious of them again” and asks “yet am I not the same Person?”<sup>297</sup> Locke solves this by asking the reader to “take notice what the Word *I* is applied to, which in this case is the Man only.”<sup>298</sup> *I*, meaning the consciousness, cannot come to be separated from its past. Recall that the self is ‘inseparable’. The Man, however, could be united with different consciousnesses, although these would have to be “incommunicable”.<sup>299</sup> Once again memory is linked with the Man and not the person.

This also strikes a blow against those who attempt to construe Locke as holding a prototype Constructivist account. These writers are impressed by Locke’s comment that “whatever

---

<sup>293</sup> See Winkler (1991b) for an interpretation of Locke along these lines. Rovane uses the moral side of Locke’s theory as inspiration for her Neo-Lockean account of personal identity, although she explicitly claims to deviate from Locke on many points. See Rovane (1999) esp. pp. 3-12.

<sup>294</sup> Although he does allow that Angels might have all their actions laid before them at all times, *Essay* 154:21-24 (Book II, Chap X, §9) This would seem to tell decisively against the role of human memory in constructing conscience.

<sup>295</sup> *Essay* 344:8-12 (Book II, Chap XXVII, §22) emphasis added.

<sup>296</sup> *Essay* 342:23-343:4 (Book II, Chap XXVII, §20)

<sup>297</sup> *Essay* 342:23-26 (Book II, Chap XXVII, §20)

<sup>298</sup> *Essay* 342:28-29 (Book II, Chap XXVII, §20)

past Actions [the *self*] cannot reconcile or appropriate to that present *self* by consciousness, it can be no more concerned in, than if they had never been done".<sup>300</sup> However, it should be clear, that consciousness (or Conscience) could not *arbitrarily* appropriate or ignore Actions. Those performed by the same consciousness belong to the person, those not so performed, do not and cannot. That Locke's theory is guaranteed by divine benevolence should not be counted against it, especially when one remembers that Locke is aiming towards an account compatible with prevalent Protestant Eschatology.<sup>301</sup>

Locke's attempt to yield a moral theory from personal identity also tempts him to find a legal theory as well. One must agree with Locke that it is the practice of law to punish the person. However, Locke notes in the Drunken and Sober Man discussion that "Humane Laws punish both with a Justice suitable to [Justice's] way of Knowledge",<sup>302</sup> that is by way of the facts established by *third-person* criteria rather than *first-person* criteria.<sup>303</sup>

Sensitivity to this distinction comes out towards the end of the personal identity discussion. Locke remarks that "[Person] is a Forensick Term"<sup>304</sup> which is to say a legal term. Locke precedes this comment by remarking that "*Person*, as I take it, is the name for this *self*. Where-ever a Man finds, what he calls *himself*, there I think **another may say** is the same *Person*."<sup>305</sup> This chimes in perfectly with Locke's comment that courts must punish the Man. It is where a Man finds himself that others might say they find him as well. Usually this will be where the Person is found, for as Locke remarks "I know that in the ordinary way of speaking, the same Person, and the same Man, stand for one and the same thing."<sup>306</sup>

---

<sup>299</sup> *Essay* 342:32 (Book II, Chap XXVII, §20)

<sup>300</sup> *Essay* 346:35-38 (Book II, Chap XXVII, §26)

<sup>301</sup> For discussion of the reaction to Locke's theory among theologians see Tennant (1982).

<sup>302</sup> *Essay* 343:38, 344:1 (Book II, Chap XXVII, §22) the reading 'Justice's way of Knowledge' rather than 'the accused's way of Knowledge' is widely accepted to be the correct way to settle the ambiguity in Locke's text here. See Allison & Jolley (1981).

<sup>303</sup> Although consider this case collected in *The Fortean Times* (No. 124) originally reported in *The Melbourne Herald* of 25<sup>th</sup> February 1999: "Two elderly Alzheimer's patients had 'a bit of a scrap' at a care centre in New Zealand's North Island. When police arrived 15 minutes later, one of the men felt sore and the other couldn't remember the fight. In court last February, the police explained to judge Pat Treston that they were obliged to withdraw the case because neither man could recall the incident. Quipped Judge Treston: 'What incident?'" Some would detect a Lockean tone in the judge's sense of humour.

<sup>304</sup> *Essay* 346:26 (Book II, Chap XXVII, §26)

<sup>305</sup> *Essay* 346:24-26 (Book II, Chap XXVII, §26) emphasis added. Locke's shift between first-person and third-person identification is of interest here, especially given the role Hume later gives this distinction. See next chapter, Section X. Locke's comment here that "*Person*, as I take it, is the name for this *self*" is not in dissonance with my claim the two terms are distinct in Locke's usage. The self is the person at a moment. In his discussion of legal practice Locke is asserting that normally we take the presence of a self to entail the presence of a person. That is a conscious being will remain the same conscious being in a reliable fashion. Which, for purposes of the law means we can rightly punish the same man with confidence that we are punishing the same person. This explains why where a man finds himself another may say he finds the same person. To find oneself is to be conscious. Being so conscious is enough for another to base their identity judgements upon. All this, of course, presumes the Compositional Account.

<sup>306</sup> *Essay* 340:18-20 (Book II, Chap XXVII, §15)

## VII

### LOCKE & ANIMALISM & THE MAN AGAIN

Locke's legal interpretation of the person rests clearly on the supposition that the Man and the Person are reliably one. If, ultimately, this is not true it does not matter for the apportioning of blame on judgement day is God's business, and not ours. Man should be suitable for our day to day business. As has been shown, though, in the published *Essay* Locke has all but thrown up his hands in dismay when trying to define Man. Locke, therefore, is trying to bring together two instincts. He does not want to premise any philosophical or moral conclusion on an account of Man that may be shown to be incorrect. Equally he wants to give an account of how we feel ourselves to be identical and persisting beings.

Locke does have instincts about what Man might turn out to be, and the implications this would have for his theory. Throughout the personal identity discussion he remarks on what he feels is most likely:

I know that in the ordinary way of speaking, the same Person, and the same Man, stand for one and the same thing.<sup>307</sup>

I agree the more probable Opinion is, that this consciousness is annexed to, and the Affection of one individual immaterial Substance.<sup>308</sup>

I am apt enough to think I have in treating of this Subject made some Suppositions that will look strange to some Readers, and possibly they are so in themselves. But yet I think, they are such, as are pardonable in this ignorance we are in of the Nature of that thinking thing, that is in us, and which we look on as our *selves*. Did we know what it was, or how it was tied to a certain System of fleeting Animal Spirits; or whether it could, or could not perform its Operations of Thinking and Memory out of a Body organised as ours is; and whether it has pleased God, that no one such Spirit shall ever be united to any but one such Body, upon the right Constitution of whose Organs its Memory should depend, we might see the Absurdity of some of those Suppositions I have made.<sup>309</sup>

In other words, if we could be sure of the Composition of Man, then one could adopt a theory of personal identity that identified the Man and the Person. Locke without these

---

<sup>307</sup> *Essay* 340:18-20 (Book II, Chap XXVII, §15)

<sup>308</sup> *Essay* 345:25-27 (Book II, Chap XXVII, §25)

<sup>309</sup> *Essay* 347:13-24 (Book II, Chap XXVII, §27)

doubts would adopt Descartes' Compositional Account and would be far more an animalist than a psychological theorist.

With his theory of personal identity Locke believes he has achieved the aims he had with the Moral Man account: moral attribution safe from the vagaries of substance. His account is one where "the precise real Essence of the Things moral Words stand for, may be perfectly known".<sup>310</sup> Locke ends the chapter 'Of Identity & Diversity' by returning once more to his initial point of departure, the definition of Man. He reviews three possible definitions. That if "a rational Spirit be the *Idea* of a *Man*, 'tis easie to know, what is the *same Man*, viz. the *same Spirit*"<sup>311</sup> or while "a rational Spirit vitally united to a Body of a certain conformation of Parts"<sup>312</sup> remain in union it is the same Man. Finally the Idea of Man might be "but the vital union of Parts in a certain shape".<sup>313</sup> How we choose between these accounts does not matter.

For whatever be the composition whereof the complex *Idea* is made, whenever Existence makes it one particular thing under any denomination, the same Existence continued, preserves it the same individual under the same denomination.<sup>314</sup>

If any of these beings are thinking intelligent Beings then they would of course also be persons.

---

<sup>310</sup> *Essay* 516:23-24 (Book III, Chap XI, §16)

<sup>311</sup> *Essay* 348:11-13 (Book II, Chap XXVII, §29)

<sup>312</sup> *Essay* 348:13-14 (Book II, Chap XXVII, §29)

<sup>313</sup> *Essay* 348:17-18 (Book II, Chap XXVII, §29)

<sup>314</sup> *Essay* 348:21-25 (Book II, Chap XXVII, §29)

## CHAPTER 3

### HUME ON PERSONAL IDENTITY

Hume's views on personal identity are notorious for two reasons. One is his supposed scepticism about personal identity. The second is his admission of confusion over the issue. Nonetheless Hume still has an important contribution to make to the personal identity debate.

Hume's views on personal identity are profoundly influenced by Locke.<sup>315</sup> Locke's quandaries arise from the doubts his empiricism throws up for the Compositional Account favoured by the Cartesians. Hume shares Locke's empiricist drive, but his principles are more extreme and his conclusions more dramatic.<sup>316</sup> Locke was prepared to resolve personal identity by way of a quasi-substantial entity, 'a suitable concrete'. Hume, though, rejected outright the notion of substance and had serious misgivings about the very notion of identity.

The extremity of Hume's account means that it is typically regarded as an interesting failure.<sup>317</sup> Penelhum describes Hume's theory as "radically defective"<sup>318</sup> while Waxman goes as far as to remark that "it is little wonder that [Hume's] account of personal identity tends to be regarded as a splendid example of how *not* to handle this problem."<sup>319</sup> However, as Hume's conclusions are frequently misappropriated and self-confessedly confused, it is difficult to judge whether his account does fail. To judge this aright it is necessary to establish what Hume's aims were in giving an account of personal identity. Just as Descartes and Locke have suffered, Hume too has been bent into the shape of contemporary

---

<sup>315</sup> Roland Hall has persuasively suggested that Hume must have written with Locke's *Essay* in front of him, so many are the co-incidences. The broad structure, a treatment of identity and then a treatment of personal identity, matches Locke. Also Hume uses examples of near exact similarity. For more see (Hall 1974).

<sup>316</sup> Broadly, therefore, the picture of Locke's influence on Hume assumed here is the traditional one. Laird's opinion on the matter is worth quoting, being both insightful and instructive: "The traditional doctrine according to which Hume's highly developed empiricism and thorough-going phenomenalism show what Locke should have held had he been persistently clearheaded is, up to a point, tolerably accurate. It fails, however, to perceive that Locke, in his own way, was as stubborn a rationalist as ever he was an empiricist, and that a distinct tincture of his rationalism remained in Hume's philosophy, essentially unmodified if ultimately disowned, although the range of it was much more rigorously restricted than with Locke." Laird (1932) p. 50. Interestingly, the normally brutally precise Laird is unable to source this 'traditional' view. It is a curious fact that although philosophy's traditional views are often alluded to their origins remain obscure. Consequently one cannot help but entertain the possibility that no-one ever held them sincerely. Instead they are, in fact, fraudulent traditions, nourished not by heritage or pedigree, but laziness and convenience.

<sup>317</sup> See for example, Noonan (1989) p. 102-3, Perry (1975a) p. 26ff.

<sup>318</sup> Penelhum (1967) p. 98

<sup>319</sup> Waxman (1994) p. 223

issues.<sup>320</sup> Hume is probably more in sympathy with modern concerns than either Locke or Descartes, but his concerns are nonetheless quite distinct from present-day ones.

Key to understanding both Hume's aims and the subsequent labyrinth of the Appendix is Hume's extreme empiricist outlook in Book I of the *Treatise*.<sup>321</sup> This is best achieved by attributing to Hume a principle that, although he never explicitly states, I believe is well supported by the text. The principle, which I call the Humean Insight, is this: that whatever it is that constitutes personal identity this should be accessible to us in a suitable manner (I say suitable manner to avoid pre-judging what this involves).

Hume's difficulties arise because personal identity is the fulcrum upon which the rest of his philosophy refuses to balance. On the one hand he claims that perceptions are bonded by a "secret connexion"<sup>322</sup> that we cannot ever know. On the other hand he speaks of

This object [...] self, or that succession of related ideas and impressions, of which we have an intimate memory and consciousness.<sup>323</sup>

The Humean Insight, I believe, captures this tension between the intimacy of the self and the distance of causation. It captures what is still valuable in Hume's account and explains his quandary in the Appendix.

Hume presents not one, but two accounts of personal identity in the *Treatise*. The explicit theory is contained in Book I, but there is a complimentary account in Books II & III.

Hume draws attention to his differing concerns when he writes:

[W]e must distinguish betwixt personal identity, as it regards our thought or imagination, and as it regards our passions or the concern we take in ourselves.

---

<sup>320</sup> For a discussion of the state of Hume scholarship see Capaldi (1992). Shoemaker (1986) p. 3 has a perceptive survey of the ways Hume has been misappropriated (not least by Shoemaker himself).

<sup>321</sup> Madell is alone in also emphasising this writing "Hume's account of personal identity is very obviously an empiricist one in its desire to see personal identity as unmysterious, and in the view that it is to be explained in terms of observable connections between experiences" Madell (1981) p.2ff. However, Madell condemns both Hume and what he identifies as the Empiricist Tradition of personal identity. On the contrary I feel the Humean Insight provides a useful test that any theory of personal identity should be able to pass. If personal identity is mysterious serious argument is needed to justify accepting such a position. Madell believes he can offer this, following Butler, Reid and Swinburne.

<sup>322</sup> *Enquiry* 66:12 (VII, I) All references are to the Nidditch edited edition of the *Enquiry*, Hume (1975). Each has a page and line numbers marking the beginning and end of the quote (e.g. *Enquiry* 144:3-12). These are followed by numerals indicating Hume's own textual divisions (e.g. VII, II).

<sup>323</sup> *Treatise* 277:17-19 (II, I, I) All references are to the Nidditch edited edition of the *Treatise*, Hume (1978). Each has a page and line numbers marking the beginning and end of the quote (e.g. *Treatise* 144:3-12). These are followed by numerals indicating Hume's own textual divisions (e.g. I, II, II). In the case of the chapter 'Of Personal Identity' there is a forth numeral indicating the paragraph number (e.g. I, IV, VI, V).

The first is our present subject; and to explain it perfectly we must take the matter pretty deep[.]<sup>324</sup>

Hume's theory concerning personal identity 'as it regards our thought or imagination' has involved critics far more than Hume's scattered comments about personal identity 'as it regards our passions or the concern we take for ourselves.' The always pertinent Kemp Smith is right to insist that we must ask:

Why it is that in Book I of the *Treatise* the existence of an impression of the self is explicitly denied, while yet his theory of the 'indirect' passions, propounded at length in Book II, is made to rest on the assumption that we do in fact experience an impression of the self, and that this impression is ever-present to us?<sup>325</sup>

That there is a question is partly Hume's own fault. He offers a whole chapter on personal identity as a metaphysical quandary, but no explicit systematic account of personal identity as it concerns the passions.<sup>326</sup> However, this omission on Hume's part is in itself revealing. The chapter 'Of Personal Identity' takes a *specific* metaphysics of personal identity to task. Hume wishes to rid philosophy of otiose concepts and explanations. Hume is not antagonistic towards the notion of a person as an everyday expedient, nor does he deny that we have concern for ourselves. Hume's negative project needs careful and explicit statement. On the contrary Hume's positive theory is (at least in Hume's own assessment) so straightforward as to be obvious. A theory that emphasises the experience of concern for ourselves, and places everything of weight with that experience, has no need to further explain itself. This lack of developed positive theory in part explains Hume's confusion in the Appendix.

Few parts of the philosophical canon are as notorious and unexpected as the Appendix recapitulation. Any interpretation of Hume's theory of personal identity must account for Hume's self-doubts. The Humean Insight does provide one such amelioration. This will

---

<sup>324</sup> *Treatise* 253:19-23 (I, IV, VI, V). Waxman suggests that the "self of 'thought or imagination' is the condition for the representation of all imperfect identities (bodies, other minds, substances)" Waxman (1994) p. 224. This suggests Hume's interest is not the traditional problem of personal identity at all, at least not in Book I. As will be discussed later, Hume does not make use of a notion of imperfect identity. Therefore, he has no need for personal identity to underwrite this practice. The discussion here will proceed in the belief that Hume's concern was more traditional, i.e. what constitutes a person, and how arrive at the idea of that constitution.

<sup>325</sup> Kemp-Smith (1941) p. v. Kemp Smith, in this quote at least, appears to conflate the self being intimately known with having an *impression* of the self.

<sup>326</sup> Indeed in Book II he says of personal identity "'Tis evident, that the idea, or rather impression of ourselves is always intimately present with us, and that our consciousness gives us so lively a conception of our own person, that 'tis not possible to imagine, that any thing can in particular go beyond it.'" *Treatise* 317:26-30 (II, I, XI)

also be achieved by looking beyond the chapter 'Of Personal Identity' and the Appendix. Garrett is typical when he says; "Although [Hume] devotes several pages in the Appendix to stating his misgivings, Hume does not succeed in clearly stating any specific problem with his earlier account. And no trace of the topic recurs in the first *Enquiry*, nor in any of Hume's other writings."<sup>327</sup> In fact Hume is quite clear where he feels his difficulty to lie, and the Humean Insight captures this difficulty.

By adopting this approach it is possible to show that Hume would tolerate a version of the Compositional Account. It may even be that Hume would more than tolerate such an account, but does in fact embrace one. What is more his account is even closer to a modern Animalism than either Descartes' or Locke's. Such a reading is contrary to a long line of critics who perceive of Hume as holding a purely psychological criterion of personal identity.<sup>328</sup>

Clearly the Naturalist Hume presented by Kemp-Smith is more likely to be an Animalist than the arch-sceptical Hume of the Logical Positivists. In what follows I side with 'The New Hume' rather than with 'The Old Hume'.<sup>329</sup> The conflicts between the Old and New Hume are the key Humean themes: causation, external objects, and persons. The Old Hume is a sceptic about all of these. The New Hume is a causal realist and a realist about objects according to his assorted champions.<sup>330</sup> The New Hume's opinion of personal identity has not been widely discussed but it reasonable to assume the New Hume is also a realist about persons. The New Hume's realist commitments are clearly a package. The Old Hume's alleged rejection of the existence of persons is supposedly the final flourish of Book I's sceptical tirade. The New Hume, all things being equal, should end with an account of personal identity that although Sceptically Realist is nonetheless Naturalistic about persons.

As will emerge in what follows just such a Naturalism and Realism about persons is contained in the first book of the *Treatise*. In part this can only be established by accepting that Hume is a realist about both causation and objects. However, quite independent investigation can present a strong case for Hume being firstly a realist about persons, secondly accepting that persons are natural and thirdly that they are natural because they are

---

<sup>327</sup> Garrett (1997) p. 164. For more on Hume's comments on personal identity outside the *Treatise* see McIntyre (1993). Although Garrett is right that Hume makes no explicit comments about personal identity several of the themes return particularly the role of introspection and intimacy. See *Enquiry* 64ff. (VII, I).

<sup>328</sup> Flew regards Hume's theory as utterly erroneous for just this reason. See Flew (1986) pp. 91ff. Significant exceptions are Baier (1979) and Capaldi (1975). For others who ascribe a purely psychological account to Hume see note 4 in the Introduction.

<sup>329</sup> The terms are borrowed from Winkler (1991b) pp. 541-44 & 575-78.

human animals. This account will be stronger when taken with the other doctrines of Sceptical Realism. Hopefully the account below will be persuasive enough, though, to stand alone as it is not possible to fully defend the New Hume within the scope of this work.

The *Treatise* is a notoriously difficult work, and the chapter 'Of Personal Identity' a famously difficult chapter. Given not only the amount, but the diversity, of secondary material about Hume the following treatment is inevitably lengthy. The first section establishes which position Hume is attacking. His target is dubbed here The Metaphysicians' Self. Consequently Hume is attacking the notion that the self is simple and perfect in its identity. §II sketches Hume's view of identity as systematic fiction. §III examines the psychological mechanisms Hume develops to explain how the fiction of identity arises. §§IV-VI explore Hume's account of identity in depth, looking first at identity in material bodies, and then in more complex cases. Having established the arguments Hume made use of prior to his arguments about personal identity §VII looks at personal identity itself. It is argued that Hume's central difficulty with the Metaphysicians' Self is with the role of introspection and §VIII deals with these issues. §§IX & X then present a new interpretation of Hume based on the Compositional Account. Finally the Appendix is looked at. §XI explains why prevailing interpretations are insufficient and §XII attempts to explain the Appendix in terms of the Compositional Account. The purpose of this is twofold. It both demonstrates that the Compositional Account was at the heart of Hume's original account and establishes that Hume's worries arise from this position and his beliefs about introspection.

---

<sup>330</sup> Primarily Strawson, G. (1989) and Wright (1983) who both follow in the spirit of Kemp-Smith (1941). See Winkler (1991b) for useful discussion as well as Capaldi (1991) & (1992)

## I

### HUME'S COMPLAINTS & THE SELF OF THE METAPHYSICIANS

The chapter 'Of Personal Identity' has both a negative and a positive stage. The negative phase is far clearer than the positive. Hume's positive position is normally taken to be typified by his claim that "I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions".<sup>331</sup> This is, in fact, quite mistaken. However, the context of Hume's apparent positive claims makes it difficult to tell when Hume is still on the attack and when he is not. Hume does hold a bundle theory of the *mind* but this does not imply or commit him to a bundle theory of *personal identity*. These questions of interpretation will be dealt with later. First Hume's negative account of personal identity will be outlined.

Hume levels his complaints against 'certain metaphysicians', although no one is named.<sup>332</sup> This picture of the Self will be called The Metaphysicians' Self. Hume lists three distinct claims:

- 1) "We are every moment intimately conscious of what we call our self."<sup>333</sup>
- 2) "We feel its [the self's] existence and its continuance in existence."<sup>334</sup>
- 3) "[We] are certain, beyond the existence of a demonstration, both of its perfect identity and simplicity."<sup>335</sup>

Hume then distinguishes two defences he feels are used to justify these three claims:

- 1D) 'Sensations' *increase* our belief in the self.<sup>336</sup>

---

<sup>331</sup> *Treatise* 252:31-33 (I, IV, VI, IV)

<sup>332</sup> Except Shaftesbury at *Treatise* 254n1. It is likely that Butler is a major influence. Butler's *Dissertation on Personal Identity* was published in 1736, three years before the *Treatise* (as von Leyden notes, 1957, p. 347). Greig (1931) states that "[Hume] desired both the acquaintance and the approbation of the Rev. Dr Butler, whose *Analogy of Religion* had been published only the year before... [however] he failed to meet Butler, though he tried at least once." p. 95-6. Church draws an interesting link to Malebranche "One of the metaphysicians whom Hume sets apart from this conclusion concerning the generality of mankind may have been Malebranche. The 'intimate consciousness' of the self that Hume fails to find would then be his translation of 'sentiment intérieur' Malebranche thought us to have of the existence, but not the essence, of the self." Church (1935) p. 104ff. However, as Church points out, Malebranche did not believe we could introspect the soul. See also Laird (1932) pp. 165-6 for the Malebranche connection. Laird also adds Anthony Collins, Henry Dodwell and Clarke into the mix (1932) p.166ff. For more historical colour see Thiel (1998b).

<sup>333</sup> *Treatise* 251:4-5 (I, IV, VI, I)

<sup>334</sup> *Treatise* 251:6 (I, IV, VI, I)

<sup>335</sup> *Treatise* 251:7-8 (I, IV, VI, I)

<sup>336</sup> *Treatise* 251:8-12 (I, IV, VI, I)

2D) “To attempt a farther proof of this were to weaken its evidence; since no proof can be derived from any fact, of which we are so intimately conscious; nor is there any thing, of which we can be certain, if we doubt of this.”<sup>337</sup>

(1) Hume agrees with, repeatedly commenting on the intimacy we have with ourselves.<sup>338</sup> However, these comments come in Book II of the *Treatise* by which time Hume has established that the self is only equivalent with a “succession of related ideas and impressions, of which we have an intimate memory and consciousness.”<sup>339</sup> Hume’s aim in this chapter is to explore how the vulgar and other philosophers respond to the basic fact of this intimacy. Hume is not suspicious of the claim that we are intimate at every moment with ourselves *as long as this is properly understood*. (1) is compatible with Hume’s own wider account (that we are never aware of anything other than individual perceptions<sup>340</sup>). What he seeks to expose are the wrong moves made on the basis of (1). The first of these is the shift to holding (2).

Hume cannot embrace (2), because he does not hold that the self continues perfectly identical through its existence. He does feel, though, that *if* there is a thing called the self, then it should be the *sort* of thing of which we *are* intimately conscious, and that we should be aware of how it exists over time. Of course, properly understood, we do feel the continuance and existence of the self over time, because we are intimately aware of perceptions. Hume, though, is highlighting the road to the erroneous belief in a simple identical self. (2) is acceptable to Hume in as much as being intimately aware of a perception is as much as we ever do know the self. However, (2) is unacceptable to Hume as commonly understood because it remains ambiguous whether the perception or the self itself is felt. Hume holds quite clearly that we cannot find our self as a distinguishable part of our perceptions.<sup>341</sup>

In the Appendix capitulation he states that “no connexions among distinct experiences are ever discoverable by human understanding. We only *feel* a connexion or a determination of the thought”.<sup>342</sup> Hume does not seem to distinguish between self-knowledge in the sense of knowing the self, and self-knowledge in the sense of knowing the contents of our own mind. Of course, for Hume both of these are merely constituted by the perceptions, but one could know individual perceptions without being aware of how they are part of the same

---

<sup>337</sup> *Treatise* 251:12-15 (I, IV, VI, D)

<sup>338</sup> E.g. *Treatise* 277:17 (II, I, D), 286:9 (II, I, V), 317:27 (II, I, XD), 320:16 (II, I, XD), 339:14 (II, II, II), 354:13 (II, II, IV), 427:25 (II, II, VII).

<sup>339</sup> *Treatise* 277:17-19 (II, I, D)

<sup>340</sup> *Treatise* 252:15-17 (I, IV, VI, III)

<sup>341</sup> *Treatise* 189:17-190:3 (I, IV, II)

succession. Not fully understanding the consequence of this distinction is a weakness in Hume's initial account. Once again intimacy plays a role in Hume's worries. Feeling a connexion is itself a type of intimacy. Yet we cannot be intimate with a causal connexion, because these connexions are ultimately mysterious.

(3) Hume rejects. The Metaphysicians' Self is alleged to be perfect and simple. Refuting this is the central drive of Hume's attack. Further, this is also part of the Humean Insight. Any firm metaphysical claim about our personal identity cannot be mysterious and indemonstrable. Hence, by the same token, Hume rejects (2)D, that any attempt of a proof will undermine knowledge of our own personal identity. And Hume rejects the claim that sensations deliver up any idea of personal identity, for it will be his very claim that *no* impression of the self is to be found in our perceptions.

(3) is the crux of the Humean dilemma. Although no proof can be offered for the existence of the self, its existence is something intimately known to us in our perceptions. The naturalness with which we know the self is the only proof Hume feels might be forthcoming. Indeed in Book II he will come to suggest that:

'Tis evident, that the idea, or rather impression of ourselves is always intimately present with us, and that our consciousness gives us so lively a conception of our own person, that 'tis not possible to imagine, that any thing can in particular go beyond it.<sup>343</sup>

At this stage, though, Hume is still committed to analysing why we feel we can go beyond the intimacy of self-perceptions to some other truth.

(1)D, like (1), Hume is in part in agreement with. Such a claim is compatible with a bundle theory of the mind. Hume would associate this claim with Cartesians who believed that sensations of the body confirmed the union of mind and body, and therefore confirmed knowledge of the soul.<sup>344</sup> Hume makes no further use of this objection. His comments only two paragraphs later that upon introspection all one can find is perceptions would seem to agree with this claim. If the self is just a bundle of perceptions then sensations cannot help but increase our belief in the self. Sensations are a source of self-awareness. However, they are impugned here because they are inclined to increase our belief in the fiction of a perfectly identical self.

---

<sup>342</sup> *Treatise* 635:23-26 (App, X)

<sup>343</sup> *Treatise* 317:26-30 (II, I, XI)

<sup>344</sup> See previous note 332 for the importance of Malebranche in this connection.

Hume, of course, is quite aware of the difference between 'self' used in the Metaphysician's sense and used in self-reference. The first two-paragraphs discuss the role of intimacy in apprehending the self, but are not spelt out in first-person terms. Intimacy, as already noted, plays a crucial role for Hume. Therefore, introspection also plays a role. Hume reserves the terms "*myself*"<sup>345</sup> and "*himself*"<sup>346</sup> to describe introspection. The first-person introspective passage<sup>347</sup> remains entirely non-committal about what relation the perceptions discovered in introspection might bear to selves, persons, minds or any other entity. In this passage Hume is merely asserting the impotence of introspection in discovering anything about our own composition. Significantly this is the only part of Hume's discussion that is carried out in exclusively first-person terms. This is not surprising given the role of introspection in Hume's positive theory.

Hume begins the discussion by describing the theory of self he attributes to other philosophers, the Metaphysicians' Self. In the first two paragraphs he uses the word 'self'<sup>348</sup> exclusively *except* on one occasion when he writes "But self or person is not any one impression, but that to which our several impressions and ideas are suppos'd to have reference."<sup>349</sup> Obviously it is of great weight whether Hume here establishes an equivalence between self and person. Arguably he does not. Instead 'self' is used to describe the theory of other philosophers. 'Person' is only used when Hume begins his analysis. 'Person' is introduced as Hume's own term, to be contrasted with 'self'. Hume regards 'self' when used to refer to a putative metaphysical entity as a bankrupt term. 'Person', however, he seeks to adopt and rehabilitate. Paraphrased, then, Hume asserts that we have no single impression either of the Metaphysician's Self nor of our Person. Two things need to be noted. One that Hume asserts no equivalence, and that his usage of 'person' is introduced without definition. This rather implies that its meaning is a natural and common-sense one.

Hume's attack, therefore, is precise and deliberate. He attacks a certain view of the self, The Simple view.<sup>350</sup> The question remains, though, as to whether Hume is carrying out an attack on *all* notions of the self. Clearly, he cannot be doing this, because there are theories of the self other than the Simple View. However, Hume might have taken himself to be attacking all notions of the self if he believed that the Simple View exhausted the

---

<sup>345</sup> *Treatise* 252:12, 252:15, 252:19 (I, IV, VI, III)

<sup>346</sup> *Treatise* 252:25, 252:29 (I, IV, VI, III)

<sup>347</sup> *Treatise* 252:6-30 (I, IV, VI, III)

<sup>348</sup> *Treatise* 251:5, 251:11 (I, IV, VI, I), 251:18, 251:22, 251:27, 251:29, 252:4 (I, IV, VI, II)

<sup>349</sup> *Treatise* 251:24-6 (I, IV, VI, II)

alternatives to his own theory. Hume is taken to be making just such a claim by Twentieth Century empiricists who adopt Hume as their own.<sup>351</sup> They take Hume to be defending a 'No-self' account of personal identity. The 'No-self' view claims that talk of selves is premised on a logical error, or that the self is a logical construct and not a real metaphysical entity. Hume does feel that identity is a *psychological* error, and that ideas of perfectly identical selves are constructs of the imagination. It is less clear where Hume feels this leaves the *metaphysical* status of persons. However, none of this commits Hume to a 'No-self' view. Broadly this is because the 'No-Self' view is anti-realist and Hume's theory of personal identity is a realist one. Strawson pithily captures the faults of this interpretation of Hume:

[The] current misinterpretation – travesty – of Hume as some sort of prototypical logical positivist rests almost entirely on supposing him to restrict his view of what could exist to what his epistemology admits as knowable or directly experienceable. And so it is that the great sceptical expositor of the vast extent of human ignorance is held to believe that there is definitely nothing we cannot know about (or at least considered as a heroic foreshadower of this view). I can think of no greater irony in the history of philosophy.<sup>352</sup>

Clearly it greatly affects our view of Hume if we decide his attack is a general rather than a limited one. It is clear that Hume is attacking a certain picture of personal identity, the Metaphysicians' Self. It is also obvious how one could take the bundle theory to be a *denial* of the existence of persons.<sup>353</sup> However, this would be too glib. Hume's denial can only be understood along with his own theory of identity. The idiosyncrasies of Hume's version of identity means that Hume could consistently deny the existence of perfectly identical and simple persons while asserting the existence of persons. Stated so boldly this might seem paradoxical. This appearance says more about Hume's account of identity than it does about the personal identity debate at large. However, if this is correct, Hume is not involved in a general attack against persons. Hume should be seen as conducting a limited attack against the theories prevalent while he was writing. These views he combined (and slightly caricatured) as the Metaphysicians' Self. Hume cannot be claimed as a proto-positivist hero. Concerning bodies Hume adopts a limited attack, stating "We may well ask, *What causes induce us to believe in the existence of body?* but 'tis vain to ask, *Whether*

---

<sup>350</sup> Parfit (1984) coins this term to describe theories which rely on a single immutable entity to ground personal identity. These he contrasts with the Complex View, in which persons are taken to have proper parts. For extensive discussion of the Simple View see the excellent Southgate (1994).

<sup>351</sup> See further Capaldi (1992) for discussion of the offences of the Logical Positivist reading of Hume. Carruthers persists in describing Hume as holding a No-Self theory. See Carruthers (1986) p. 51.

<sup>352</sup> Strawson, G. (1989) p. 66-7.

*there be body or not?*<sup>354</sup> One might suggest that the Humean question of personal identity should be why we believe in selves, not whether or not selves exist.

Regarding Hume as making a limited attack also rules out one of the more persistent objections to Hume. Hume is taken to claim that when he introspects he cannot find the thing that would be his self, and that he has no idea of the self (although no part of the text exactly fits with this reading). It is objected that if Hume has no idea of the self, then he cannot be sure what he is looking for, and cannot claim therefore to have not found it.<sup>355</sup> Yet, Hume does have an idea of the self, the same idea of an identical simple self that all of us are inclined to have. It is this that he cannot find when introspecting. Hume's limited attack is on this notion of the self, and it is this notion of the self that Hume claims he can find no evidence for when introspecting. Seen like this Hume's claim is consistent, and the objection seems quite trite.<sup>356</sup>

Having established that it is the Metaphysicians' Self that Hume attacks, a lengthy diversion is necessary into Hume's account of identity. This done, Hume's explanation of the emergence of the fallacy of the Metaphysicians' Self can be put in perspective. This in turn will finally allow Hume's positive theory of personal identity to emerge.

---

<sup>353</sup> For example Perry (1975a) introduces extracts from Hume under the title "Abandonment of Personal Identity".

<sup>354</sup> *Treatise* 187:25-27 (I, IV, II)

<sup>355</sup> Shoemaker (1986) p. 3-4 & Noonan (1989) p. 83. Although every philosopher I know is familiar with this 'traditional objection', and undergraduates seem able to produce it with tedious regularity under the guise of a standard response, it seems to rarely make it into print. Noonan, for example, cites the objection as one "noted by many commentators" but cites no specific authors. Typically both Noonan and Shoemaker note the opinion but do not hold it themselves. Philosophy's traditional arguments often seem to exist as an oral tradition, perpetuated by uninspired teachers and students alike, and given currency by frequent repetition in dull discussions at duller dinner parties.

<sup>356</sup> All the more so considering Hume's clear description of his procedure at *Treatise* 633:24-29 (App. II)

## II

### HUME ON IDENTITY

Hume regards the problem of personal identity as peculiarly difficult. He remarks more than once that it needs separate treatment apart from a general account of identity.<sup>357</sup>

However, by the chapter 'Of Personal Identity' Hume asserts the problem can be explained in the same manner as "the identity of plants, and animals, and ships, and houses, and all the compounded and changeable productions if either art or nature."<sup>358</sup> However, he still devotes a whole chapter to establishing this claim.

Therefore, before discussing Hume's account of personal identity his general account of identity needs to be examined.<sup>359</sup> Hume regards identity as a perfectly legitimate relation listing it as one of the seven philosophical relations:

*Identity* may be esteem'd a second species of relation. This relation I here consider as apply'd in its strictest sense to constant and unchangeable objects: without examining the nature and foundation of personal identity, which shall find its place afterwards. Of all relations the most universal is that of identity, being common to every being, whose existence has any duration.<sup>360</sup>

The difference between identity and personal identity is previewed here. Hume's general account of identity is concerned primarily with the identity of 'constant and unchangeable objects'. This is the strict sense of identity, or as Hume typically calls it *Perfect Identity*. Persons, however, are changeable and their identity therefore particularly requires explanation. Hume's account of the identity of changeable objects grows out of his account of the identity of unchangeable objects. Hume's discussion of Perfect Identity as 'strict' has encouraged some commentators to suppose Hume thought of other identities as being

---

<sup>357</sup> At *Treatise* 14:27-29 (I, I, V) and 189:30-190:3 (I, IV, II). Hume notes the question is "abstruse" (189:31) and therefore demands "recourse to the most profound metaphysics" (189:34-5) and further that the meaning of person "in common life 'tis... never very fix'd nor determinate." (189:35-6 & 200:1). It is the difficulty of the question, therefore, that forces its separate consideration, not any difference in the problem's nature.

<sup>358</sup> *Treatise* 259:6-9 (I, IV, VI, XV).

<sup>359</sup> Waxman also adopts this approach, but in a moment of self-aggrandisement suggests "It also puts one in a better position to reckon with a possibility that, to my knowledge, has been neglected by every previous interpreter: that personal identity, thought the subject of *Treatise* I/iv/§6, is nevertheless premised in the account of identity of bodies in iv/§2 and substances in iv/§§3-5." Waxman (1994) p. 201. In fact many commentators have been aware of Hume's earlier comments on identity. Waxman is right, though, that the connection between the two discussions has been under-appreciated. Waxman may rest easy that such a mistake will not occur here.

<sup>360</sup> *Treatise* 14: 25-31 (I, I, V)

Imperfect or loose and popular identities (following Butler's distinction).<sup>361</sup> This, it will be argued, is not the case.

This section will examine Hume's general account of identity by looking at his comments concerning the identity of bodies. Then Hume's re-examination of identity as part of his discussion of personal identity will be considered. A large number of issues need to be examined. First, Hume's use of Perfect Identity needs to be clarified. Once this is clear it will be possible to settle whether Hume believes we ever encounter anything that is Perfectly Identical. This done Hume's talk of improper, imperfect, supposed, fictional and attributed identities will be considered, as well as the occasions when he uses the word 'identity' alone and without seeming qualification. The aim of this is to establish if Hume does propose a second legitimate sense of identity along with Perfect Identity.

Everyone must concede that some identity statements are erroneous. Hume, however, might go further and claim that *all* identity statements are errors, fallacies or fictions.

Strawson explains clearly why Hume could be thought to hold this:

Clearly, anyone who accepts a basically Humean (or more generally strict classical empiricist) epistemology is going to be led or forced... to consider... an account according to which all that ultimately exists are perceptions. For however dubious such a Strict Idealist theory is, it is (given Humean epistemology) quite clearly the *only way* not to go beyond the kind of thing one can know to exist, and know the nature of, or have direct experience of, in giving an account (and an 'ontologically outright' account) of the nature of whatever one is trying to give an account of. It is, in fact, the *only way not to get involved in 'fiction'*.<sup>362</sup>

Hence Strawson concludes "Indeed it is the only way to avoid committing oneself to the view that some 'fiction' or other is actually *true*."<sup>363</sup>

Hume is not a Strict Idealist (though, of course some might argue he *should* be). Identity could be one of the fictions that Hume chooses to tolerate or believe in. The generality (and ontological generosity) of the above quoted definition suggests that Hume does tolerate identity. However, this only commits him to tolerating fictions if *all* identity statements are fictions. The question turns once more upon whether Hume does or does not give identity

---

<sup>361</sup> See Butler's comments in Perry (1975a) p. 100-101. This is not to dispute the general influence of Butler on Hume.

<sup>362</sup> Strawson, G. (1989) p. 66

<sup>363</sup> Strawson, G. (1989) p. 66 Strawson would seem to be committed to saying Hume does not disparage fictions. It is one of the tasks of proponents of the New Hume to explain the many occasions when Hume clearly talks of fictions also being errors or fallacies.

one or more senses. Of course Hume may not distinguish between legitimate and illegitimate cases of identity. He may merely claim that all identities arise as a result of the imagination and there is nothing our enquiries can discover beyond this fact.

### III

#### THE CONTRASTING ROLES OF REASON AND THE IMAGINATION

Hume's account of identity first arises during his discussion of the existence of external body. Identity, therefore, is not Hume's prime subject. In fact Hume's prime concern is with the two competing explanations for our belief in the existence of external objects: the vulgar and the philosophical. The vulgar do not distinguish between internal perceptions and external objects. The philosophical system suggests a double existence in which perceptions represent objects. Hume devotes the closing passages of 'Of Scepticism with Regard to the Senses' to discussing this conflict.<sup>364</sup> If Hume's version of identity is forged in the white heat of a conflict it is wise to be clear about what fuels the disagreement.

Famously, Hume declares that whatever the defects of vulgar thinking it is nonetheless the practice of "the unthinking and unphilosophical part of mankind, (that is, all of us, at one time or other)".<sup>365</sup> However, "the philosophical system is found by experience to take hold of many minds"<sup>366</sup>, but "Nature is obstinate, and will not quit the field, however strongly attack'd by reason".<sup>367</sup> The conflict, though, is yet more complicated for the "*philosophical hypothesis has no primary recommendation, either to reason or the imagination*"<sup>368</sup>. It must therefore be the case "*that the philosophical system acquires all its influence on the imagination from the vulgar one*".<sup>369</sup>

The vulgar way of thinking is therefore prior to any conclusions reason may draw. The faculty of vulgar thinking is the imagination. Hume, recalling his discussion of mathematics,<sup>370</sup> notes that "the imagination, when set into any train of thinking, is apt to continue, even when its object fails it, and like a galley put in motion by the oars, carries on its course without any new impulse."<sup>371</sup> This impulse means "the mind... once in the train of observing an uniformity among objects, [...] naturally continues, till it renders the uniformity as complete as possible."<sup>372</sup>

---

<sup>364</sup> One third of the chapter's thirty pages. It is also the longest chapter in the *Treatise* by some margin. Such prolixity in as precise a writer as Hume is in itself indicative of crisis.

<sup>365</sup> *Treatise* 205:2-4 (I, IV, II)

<sup>366</sup> *Treatise* 213:23-24 (I, IV, II)

<sup>367</sup> *Treatise* 215:24-5 (I, IV, II)

<sup>368</sup> *Treatise* 212:3-5 (I, IV, II)

<sup>369</sup> *Treatise* 213:18-20 (I, IV, II)

<sup>370</sup> *Treatise* 198:14-15 (I, IV, II) citing (I, II, IV)

<sup>371</sup> *Treatise* 198:15-18 (I, IV, II)

<sup>372</sup> *Treatise* 198:27-30 (I, IV, II)

The vulgar way of thinking is therefore greatly different to philosophical thinking. As Hume believes identity arises as a psychological fiction, this cleave is worthy of investigation. He writes:

There is a great difference betwixt such opinions as we form after a calm and profound reflection, and such as we embrace by a kind of instinct or natural impulse, on account of their suitableness and conformity to the mind. If these opinions become contrary, 'tis not difficult to foresee which of them will have the advantage. As long as our attention is bent upon the subject, the philosophical and study'd principle may prevail; but the moment we relax our thoughts, nature will display herself, and draw us back to our former opinion. Nay she has sometimes such an influence, that she can stop our progress, even in the midst of our most profound reflections, and keep us from running on with all the consequences of any philosophical opinion. Thus tho' we clearly perceive the dependence and interruption of our perceptions, we stop short in our career, and never upon that account reject the notion of an independent and continu'd existence. That opinion has taken such deep root in the imagination, that 'tis impossible ever to eradicate it, nor will any strain'd metaphysical conviction of the dependence of our perceptions be sufficient for that purpose.<sup>373</sup>

Reason is at a double disadvantage. Firstly its conclusions are unable to dislodge those which have taken 'deep root in the imagination'. Secondly, the act of reasoning itself is prone to being corrupted and usurped by the imagination even while it takes place.<sup>374</sup>

Nonetheless reason does have a role and some power with which to *interrogate* the progress of the imagination. The imagination moves from impressions to belief with such ease and rapidity that it does not spontaneously reveal its own workings. Reason, however, can discern the imagination's machinations and therefore expose which of our beliefs are fictions.

Reason can interrogate imagination via reflection. Reflection is itself an act of reason and Hume appeals to the power of reflection to expose the imagination's fictions.<sup>375</sup> Reflection can achieve this interrogation because the imagination ascribes qualities that are not present

---

<sup>373</sup> *Treatise* 214:16-35 (I, IV, II) Compare also these remarks "Nothing is more dangerous to reason than the flights of the imagination, and nothing has been the occasion of more mistakes among philosophers. Men of bright fancies may in this respect be compar'd to those angels, whom the scripture represents as covering their eyes with their wings" *Treatise* 267:18-23 (I, IV, VII)

<sup>374</sup> Hume also distinguishes two types of imagination. The first's principles "are permanent, irresistable, and universal" *Treatise* 225:11-12 (I, IV, IV) the other "changeable, weak and irregular" 225:14-15 (I, IV, IV). The first are "the foundation of all of our thoughts and actions, so that upon their removal human nature must immediately perish and go to ruin" 225:16-18 (I, IV, IV) and therefore "received by philosophy" while "the latter are rejected" 225:24-25 (I, IV, IV). The weaker sort of the imagination is prone to extravagant explanation, e.g. explaining a disembodied voice in the dark as a ghost instead of merely an unseen human being 225:25-225:5 (I, IV, IV). It is the former sense of imagination that is Hume's object in this discussion.

<sup>375</sup> E.g. at *Treatise* 209:14-17 (I, IV, II), and the remark at 210:13-15 (I, IV, II) that "a very little reflection and philosophy is sufficient to make us perceive the fallacy of that opinion [of the imagination]."



in our perceptions. Reflection does not make this error because it only surveys the true contents of our consciousness. Hume takes consciousness to be transparent “For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear.”<sup>376</sup> The imagination being “permanent, irresistible, and universal”<sup>377</sup> the generation of fictions is constant. Yet reason can make the effort of reflection and bring about a hiatus long enough to realise that our perceptions do not justify our imaginative fictions.

This dual account is crucial to Hume. With it he can explain how our ideas are mostly fictions, yet also explain how we acquire various terms of rational enquiry. Broadly, Hume’s explanation is that we inevitably acquire some ideas through the natural action of the imagination and can then examine these using reason. Further, reason is itself aware of the transparency of consciousness. Knowing that perceptions ‘must necessarily appear in every particular what they are, and be what they appear’ reason can derive notions of what perceptions *would* cause the experiences *actually* caused by imaginative fictions. These notions, therefore, need never be encountered in experience for Hume to legitimately claim we can reason with them. Indeed, it might be the case that we could *never* experience them and yet they can play a role in our reasoning.

Unfortunately Hume nowhere explicitly claims this as his procedure. However, it does explain why Hume often appears to introduce terms out of order and offer definitions that purely empirical methods would not justify. When Hume is explaining how the imagination works he can presume all the conclusions of the imagination are already known. The imagination will already be working within any enquirer. If imagination attempted to examine itself, then circularity would inevitably follow. However, reason underwrites our examination of our beliefs, and reason, necessarily trailing in the imagination’s wake, has access to all our common beliefs. Adopting this procedure as Hume’s own allays the fears of those commentators who believe Hume frequently finds himself tugging his bootstraps or liberally helping himself to the tools of rationalism he supposedly rejects.<sup>378</sup>

Before preceding, though, a brief word of caution. By the time of the *Enquiry* Hume dramatically neuters his early account. There he merely asserts:

---

<sup>376</sup> *Treatise* 190:13-15 (I, IV, II)

<sup>377</sup> *Treatise* 225:11-12 (I, IV, IV)

<sup>378</sup> E.g. Stroud (1977) p. 100ff.

It seems evident, that... without any reasoning, or even almost before the use of reason, we always suppose an external universe, which depends not on our perception, but would exist, though we and every sensible creature were absent and annihilated. Even the animal creation are governed by a like opinion, and preserve this belief of external objects, in all their thoughts, designs and actions.<sup>379</sup>

The accompanying discussion of the conflict between reason and the imagination is entirely absent. It is natural to ask why Hume abandons the complexity of his earlier account. Hume begins this phase of his discussion with identical external objects and leads through to personal identity. This encompasses his discussion of the conflict between reason and the imagination and his discussion of identity. Famously Hume expresses reservations about his conclusions about personal identity in the Appendix. Arguably the account of personal identity is the accumulation of Hume's arbitration between reason and the imagination. If he became dissatisfied with the outcome, it is likely that he also became dissatisfied with the preceding argument. The final section will argue that Hume's dissatisfaction in the Appendix is with reason's inability to discover anything positive about personal identity. That is to say, with the failure of introspection. For present concerns, though, one must consider that Hume's self doubts cast a long shadow across these chapters. Although the procedure suggested above appears to be Hume's (and serves him well) his failure to explicitly proclaim it, combined with his later doubts, suggests it is not without problems. Nonetheless, I will proceed with it because it not only explains Hume's initial discussion, but also makes best sense of his later doubts. For if one's later dissatisfaction is with the conflict between reason and imagination what procedure is more likely to prompt this malaise than one that draws the conflict so starkly?

---

<sup>379</sup> *Enquiry* 151:21-29 (XII, I). Stroud also cites this passage, Stroud (1977) p. 110-1.

## IV

### HUME'S ANALYSIS OF THE IDENTITY OF BODY

Hume first notes "We may well ask, What causes induce us to believe in the existence of body? but 'tis vain to ask, Whether there be body or not?"<sup>380</sup> The proper question, then, is "Why we attribute a CONTINU'D existence to objects, even when they are not present to the senses; and why we suppose them to have an existence DISTINCT from the mind and perception."<sup>381</sup>

Only one solution appeases both demands

This hypothesis is the philosophical one of the double existence of perceptions and objects; which pleases our reason, in allowing, that our dependent perceptions are interrupted and different; and at the same time is agreeable to the imagination, in attributing a continu'd existence to something else, which we call *objects*. This philosophical system, therefore is the monstrous offspring of two principles, which are contrary to each other, which are both at once embrac'd by the mind, and which are unable mutually to destroy each other.<sup>382</sup>

Hume's description of the Doctrine of Double Existence as a 'monstrous offspring' strongly suggests his opinion of it is poor. The key issue presently is to discover to what extent Hume regards identity as being a partner in crime with this monstrous offspring.

Our senses "convey to us nothing but a single perception"<sup>383</sup> and "A single perception can never produce the idea of a double existence".<sup>384</sup> Instead, "fallacy or illusion" must construe any double existence.<sup>385</sup> If "the senses presented our impressions as external to, and independent of ourselves, both objects and ourselves must be obvious to our senses".<sup>386</sup> Consequently Hume is prompted to ask "how far we are *ourselves* the object of our senses."<sup>387</sup>

Hume is quick to assert that as the notion of person is "in common life... never very fix'd nor determinate. 'Tis absurd, therefore, to imagine the senses can ever distinguish betwixt

---

<sup>380</sup> *Treatise* 187: 25-7 (I, IV, II)

<sup>381</sup> *Treatise* 188:6-9 (I, IV, II)

<sup>382</sup> *Treatise* 215:6-15 (I, IV, II)

<sup>383</sup> *Treatise* 189:7 (I, IV, II)

<sup>384</sup> *Treatise* 189:9-10 (I, IV, II)

<sup>385</sup> *Treatise* 189:19 (I, IV, II) Here Hume does not employ the word 'fiction'. However, he appears to use 'fallacy' and 'illusion' as simple cognates.

<sup>386</sup> *Treatise* 189:24-8 (I, IV, II)

<sup>387</sup> *Treatise* 189:28-9 (I, IV, II)

ourselves and external objects.”<sup>388</sup> Hume’s reasoning seems less than forceful. An unclear notion of something does not seem to preclude distinguishing it. Something’s being a mere something would suffice. Admittedly such opacity would be unsatisfactory in an account of personal identity, but this is not the objection that Hume makes. His objection only carries weight when backed by his own version of the Transparency Thesis. Hume insists that

[S]ince all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Every thing that enters the mind, being in *reality* as the perception, ‘tis impossible any thing shou’d to *feeling* appear different. This were to suppose, that even where we are most intimately conscious, we might be mistaken.<sup>389</sup>

One consequence of this is to dispense with one of the classic objections to Hume’s theory of personal identity. If Hume is taken to hold a bundle theory of personal identity, then he has to explain how we know where one bundle ends and another begins. However, Hume doesn’t have to answer this question. He boldly asserts that we cannot distinguish between perceptions of ourselves and perceptions of the external world within our own perceptual arrays. This would imply that if one’s perceptions were of another person one could not *by principle of reason alone* tell where the other person began and one’s own self ended. Clearly, it can be objected that this is nonsense. We do know where we end and others begin. Hume, however, is not saying that we do not regularly make this judgement (that would be absurd). He is instead claiming that such judgements about persons are governed by custom and habit as much as are judgements about objects. Appreciating this will be essential to understanding the later account of personal identity.

Hume does seem to make an almost instant slip and disregard his own doctrine here, when he notes that “[We] think an object has a sufficient reality, when its Being is uninterrupted, and independent of the incessant revolutions, which we are conscious of in ourselves.”<sup>390</sup> If this were Hume’s sincere opinion this would count severely against the interpretation above. The ‘incessant revolution’ can only be of perceptions. If we are aware of this being ‘in ourselves’, i.e. internal, then surely we *can* distinguish ourselves from external existences.

---

<sup>388</sup> *Treatise* 189:35-200:3 (I, IV, II)

<sup>389</sup> *Treatise* 190:13-19 (I, IV, II) Hume recaps in the Appendix “As long as we confine our speculations to *the appearances* of objects to our senses, without entering into disquisitions concerning their real nature and operations, we are safe from all difficulties” *Treatise* 638:18-21 (App.)

<sup>390</sup> *Treatise* 191:26-9 (I, IV, II)

There is, however, no slip. The use of 'ourselves' here is merely reflexive. As we do not experience the divide between internal perceptions and external objects we can only be said to experience perceptions as being internal. It could be objected that Hume is not entitled to a notion of internal without the collateral notion of the external. However, the Transparency Thesis deals with this difficulty. Perceptions necessarily present themselves as parts of our minds if we experience them at all. All Hume is committed to saying is that we are transparently aware of the coming and going of our own perceptions. The use of 'ourselves' carries no ontological weight. Again this seems to preclude Hume's supposed bundling difficulty. Experience of a perception must be complete and transparent. Further as perceptions are "perishing existences"<sup>391</sup> there can be no problem of possible co-ownership. A perception comes into existence when it becomes part of a pre-existing bundle (a mind) and subsequently perishes.<sup>392</sup> The numerically same perception cannot, therefore, migrate to another bundle.<sup>393</sup> Nor could it belong to two minds simultaneously for its brief perishing existence. If this happened the two minds would have to be transparent to each other, because each mind always has complete knowledge of every perception currently in its bundle. There could only be two minds if they were qualitatively identical and superimposed (for want of a better phrase) upon each other. Parsimony tells against such an elaborate objection. Such minds are numerically identical.<sup>394</sup>

Returning to present concerns, though, Hume has to reconstruct the distinction we do experience between internal and external. Hume has denied this is readily present in our perceptions as awareness of our own nature. Instead, Hume continues that "as far as the senses are judges, all perceptions are the same in the manner of their existence."<sup>395</sup> Because all perceptions are equal we are unable to make an effective division between internal perceptions and external objects. This leaves reason (and philosophy) in radical disagreement with vulgar practice,

For philosophy informs us, that every thing, which appears to the mind, is nothing but a perception, and is interrupted, and dependent on the mind; whereas the vulgar confound perceptions and objects, and attribute a distinct continu'd existence to the very things they feel or see. This sentiment, then, as

---

<sup>391</sup> *Treatise* 194:1 (I, IV, II)

<sup>392</sup> It will not be discussed here how long a perishing thing exists when compared with an instantaneous thing, or momentary thing.

<sup>393</sup> Presuming its perishing career is very short.

<sup>394</sup> Anscombe famously objects to the *cogito* that it offers no guarantee that each of us is *not* many minds in union.

<sup>395</sup> *Treatise* 193:3-4 (I, IV, II)

it is entirely unreasonable, must proceed from some other faculty than the understanding.<sup>396</sup>

It is, of course, the Imagination that Hume identifies as the source of these beliefs.

Although “all impressions are internal and perishing”<sup>397</sup> we are still prone to regard some as having external existence. Hume asserts this is because impressions of objects like “mountains, and houses, and trees”<sup>398</sup> exhibit “COHERENCE and CONSTANCY”<sup>399</sup>. In contrast “Our passions are found by experience to have a mutual connexion with and dependence on each other, but on no occasion is it necessary to suppose, that they have existed and operated, when they were not perceiv’d”.<sup>400</sup> This, of course, would follow from the Transparency Thesis. An individual pain does not exhibit the sort of coherence and constancy that might prompt one to think the pain needs to exist beyond one’s experience of it. Our complete knowledge of it leads us to believe this. Conversely, impressions of constant and coherent external objects by virtue of their transparency feel incomplete. This produces the propensity to believe that coherence and constancy are indicative of a continued existence.

Hume is, not unexpectedly, circumspect about this propensity.

But ‘tis evident, that whenever we infer the continu’d existence of the objects of sense from their coherence, and the frequency of their union, ‘tis in order to bestow on the objects a greater regularity than what is observ’d in our mere perceptions.<sup>401</sup>

This obviously follows from what Hume has said before. All perceptions are on an equal footing. Therefore the mind cannot distinguish which are of external objects from the contents of any individual perception. Any distinction must therefore arise from the perceptions *in combination*. Sensations do not exhibit the same coherence and constancy as impressions of external objects. This coherence belongs to external objects only.

Stroud suggests Hume is careless in his discussion of how coherence and constancy ground the shift in belief to identical external objects.<sup>402</sup> Hume talks about “mountains, and houses,

---

<sup>396</sup> *Treatise* 193:17-24 (I, IV, II)

<sup>397</sup> *Treatise* 194:1 (I, IV, II)

<sup>398</sup> *Treatise* 194:32 (I, IV, II)

<sup>399</sup> *Treatise* 195:22 (I, IV, II)

<sup>400</sup> *Treatise* 195:29-33 (I, IV, II)

<sup>401</sup> *Treatise* 197:31-35 (I, IV, II)

<sup>402</sup> Stroud (1977) p. 100 *ff.* Stroud dismisses H.H. Price’s attempt to explain Hume’s project in terms of a naturalism about our material object terms (p. 109). This is surely correct. Hume is constructing a pathology

and trees”<sup>403</sup> *before* he has established how we believe in external objects. As argued above, Hume is perfectly entitled to make use of these concepts because they are universally introduced by the imagination. Hume is in effect claiming that our imagination acts on some of our perceptions and not on others. The difference he finds between perceptions of external objects and sensations is that in the former the imagination can seize upon coherence and then constancy.<sup>404</sup>

It is the imagination that provides what cannot be found in our perceptions. Before going into the discussion of identity Hume gives a preliminary account. As already noted, Hume says “the imagination, when set into any train of thinking, is apt to continue, even when its object fails it, and like a galley put in motion by the oars, carries on its course without any new impulse.”<sup>405</sup> This impulse means “the mind... once in the train of observing an uniformity among objects, [...] naturally continues, till it renders the uniformity as compleat as possible.”<sup>406</sup>

If this were the end of Hume’s discussion he would save himself a great deal of controversy. However, the coherence of our perceptions even when completed by the imagination is “too weak to support alone so vast an edifice, as is that of the continu’d existence of all external bodies”<sup>407</sup>. We must have regard for the constancy of our impressions and this is where Hume’s difficulties begin:

When we have been accustom’d to observe a constancy in certain impressions... we are not apt to regard these interrupted perceptions as different, (which they really are) but on the contrary consider them to be individually the same, upon account of their resemblance. But as this interruption of their existence is contrary to their perfect identity, and makes us regard the first impression as annihilated, and the second as newly created, we find ourselves somewhat at a loss, and are involv’d in a kind of contradiction.<sup>408</sup>

---

of *thought* not language. He is not asserting that we *talk* about objects as if they are as real as our perceptions, but *think* about them in this way.

<sup>403</sup> *Treatise* 194:32 (I, IV, II)

<sup>404</sup> It could be objected that some sensations, e.g. a long-term pain, are as coherent and constant as any external object, indeed more so. One might ask if Hume’s distinction on the grounds of coherence and constancy is one of kind or degree. Evidence favours degree, as the imagination can smooth over many different interruptions and irregularities. It might also allow that certain coherent and constant perceptions must themselves be internal, such as long-term pains.

<sup>405</sup> *Treatise* 198:15-18 (I, IV, II)

<sup>406</sup> *Treatise* 198:27-30 (I, IV, II)

<sup>407</sup> *Treatise* 198:35-199:1 (I, IV, II)

<sup>408</sup> *Treatise* 199:13-24 (I, IV, II)

The solution to this problem comes by “supposing that these interrupted perceptions are connected by a real existence, of which we are insensible.”<sup>409</sup>

Stroud remarks that “Hume admits that ‘coherence’ plays only a supplementary role”<sup>410</sup> to constancy in generating this fiction. However, Hume does not, as Stroud suggests, move so casually here. One must ask why Hume starts his discussion of identity with a discussion of constancy and coherence. Hume clearly thinks that constancy and coherence are the foundation of our belief in permanent external objects. However, he does not think this is the entire story, for coherence and constancy have a ‘vast edifice’ built upon them. Central to this edifice is identity. So the question arises why Hume thinks identity differs from the belief in external objects merely based on the coherence and constancy of our impressions.

Hume’s project in these passages is to sketch a progression from our impressions to the belief in external objects.<sup>411</sup> In this coherence is not supplementary, but primary. It is the coherence of our impressions with which the imagination first engages. The imagination’s first step is only a small one. Our impressions appear to exhibit a predictability and reliability, and this coherence prompts the belief that something lies behind this coherence. Alone coherence would only prompt the notion of an ordered world. This Hume can exercise a liberal tolerance towards, for surely the world is ordered. The imagination’s only role here is to make us prone to ignore inconsistency and dwell more on consistency.

Constancy gains its role *prompted* by coherence. Constancy is a stronger condition than coherence. To say the world is coherent does not limit the amount of change in the world. Coherence is only a probable judgement that perceptions are more likely to be the same in the future. Constancy is the claim that they will be the same. This is a step further towards the belief in external objects as the cause of our perceptions. This is why Stroud is wrong to suggest Hume gives coherence only a supplementary role. Coherence is the most easily and naturally formed opinion about our perceptions. Our perceptions are surely less constant than they are coherent. Constancy, therefore, can only provide a foothold for the imagination once coherence has provided a step-up.

This does not, however, explain why Hume thinks the imagination carries us even further into the vast edifice of identical external objects. Part of the explanation is Hume’s above

---

<sup>409</sup> *Treatise* 199:26-28 (I, IV, II)

<sup>410</sup> Stroud (1977) p. 100

<sup>411</sup> In the *Enquiry* the entire detail of this project is (sadly) absent. There Hume merely remarks “It is evident that there is a principle of connexion between different thoughts or ideas of the mind, and that, in their

cited contention that the imagination moves inevitably towards its conclusion once set under way. To this Hume adds the claim that constancy gives us the notion of objects' Perfect Identity. Although not spelt out, Hume commits himself to another progression. If our view of an object is interrupted we imagine the broken periods of observation to be not only qualitatively identical, but numerically identical. Hume suggests we do not do this by any inference, but naturally. Hence philosophy and vulgar experience find themselves in inevitable conflict. Reason can only tell us that interrupted perceptions are distinct existences. Yet naturally we confound qualitative identity with numerical identity.

Hume's account would be simpler if he merely asserted that we regularly confused qualitative identity with numerical identity. However, this would leave it unexplained why we make this mistake so regularly, and also not explain how we acquire the notion of perfect identity. Hume's account would also be simpler if he merely remarked that this mistake was caused by errors generated by the coherence and constancy we observe in our perceptions. However, this would not explain the edifice we build upon them, or the errors of ancient and modern philosophy Hume attacks.

Hume summarises the problem as follows

When we have been accustom'd to observe a constancy in certain impressions, [...], we are not apt to regard these interrupted perceptions as different, (which they really are) but on the contrary consider them as individually the same, upon account of their resemblance. But as this interruption of their existence is contrary to their perfect identity, and makes us regard the first impression as annihilated, and the second as newly created, we find ourselves somewhat at a loss, and are involv'd in a kind of contradiction. In order to free ourselves from this difficulty, we disguise, as much as possible, the interruption, or rather remove it entirely, by supposing that these interrupted perceptions are connected by real existence, of which we are insensible.<sup>412</sup>

It is this supposition of a 'real existence, of which we are insensible' that prompts Hume's further enquiry. Hume offers this plan of analysis:

In order to justify this system, there are four things requisite. *First*, To explain the *principium individuationis*, or principle of identity. *Secondly*, Give a reason, why the resemblance of our broken and interrupted perceptions induces us to attribute an identity to them. *Thirdly*, Account for that propensity, which this illusion gives, to unite these broken appearances by a continu'd existence.

---

appearance to the memory or imagination, they introduce each other with a certain degree of method and regularity." *Enquiry* 23:1-5 (III)

<sup>412</sup> *Treatise* 199:13-14... 17-28 (I, IV, II)

*Fourthly* and lastly, Explain that force and vivacity of conception, which arises from the propensity.<sup>413</sup>

As this is Hume's procedure it will be followed through here.

---

<sup>413</sup> *Treatise* 199:35-200:7 (I, IV, II)

## HUME'S ACCOUNT OF THE *PRINCIPIUM INDIVIDUATIONIS*

Hume's establishing of the notion of identity has interminably puzzled critics. Stroud is typical when he remarks:

It is difficult, to say the least, to see how [Hume] explains how we originally get the idea of the identity of an object through time. We get it, Hume says, by 'conceiving' of a moment of time and an object existent at that time, and then 'imagining' a change in the time without any variation or interruption in the object. But how are we able to 'imagine' such a thing unless we already have the idea of the invariableness and uninterruptedness of an object through time?<sup>414</sup>

This confusion can be dissolved in the manner previously sketched.<sup>415</sup> The imagination's universally provided experience of the fiction of identity is investigated by reason. The relation between imagination and reason is complex here, and consequently of considerable interest. It is worth, therefore, investigating these passages very closely.

Hume begins by dismissing two possible ways by which we might acquire the idea of identity.<sup>416</sup> Consideration of a single object delivers only the idea of *unity*,<sup>417</sup> while a multiplicity of objects can only deliver the idea of *diversity*.<sup>418</sup> The second of these is uncontroversial. However, Hume's imposition of unity over self-identity is obviously controversial. Curious as self-identity might seem, it is not a principle of logic Hume can simply dismiss. He can, of course, dismiss conclusions assumed to follow from the proposition that every thing is identical with itself. If this is what he does then one should be able to extract what Hume's worries are. Otherwise Hume has to suffer the undignified suggestion that he miscomprehends a simple principle of logic.

Hume's criticism has two parts. Firstly,

---

<sup>414</sup> Stroud (1977) p. 103

<sup>415</sup> I therefore find myself in deep disagreement with Hirsch's attempts to establish that Hume argues that the idea of identity is found by the senses. See Hirsch (1983).

<sup>416</sup> Hume writes "we may observe" 200:8 (I, IV, II) which I understand to entail an act of clear perception guided only by reason. That is to say the complete and transparent consideration of the "view of any one object" 200:9 (I, IV, II) uninterfered with by imagination cannot give the idea of identity, only unity. Equally, when Hume writes when discussing a multiplicity of objects "The mind always pronounces the one not to be the other" 200:19 (I, IV, II) I take 'mind' to be a synonym for reason. Hume cannot mean mind in a broader sense that would include the imagination. If he did the imagination would proceed to confuse a resembling multiplicity with an identity.

<sup>417</sup> *Treatise* 200:8-16 (I, IV, II)

<sup>418</sup> *Treatise* 200:17-22 (I, IV, II)

For in that proposition, *an object is the same with itself*, if the idea express'd by the word, *object*, were no ways distinguish'd from that meant by *itself*, we really shou'd mean nothing[.]<sup>419</sup>

And

[N]or wou'd the proposition contain a predicate and a subject, which however are imply'd in this affirmation.<sup>420</sup>

The latter grammatical objection is of little interest. The former objection is where Hume's opinion really lies.

Hume regards identity as a relation. Therefore it must relate two things to have any significance. Hume objects that one object cannot stand in a relation with itself. Either there is a way to distinguish the two parts of a relation, or there is not. If there is no way of making a distinction, then only one thing can be present, and the idea given by this is merely unity.

Hume would be guilty of a bad error if he claims there are no reflexive relations. Hume, however, is concerned here not with metaphysics, but with the psychology of reasoning about identity. This is entirely proper and consistent with his project to use reason to interrogate the imagination. Self-identity is not under trial as an abstract metaphysical principle *per se*. Hume is claiming reason cannot extrapolate from a single perception the idea of something related. The construction of relation requires the imagination.

It is no accident that Hume considers "the view of any one object"<sup>421</sup>. It is our ability to *picture* in our mind's eye that is of interest to Hume. Recall, for purposes of comparison, Hume's comments on the difficulty of imagining something to be infinitely divisible. Hume draws our attention to the way we are obliged to picture the divisions of a grain of sand:

When you tell me of the thousandth and ten thousandth part of a grain of sand, I have a distinct idea of these numbers and of their different proportions; but the images, which I form in my mind to represent the things themselves, are

---

<sup>419</sup> *Treatise* 200:10-13 (I, IV, II) Hence Hume's later comment that "We cannot, in any propriety of speech, say, that an object is the same with itself, unless we mean, that the object existent at one time is the same with itself existent at another." *Treatise* 201:22-25

<sup>420</sup> *Treatise* 200:13-15 (I, IV, II)

<sup>421</sup> *Treatise* 200:9 (I, IV, II)

nothing different from each other, nor inferior to that image, by which I represent the grain of sand itself, which is suppos'd to so vastly exceed them.<sup>422</sup>

Hume's objection to self-identity has the same basis. When we think of a thing being the same as something we have to picture *two* things to ask ourselves if they are related in some way. There is no way of picturing in our mind one thing in a relation with itself.

There is something temptingly correct about what Hume says when considered this way. We have no way of comprehending self-identity even (or especially) as an abstract metaphysical principle. It is not natural to state, and can only be imagined in the most slippery of ways.<sup>423</sup>

Instead Hume claims that "One single object conveys the idea of unity, not that of identity."<sup>424</sup> Unity, as Hume uses it, is meant to exclude identity. Unity is to be 'one' while identity is to be 'one and the same'. This exclusion, though, is only justified on Hume's own rejection of self-identity. Hume insists that identity demands diversity yet purports to be about what is in fact unity. Understood this way Hume's attack makes clear sense. He is claiming that the perceptions we have of objects do not reveal objects' metaphysical constitutions. This means we do not know how objects remain identical (even synchronically).

Even if Hume succeeds in offering psychological reasons to separate Unity from Synchronic Identity it can still be the case that Humean Unity is no different from Synchronic Identity otherwise understood. Obviously each perception is identical with itself. Whatever interpretation one chooses one must accept that. However, Hume's dispute is not with this claim stated as a bald rule of metaphysics. His argument is that we have no way of reasoning about any perception such that we can know *about that perception* that it is identical with itself. To know this, the fact of a perception's self-identity would have to be contained within each perception. Hume can find no principle present in each, just as he cannot find the distinction between the internal and external, or the perception and ourselves.

---

<sup>422</sup> *Treatise* 27:14-20 (I, II, I) when Hume here says 'nothing different' he of course means to suggest that the two images are qualitative exact similar not that they are numerical identical.

<sup>423</sup> Of course, the fact that we must *imagine* it is exactly Hume's point, i.e. exercise the imagination. Hume's distinction between reason driven picturing of states of affairs, and imaginings is an interesting one. Hume might suggest that reason can, temporally at least, enslave the imagination to do reason's bidding. Alternatively he could be suggesting that reason has its own ability to picture things in the mind's eye, in parallel to the imagination. Finally, he could rely on *imaging* being separate from both reason and the imagination and employed by both. Settling such details is beyond the scope of the present work.

<sup>424</sup> *Treatise* 200:15-16 (I, IV, II)

Further, knowing a perception to be self-identical would imply knowing something of perceptions' true nature. If this were possible we could know something of objects' true nature. And it is just this kind of knowledge that Hume denies we have. Hume's sceptical realism means he must be neutral about what perceptions are. They might be individual existences (as self-identity would imply) or they might be dependent parts of a causally ordered system. As it is Hume's claim that we cannot settle this issue he must suppress the possibility of self-identity while continuing to explain how perceptions are individuated. Unity does just this without having recourse to the metaphysical certainties to which self-identity pretends.<sup>425</sup>

This is certainly Hume's conclusion by the time of the *Enquiry*. There the discussion of unity and diversity is absent, and Hume only discusses the senses as a possible source of knowledge of objects. Nonetheless, he is scathing:

In vain do you pretend to have learned the nature of bodies from your past experience. Their secret nature, and consequently all their effects and influence, may change, without any change in their sensible qualities. This happens sometimes, and with regard to some objects: Why may it not happen always, and with regard to all objects?<sup>426</sup>

Likewise it would seem reasonable to conclude that unity cannot be construed as discovering the secret nature of objects, even when buttressed with the metaphysical principle of self-identity.

What is required to explain identity is how "we make a difference, betwixt the idea meant by the word, *object*, and that meant by *itself*, without going the length of number, and at the same time without restraining ourselves to strict and absolute unity."<sup>427</sup> The answer will be found in "recourse to the idea of time or duration."<sup>428</sup> Of course, Hume's theory of time is

---

<sup>425</sup> Hume appears to believe that simples can be defined psychologically rather than metaphysically. As his simples are perceptions the minimal unit of perception delimits what counts as a simple. This is supposedly a unity, but this projection seems unjust. The minimal unit of perception is variable and the amount of the world it individuates variable too. Hence whatever part of the world a moment's perception might delimit is not a unity in any independent sense. Hume's assumption that perception can provide these units not only leaves his metaphysics curiously grounded, but seems to lead to a collapse into idealism. This problem is made worse by Hume's habit of appearing to speak of perceptions as having duration, or treating as one perception what he elsewhere treats as a succession.

<sup>426</sup> *Enquiry* 38:9-15 (IV, II)

<sup>427</sup> *Treatise* 201:25-28 (I, IV, II). Hume initially remarks that "at first sight this seems utterly impossible" 200:25-6 (I, IV, II). This is one of many occasions when reason is implied by the metaphor of vision. It would be an interesting discussion for another time to tally Hume's comparisons between reason and having a clear view of a subject. Hume's choice of language is in part influenced by his pictorial theory of the imagination.

<sup>428</sup> *Treatise* 200:32-33 (I, IV, II)

not without its own complications. It would be an intolerable diversion to fully examine Hume's account of time here.<sup>429</sup> However, a brief sketch is nonetheless necessary.

A central Humean claim is "*That all our simple ideas in their first appearance are deriv'd from simple impressions*".<sup>430</sup> The honourable exception to this is the 'missing shade of blue'.<sup>431</sup> Yet it is clear that Hume maintains that some ideas are obtained without first being impressions. Of our idea of time, Hume says:

Five notes play'd on a flute give us the impression and idea of time; tho' time be not a sixth impression, which presents itself to the hearing or any other of the senses. Nor is it a sixth impression, which the mind by reflection finds in itself.<sup>432</sup>

Time, Hume concludes "since it appears not as any primary distinct impression, can plainly be nothing but different ideas, or impressions, or objects dispos'd in a certain manner, that is, succeeding each other."<sup>433</sup>

Similarly the idea of extension is a *compound* idea, abstracted from the senses of touch and sight.<sup>434</sup> Extension is derived from "several lesser impressions"<sup>435</sup> and as such no simple idea of this exists. Hume has to regard space and time as not being capable of simple perception. This is because he wants to claim that space and time do not exist apart from the objects that exist in space and time. Hence

We have therefore no idea of space or extension, but when we regard it as an object either of our sight or feeling.

The same reasoning will prove, that the indivisible moments of time must be fill'd with some real object or existence, whose succession forms the duration, and makes it be conceivable by the mind.<sup>436</sup>

This latter statement about "a real object ... whose succession forms the duration"<sup>437</sup> needs close analysis. In a passage just before this Hume states his opinion concerning duration. It is held by both the vulgar and philosophers that "the idea of duration is applicable in a

---

<sup>429</sup> Those who wish to be so diverted will be ably engaged by Baxter (1987)

<sup>430</sup> *Treatise* 4:19-21 (I, I, I)

<sup>431</sup> *Treatise* 5:32-6:28 (I, I, II)

<sup>432</sup> *Treatise* 36:31-35 (I, II, III)

<sup>433</sup> *Treatise* 37:13-16 (I, II, III)

<sup>434</sup> *Treatise* 38:6-39:15 (I, II, III)

<sup>435</sup> *Treatise* 38:28 (I, II, III)

<sup>436</sup> *Treatise* 39:9-11,12-15 (I, II, III)

<sup>437</sup> *Treatise* 39:13-14 (I, II, III)

proper sense to objects, which are perfectly unchangeable".<sup>438</sup> However the idea of duration "can never be convey'd to the mind by any thing stedfast and unchangeable."<sup>439</sup> Hume refers the reader to elsewhere in the *Treatise* to discover the source of this fiction.<sup>440</sup> There Hume suggests that there is a "continual succession of perceptions in our mind; so that the idea of time [is] ever present with us".<sup>441</sup> When we compare this succession with an object that has remained unchanged we are apt to assume that it has, despite appearances, undergone a succession of changes as well.<sup>442</sup>

Hume appears to link object identity to survival of change. Seen in Humean terms, this is the linking of a diversity of units into a succession. This also gives us the notion of time. Of course, to modern eyes this is arguably a perfectly good notion of identity. Yet Hume finds this account to be antagonistic to the notion of identity. This is because Hume takes identity to imply simplicity and immutability. Hume's target is substance and the soul, the twin conceits of both the ancient philosophy and modern rationalism. In both these systems identity is underwritten by something that supports change but itself remains immutable.<sup>443</sup>

Hume sums up his own argument, writing "I have already observ'd (I, II, V), that time, in a strict sense, implies succession, and that when we apply its idea to any unchangeable object, 'tis only by a fiction of the imagination, by which the unchangeable object is suppos'd to participate of the changes of co-existent objects, and in particular of that of our perceptions."<sup>444</sup> Hume continues: "This fiction of the imagination almost universally takes place; and 'tis by means of it, that a single object, plac'd before us, and survey'd for any

---

<sup>438</sup> *Treatise* 37:17-19 (I, II, III)

<sup>439</sup> *Treatise* 37:23-25 (I, II, III)

<sup>440</sup> *Treatise* 37n1 (I, II, III) refers the reader to *Treatise* 65:11-31 (I, II, V)

<sup>441</sup> *Treatise* 65:15-17 (I, II, V)

<sup>442</sup> This separates Hume from contemporary four-dimensional ontologists. The march of time is uniform in these ontologies, and an unchanging object is in fact still made up of a succession of exact similar temporal stages. Hume, however, insists on time progressing differently for different objects. If there is any part of Hume's philosophy that does lead him into a labyrinth and should have led him into despair then it is this. One can have sympathy with the idea that change is essential to a genuine notion of objecthood, and Hume goes some way to recognising this. Hume's position could be saved if he made a Cartesian manoeuvre and insisted that all objects are in fact always undergoing change, just as Descartes finds himself obliged to insist that minds are always thinking. However, Hume is not prepared to make any such move.

<sup>443</sup> Hume is arguably right to reject the notion of any sort of object (mental or physical) that does not have parts. A simple substance that nonetheless changes by having different qualities, such as a Cartesian soul, Hume finds nonsensical. To have different qualities is to have different parts which can become differently arranged. To have parts is not to be simple.

<sup>444</sup> *Treatise* 200:33-201:3 (I, IV, II) Those who argue Hume distinguishes between strict and loose (or perfect and imperfect) identity, might draw support from an apparently similar use of 'strict' here. However, I am aware of no-one doing so. The distinction is most often argued for on grounds of philosophical neatness, and not textual evidence. Famously, the textual evidence is slim, Hume only using the phrase 'imperfect identity' once (at *Treatise* 256:24) and even this use is queried as an error by Selby-Bigge. Of this, more later.

time without our discovering in it any interruption or variation is able to give us a notion of identity.”<sup>445</sup>

This fiction of a single unchanging object allows the further fiction of identity to flourish. The imagination suggests two scenarios. In the first we consider the object simultaneously at two times.<sup>446</sup> In the second we imagine the object remaining steadfast while time changes.<sup>447</sup> Waxman urges that “despite what one may be tempted to suppose, both viewpoints are necessary to acquiring the idea of identity.”<sup>448</sup> Indeed we do need both so we might obtain an idea between unity and number. The first gives us the notion of number, essential to relation, the second preserves unity.

What has not been commented on is the importance in Hume’s discussion of being able to *picture* these events. The first is suggestive of imagining two qualitatively identical objects, one of which is mentally tagged as earlier than the other. The second scenario suggests picturing an object and supposing it unchanging while time passes (and therefore change occurs). This ties with Hume’s earlier suspicions of self-identity based on our inability to picture something as self-identical in our mind’s eye. Hence Hume concludes “Here then is an idea, which is a medium betwixt unity and number; or more properly speaking, is either of them, *according to the view*, in which we take it”.<sup>449</sup>

Hume concludes:

Thus the principle of individuation is nothing but the *invariableness* and *uninterruptedness* of any object, thro’ a suppos’d variation of time, by which the mind can trace it in different periods of its existence, without any break of the view, and without being oblig’d to form the idea of multiplicity or number.<sup>450</sup>

Hume’s repeated reference to ‘a suppos’d variation of time’ reminds us that Perfect Identity rests on an act of the imagination. Perfect Identity, therefore, is a fiction.

---

<sup>445</sup> *Treatise* 201:3-7 (I, IV, II) The strategy, adopted on Hume’s behalf in this analysis, of presuming imagination’s actions always to be complete for reason’s examination finds support here in Hume’s comment that ‘This fiction of the imagination almost universally takes place’. The qualifying ‘almost’ detracts nothing. Hume is merely allowing for the possibility that reason *can* intervene to avoid the error. This, however, involves mental effort. Hence Hume writes ‘without our discovering in it any interruption or variation’. The fiction occurs *unless* we sufficiently exercise reason to prevent the imagination suffocating us with persuasive and effortless fictions.

<sup>446</sup> *Treatise* 201:8-13 (I, IV, II)

<sup>447</sup> *Treatise* 201:13-18 (I, IV, II)

<sup>448</sup> Waxman (1994) p. 208

<sup>449</sup> *Treatise* 201:18-21 (I, IV, II) Emphasis added.

<sup>450</sup> *Treatise* 201:29-34 (I, IV, II)

This seems clear enough, but the fact seems lost on many commentators. This is particularly the case with those commentators who assume Hume makes a distinction between Perfect and Imperfect Identity (this will be further discussed in the next section). These writers assume the title of fiction attaches only to Imperfect Identity, and that Perfect Identity is somehow more legitimate or respectable. The passages that have most prompted these interpretations are found later in Hume's discussion of personal identity. In these passages Hume is discussing how we come to attribute identity to items that are radically variable and interrupted (rebuilt buildings, rivers). These cases clearly differ from those where identity is attributed to simple, stable objects like hats and stones. However, it will not do to assume that Hume thinks there is a difference not just of degree, but of kind, here. This will be returned to shortly, once Hume's conclusions regarding the emergence of our belief in body have been dealt with. For the moment, though, it should be noted that Hume merely describes how we most easily obtain the notion of a Perfect Identity. Nowhere does he suggest we experience Perfect Identities, nor that the experience of Perfect Identity is in anyway privileged, or any less a fiction than other productions of the imagination.

The second part of Hume's system attempts to "shew why the constancy of our perceptions makes us ascribe to them a perfect numerical identity, tho' there be very long intervals betwixt their appearance, and they have only one of the essential qualities of identity, viz. *Invariableness*."<sup>431</sup> Hume has already hinted at the reason when recapping the definition of identity, which occurs on the occasions when the mind can act 'without being oblig'd to form the idea of multiplicity or number'. The imagination is a lazy faculty, choosing always the easiest view of the world.<sup>432</sup> Hence, the less likely a succession of perceptions is to oblige one to observe a difference the less likely the mind is to comprehend one. Hume subsequently makes this claim in yet stronger terms, saying "The mind readily passes from one [resembling perception] to the other, and perceives not the change without a strict attention, of which, generally speaking, 'tis wholly incapable."<sup>433</sup>

It is, of course, "a succession of related objects [that] places the mind in this disposition, and is consider'd with the same smooth and uninterrupted progress of the imagination, as attends the view of the same invariable object."<sup>434</sup> Consequently "The thought slides along the succession with equal facility, as if it consider'd only one object; and therefore

---

<sup>431</sup> *Treatise* 201:36-202:4 (I, IV, II)

<sup>432</sup> "The faculties of the mind repose themselves... and take no more exercise, than what is necessary to continue that idea" *Treatise* 203:20-22 (I, IV, II), See *Treatise* 203:12-206:25 in general.

<sup>433</sup> *Treatise* 203:8-11 (I, IV, II)

<sup>434</sup> *Treatise* 204:2-5 (I, IV, II)

confounds the succession with the identity.”<sup>455</sup> In a footnote<sup>456</sup> Hume points out that the resemblance is two-fold. Individual perceptions resemble each other, and so the act of surveying them feels the same. Also, the act of surveying a succession itself resembles that of surveying an unchanging object. The mind, therefore, is doubly confounded finding both the act and its contents resembling. Finally, Hume says

The smooth passage of the imagination along the ideas of the resembling perceptions makes us ascribe to them a perfect identity. The interrupted manner of their appearance makes us consider them as so many resembling, but still distinct beings, which appear after certain intervals. The perplexity arising from this contradiction produces a propension to unite these broken appearances by the fiction of a continu’d existence, which is the *third* part of that hypothesis I propos’d to explain.<sup>457</sup>

Although Hume is trying to explain the emergence of a fiction his aim is not entirely sceptical. He feels moved to remind his reader that:

We may begin with observing, that the difficulty in the present case is not concerning the matter of fact, or whether the mind forms such a conclusion concerning the continu’d existence of its perceptions, but only concerning the manner in which the conclusion is form’d, and principles from which it is deriv’d.<sup>458</sup>

The proper question, then, is “*First*, How we can satisfy ourselves in supposing a perception to be absent from the mind without being annihilated”<sup>459</sup> and also “*Secondly*, After what manner we conceive an object to become present to the mind, without some new creation of a perception or image”.<sup>460</sup>

Hume’s answer is simple. The mind “is nothing but a heap or collection of different perceptions, united together by certain relations”.<sup>461</sup> Each perception is distinguishable and therefore separable from this collection.<sup>462</sup> Equally, each perception can come to be related

---

<sup>455</sup> *Treatise* 204:14-16 (I, IV, II) Hume comments directly after that “We shall afterwards see many instances of this tendency of relation to make us subscribe an identity to different objects; but shall here confine ourselves to the present subject” *Treatise* 204:17-19 (I, IV, II). I take this to refer to the renewed discussion of more exotic cases of identity in ‘*Of Personal Identity*’ (see *Treatise* 253:15-258:36 (I, IV, VI)) Hume explicit directs his reader to these passages at 206:25n1. There we will “learn [the principles that explain] how the interruption in the appearance of a perception implies not necessarily an interruption in its existence” *Treatise* 206:21-23 (I, IV, II)

<sup>456</sup> *Treatise* 204:n1 (I, IV, II)

<sup>457</sup> *Treatise* 205:13-21 (I, IV, II)

<sup>458</sup> *Treatise* 206:26-31 (I, IV, II)

<sup>459</sup> *Treatise* 207:7-8 (I, IV, II)

<sup>460</sup> *Treatise* 207:9-11 (I, IV, II)

<sup>461</sup> *Treatise* 207:14-15 (I, IV, II)

<sup>462</sup> *Treatise* 207:17-22 (I, IV, II)

to any other perception or collection of perceptions, i.e. be introduced to a mind.<sup>463</sup>

Consequently “The supposition of the continu’d existence of sensible objects or perceptions involves no contradiction.”<sup>464</sup> As there is no logical contradiction in these suppositions there is no barrier to creating a fiction based upon them. This is precisely what we do, for “When the exact resemblance of our perceptions makes us ascribe to them an identity, we may remove the seeming interruption by feigning a continu’d being, which may fill those intervals, and preserve a perfect and entire identity to our perceptions.”<sup>465</sup>

Finally, Hume turns to the fourth part of his system which explains why we “not only *feign* but *believe* this continu’d existence,”<sup>466</sup> and that, therefore, “the question is, *from whence arises such a belief*”.<sup>467</sup> Belief is engendered by the “vivacity of an idea”<sup>468</sup> and Hume argues that a succession of related ideas is almost as vivid in the mind as a single idea.<sup>469</sup> Further, Hume explains how memory and resemblance combine, so that “sometimes we ascribe a continu’d existence to objects, which are perfectly new to us, and of whose constancy and coherence we have no experience.”<sup>470</sup> The reason for this “‘tis because the manner, in which they present themselves to our senses, resembles that of constant and coherent objects”.<sup>471</sup> Hume is therefore able to explain how the imagination is able to derive an ever increasing hierarchy of fictions. Perceptions resemble not only each other, but also other fictions. The imagination can progress with less and less reference to the actual contents of our perceptions. Without this efficiency the imagination could scarce be claimed to achieve all Hume would want of it. The workload of producing a unique fiction for every object would surely be mentally exhausting. What is more it would be antithetical to Hume’s conception of the imagination as a faculty that always prefers to take the line of least resistance.

---

<sup>463</sup> *Treatise* 207:23-32 (I, IV, II)

<sup>464</sup> *Treatise* 208:1-2 (I, IV, II)

<sup>465</sup> *Treatise* 208:4-8 (I, IV, II) Note here that the identity preserved fictitiously here is *perfect* and *entire*. As far as I am aware this is the only occasion on which Hume qualifies perfect identity in this way. However, the implication is clear: the perfect identity is a fiction derived from a succession. As such it cannot be contrasted with the so-called imperfect identity some writers postulate.

<sup>466</sup> *Treatise* 208:10 (I, IV, II) Hume’s use of the word ‘feign’ is a little curious here as it suggest artifice, and therefore, intent in perceiving the fiction. However, this quirk seems minor enough to be regarded as a slip or merely curious usage. Hume’s commitment to the inevitability of the fictions of imagination is well enough established elsewhere in the text.

<sup>467</sup> *Treatise* 208:10 (I, IV, II) Hume’s use of the word ‘feign’ is a little curious here as it suggest artifice, and therefore, intent in perceiving the fiction. However, this quirk seems minor enough to be regarded as a slip or merely curious usage. Hume’s commitment to the inevitability of the fictions of imagination is well-enough established elsewhere in the text.

<sup>468</sup> *Treatise* 208:13 (I, IV, II)

<sup>469</sup> *Treatise* 208:12-25 (I, IV, II)

<sup>470</sup> *Treatise* 209:6-9 (I, IV, II)

<sup>471</sup> *Treatise* 209:10-11 (I, IV, II)

The discussion of identity ends at this point, and is not returned to until the chapter '*Of Personal Identity*'. Hume finishes '*Of Scepticism with Regard to the Senses*' with his discussion of the conflict between the philosophical hypothesis of Double Existence and the Vulgar way of thinking.<sup>472</sup> In the three chapters before personal identity takes centre stage he discusses 'The Antient Philosophy'<sup>473</sup>, 'The Modern Philosophy'<sup>474</sup> and 'The Immortality of the Soul'.<sup>475</sup> These chapters prefigure the discussion of personal identity only in as much as they contain refinements of Hume's attack on the philosophical respectability of the notions of substance and the soul. These Hume dismisses as fictions created to explain the identity of objects. A full discussion of these secondary arguments would be too much of a distraction. On the occasions where Hume's later argument draws support from these chapters they will be referred back to.

---

<sup>472</sup> *Treatise* 209:14-218:34 (I, IV, II)

<sup>473</sup> *Treatise* 219:1-225:4 (I, IV, III)

<sup>474</sup> *Treatise* 225:5-231:31 (I, IV, IV)

<sup>475</sup> *Treatise* 232:1-251:3 (I, IV, V)

## VI

### IDENTITY IN HUME'S 'OF PERSONAL IDENTITY'

When Hume returns to the discussion of identity his subject is clearly meant to be the identity of the *mind*. He does not start his discussion of *personal identity* until later in the chapter.<sup>476</sup> Hume has started the chapter with his attack upon the Metaphysician's Self (see §I above) and concluded that we can only introspect our own perceptions. This leads Hume naturally into a discussion of how we ascribe identity to our grouped perceptions. The discussion is not yet of persons. Clearly it might be objected that Hume does write "I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement."<sup>477</sup> Here, Hume seems to use 'mankind' as a synonym for 'persons'. However, nothing forces us to assume this. The previous passage closed with Hume asserting that *he* could only find individual perceptions when he introspected.<sup>478</sup> Therefore, he reasons, this will be the case for the rest of mankind. The rest of mankind are, in this sense, only bundles or collections of perceptions. Indeed the term mankind seems chosen quite deliberately to be *neutral* between 'person' and 'self'. This is a debate Hume wishes to save for later.<sup>479</sup> For the moment he wants to highlight that all we have introspective knowledge of is a bundle or heap of perceptions.

The mind, Hume is quick to remind us, is in "perpetual flux",<sup>480</sup> a flux contributed to by both the senses<sup>481</sup> and the powers of the soul.<sup>482</sup> Therefore "There is properly no *simplicity* in it at one time, nor *identity* in different; whatever natural propension we may have to imagine that simplicity and identity."<sup>483</sup> Hume reminds us that this notion of the mind is quite abstract: "They are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which

---

<sup>476</sup> *Treatise* 259:1 (I, IV, VI, XV) "We now proceed to explain the nature of *personal identity*." One could not ask Hume to be clearer about where his discussion begins. Yet this simple sentence is rarely, if ever, acknowledged at face value. Certainly, I am aware of no commentator who does so.

<sup>477</sup> *Treatise* 252:31-35 (I, IV, VI, IV)

<sup>478</sup> *Treatise* 252:24-30 (I, IV, VI, III)

<sup>479</sup> Likewise Hume's mention of personal identity at 253:19 refers to the present chapter at large, and not to the specific preceding point of issue (which is just the mind and its lack of identity or simplicity).

<sup>480</sup> *Treatise* 252:35 (I, IV, VI, IV)

<sup>481</sup> *Treatise* 252:35-253:2 (I, IV, VI, IV)

<sup>482</sup> *Treatise* 253:2-4 (I, IV, VI, IV)

<sup>483</sup> *Treatise* 253:7-9 (I, IV, VI, IV)

it is compos'd."<sup>484</sup> This scepticism had already been voiced earlier in Hume's conclusions about the Soul. There he wrote;

[T]he question concerning the substance of the soul is absolutely unintelligible: All our perceptions are not susceptible of a local union, either with what is extended or unextended; there being some of them of the one kind, and some of the other: And as the constant conjunction of objects constitutes the very essence of cause and effect, matter and motion may often be regarded as the causes of thought, as far as we have any notion of that relation.<sup>485</sup>

Hume shares this determined agnosticism about what might ultimately constitute the mind with Locke. The notion of mind Hume is working with is, therefore, somewhat desiccated. This is, however, quite deliberate. It is a notion of mind stripped to its perceivable components. When Hume claims that the mind is only constituted by perceptions this is not a metaphysical claim. This much should be clear from his confession that *we do not know* what materials compose the mind. Our only *clear* notion of the mind is one of a place where individual perceptions collect and mingle. Hume's question is how we come to ascribe an identity to this mind.<sup>486</sup>

This is clearly a special problem for Hume. His initial account of identity of objects has imagination building on the foundations of coherence and constancy. Yet, the mind singularly fails to exhibit either coherence *or* constancy. Two possible strategies present themselves. Hume could attempt to bridge the gap between his earlier account of identity in terms of coherence and constancy and the identity of incoherent and fleeting objects. Alternatively he can introduce a new account of identity to deal specifically with these difficult cases.

The interpretation to be offered here takes the first route. It has already been observed that Hume regarded *all* identities as fictions, even Perfect Identity. Identity is no more than the relation that governs the duration of any being.<sup>487</sup> It is perfectly possible for Hume to extend his account of identity to include anything with duration. All that is required is an account of how we ignore even radical incoherence and inconstancy when ascribing identity.

---

<sup>484</sup> *Treatise* 253:10-14 (I, IV, VI, IV)

<sup>485</sup> *Treatise* 250:4-11 (I, IV, V)

<sup>486</sup> Compare Hume's clear comment that processes of the mind *do* depend on the physiology of the brain at *Treatise* 60:21-61:15 (I, II, V). See §XII for discussion.

<sup>487</sup> See definition at *Treatise* 12:25-31 (I, I, V)

Therefore, those accounts that suggest Hume is making use of a second kind of identity are to be rejected.<sup>488</sup> This second type of identity is usually called *Imperfect Identity*. The phrase itself is difficult enough to adopt as Hume only uses it once.<sup>489</sup> What is more this single occurrence is the subject of editorial conflict.<sup>490</sup> It is of course possible that Hume *might* employ two senses of identity, yet only refer to one of the senses on one occasion, and then without separate definition and scant elucidation. It is, however, unlikely. Nonetheless, there might be philosophically persuasive reasons for believing Hume employs two senses of identity, despite the paucity of explicit textual evidence. These considerations are of some import in understanding Hume, so they are worth examining.

Hume's previous discussion of identity had already established that Perfect Identity consists in being uninterrupted and invariable. As already mentioned, Hume's interest when he returns to discuss identity is with things that are very clearly both interrupted and variable (living things, rivers, persons).<sup>491</sup> Therefore he offers his definition this time in negative terms:

Thus the controversy concerning identity is not merely a dispute of words. For when we attribute identity, in an improper sense, to variable or interrupted objects, our mistake is not confin'd to the expression, but is commonly attended with a fiction, either of something invariable and uninterrupted, or of something mysterious and inexplicable [soul, self or substance], or at least with a propensity to such fictions.<sup>492</sup>

---

<sup>488</sup> Prime among these is Ashley & Stack (1974). Patten (1976a, esp. p670ff) follows them cautiously. Waxman (in an uncharacteristically uncritical moment) also assumes Hume makes use of the term, see Waxman (1994) p. 209-10.

<sup>489</sup> *Treatise* 256:24 (I, IV, VI, IX)

<sup>490</sup> Selby-Bigge suggests that Hume meant 'perfect' instead of 'imperfect'. Nidditch, however, finds no difficulty with the phrase and does not highlight it in his edition. For exhaustive debate see von Leyden (1957). Von Leyden suggests three solutions: to delete 'imperfect' altogether, to replace 'imperfect' with 'improper' (his own preferred solution) or to replace 'imperfect' with 'perfect' (Selby-Bigge's solution). My personal preference is for the first solution. Hume is only *qualifying* identity here to make the text more lucid. He is not marking a difference in theory or analysis. As the qualification makes the text less lucid it is best dispensed with altogether. Neither alternative qualification makes the text clearer, and von Leyden's suggestion has little precedence in the *Treatise* (primarily a single use at 255:8).

<sup>491</sup> There is a school of criticism that attacks Hume's account of personal identity because it relies on regularities in our experiences. These comments from Stroud are typical "But our experience in fact exhibits no such regularities. It is not true that we get an experience of a certain sort only when we have just had an experience of a certain other sort, or that experiences of the first sort are always followed by experiences of another sort. Our experience does not exhibit such uniformities. And for those who like a little novelty in life, that is a very good thing." Stroud (1977), p. 126. Clearly I find this type of attack to be quite mistaken. As will be argued shortly, Hume *embraces* the diversity of our perceptions (indeed repeatedly comments on the mind's flux) and deliberately extends his account of identity attribution to allow for this.

<sup>492</sup> *Treatise* 255:6-12 (I, IV, VI, VII). Hume's final concession that we merely suffer 'a propensity to such fictions' is worthy to note. On what occasions, one wonders, are we not prone to these fictions? However, Hume must allow for this possibility if judicious and concentrated reflection and reason are to have the reach he wished them to.

This passage prompts the belief that Hume is beginning to offer a different account of identity. The phrase ‘when we attribute identity, in an improper sense’ suggests that there are other occasions when it is *proper* to attribute identity. On this suggestion such occasions would, of course, be observations of Perfect Identity, i.e. uninterrupted and invariant objects. Importantly, Hume insists this is not (*contra* Locke) a matter of words or denominations, but a question of psychology. Whether we think of a billiard ball being the same, or a river being the same, the mental effort is all but equivalent. There is no difference despite the obvious truth (a truth we would all confess to observe) that the career of a river is more turbulent than that of a billiard ball. However, this similarity of effort also denotes a similarity of procedure. All identities are *equally* fictitious and therefore no distinction can be drawn between those that are proper and those that are improper.

However, Hume does not indulge here in pleonasm (Hume is too good a writer for this to be credible). There is no *philosophical* distinction between ‘Imperfect’, ‘Improper’ or ‘Loose’ Identity and ‘Perfect’, ‘Proper’ or ‘Strict’ Identity. Nonetheless, Hume is highlighting a distinction he has drawn in the previous passage, which does introduce a distinction between proper and improper *ascriptions* of identity. They, however, all remain equally suspect. Hume sketches a hierarchy of fictions we “feign”<sup>493</sup> with an identity. The worst are the complete fictions of “*soul, and self, and substance*”.<sup>494</sup> Hume then distinguishes occasions when *no* such fictions arise but “our propension to confound identity with relation is so great”<sup>495</sup> we instead “imagine something unknown and mysterious”.<sup>496</sup> Hume offers “plants and vegetables”<sup>497</sup> as examples. Importantly Hume uses the phrase “do[es] not give rise to such a fiction”<sup>498</sup> where the ‘such’ alludes to soul, self and substance. Then Hume distinguishes occasions when we ascribe identity without either an extreme fiction or the attribution of a ‘mysterious principle’ but nonetheless “are not able fully to satisfy ourselves”<sup>499</sup> by finding something to “justify our notion of identity.”<sup>500</sup> All these occasions are fictions of the imagination, and all produced in similar fashion. The phrase ‘improper’ merely highlights that “variable and interrupted objects”<sup>501</sup> are *more likely* to be attended with an extreme fiction than more steadfast objects. In this passage Hume again urges that

---

<sup>493</sup> *Treatise* 254:25 (I, IV, VI, VI)

<sup>494</sup> *Treatise* 254:27 (I, IV, VI, VI)

<sup>495</sup> *Treatise* 254:29-30 (I, IV, VI, VI)

<sup>496</sup> *Treatise* 254:31 (I, IV, VI, VI)

<sup>497</sup> *Treatise* 255:1 (I, IV, VI, VI)

<sup>498</sup> *Treatise* 254:29 (I, IV, VI, VI)

<sup>499</sup> *Treatise* 255:3-4 (I, IV, VI, VI)

<sup>500</sup> *Treatise* 255:5 (I, IV, VI, VI)

<sup>501</sup> *Treatise* 255:8 (I, IV, VI, VII)

there is a hierarchy.<sup>502</sup> What is more, the previous passage is quite clear why these special cases of fiction emerge. Hume highlights that “tho’ we incessantly correct ourselves by reflexion, and return to a more accurate method of thinking, yet we cannot long sustain our philosophy, or take this bias from the imagination.”<sup>503</sup> Finally “Our last resource is to yield to it, and boldly assert that these different related objects are in effect the same”.<sup>504</sup> It is this act of ‘boldly asserting’ Hume is interested in. In less complicated identity claims (steadfast objects) the imagination makes easy progress. In more complex cases there is also an attendant ‘assertion’. There is no difference in mental effort for the imagination is equally irresistible in its progression. However, certain identity claims are so clearly difficult that we have to feign some principle to accompany the action of the imagination. This is all Hume wishes to mark out by the phrase ‘in an improper sense’.<sup>505</sup>

Putting this distinction aside, Hume proceeds with his general argument by again repeating points made in the earlier discussion.<sup>506</sup> Succession answers to the notion of diversity, but we attribute identity to it because “the relation of parts... produces an association of idea, and an easy transition of the imagination”<sup>507</sup> and the resemblance “this act of the mind bears to that, by which we contemplate one continu’d object”.<sup>508</sup> Hume’s task, then, must be to show that things we think we observe to be identical are in fact composed of successions.<sup>509</sup>

At this point Hume offers what amounts to a second definition of Perfect Identity.

[S]uppose any mass of matter, of which the parts are contiguous and connected, to be plac’d before us; ‘tis plain we must attribute a perfect identity to this mass, provided all the parts continue uninterruptedly and invariably the same, whatever motion or change of parts we may observe either in the whole or in any of the parts.<sup>510</sup>

---

<sup>502</sup> *Treatise* 255:9-12 (I, IV, VI, VII)

<sup>503</sup> *Treatise* 254:16-19 (I, IV, VI, VI)

<sup>504</sup> *Treatise* 254:19-21 (I, IV, VI, VI)

<sup>505</sup> One should also consider the consequences of inverting the sentence’s meaning. Hence “When we attribute identity in a proper sense to invariable and uninterrupted objects,...”. If this is correct (i.e. all the change logic requires) there is little impact on my reading. For this proper sense is equally attended by fictions, etc. Improper identity is only of interest if its corollary, proper identity, can be shown not to rely on a fiction. And this cannot be done. In a similar vein, therefore, I suggest that all of the various words Hume combines with identity are merely qualifications. Each denotes the degree of imagination involved. This no more commits one to having more than one identity relation than would The Sortal Dependency Thesis commit one to Relative Identity (See earlier discussion in Chapter 2, §II)

<sup>506</sup> See *Treatise* 204n1 (I, IV, II)

<sup>507</sup> *Treatise* 255:20-23 (I, IV, VI, VII)

<sup>508</sup> *Treatise* 255:24-25 (I, IV, VI, VII)

<sup>509</sup> *Treatise* 255:26-29 (I, IV, VI, VII)

<sup>510</sup> *Treatise* 255:30-35 (I, IV, VI, VIII) The example is lifted wholesale from Locke *Essay* 330:14-20 (Bk II, XXVII, §3). For discussion of the importance of this to Locke see Chapter 2, §II in this thesis.

Crucially, Hume chooses to explain Perfect Identity here not in terms of invariance and uninterruptedness in our perceptions, but in the object itself. On all prior occasions Hume had explained identity as a fiction of the imagination that arose from coherence and constancy in our *perceptions*.

One could concede that on this occasion Hume allows himself to indulge in outright metaphysics and make a sure claim about the nature of body. Further, he claims that we could perceive the identity of such an object in a direct fashion unmediated by the imagination. Yet this seems extremely unlikely. Identity is *essentially* a fiction, being something the imagination creates betwixt unity and diversity. It simply cannot be observed. Further, it would seem incredible that Hume should labour so much effort on *denying* that the senses comprehend identity, only to latterly claim that certain objects *could* be observed to be identical.<sup>311</sup>

Hume would also be courting further disaster. He continues "But supposing some very *small* or *inconsiderable* part to be added to the mass, or subtracted from it; [...] this absolutely destroys the identity of the whole, strictly speaking".<sup>312</sup> Defining Perfect Identity in the manner above has the effect of making change and identity incompatible. It seems entirely antagonistic to Hume's carefully constructed Sceptical Realism to suggest that he embraces the extreme sceptical view that nothing at all can survive change. Not only would this run counter to common-sense, it would involve a positive metaphysical claim about the relations between the parts of objects over time. It is surely Hume's opinion that we are in no position to decide such matters.

Fortunately, I believe, there is a better explanation at hand for what Hume is doing here. Recall the importance Hume had placed on being able to *visually* conceive identity in his initial discussion. In this passage Hume is offering a *visual thought-experiment*. Hence his stipulation to 'suppose'. Hume has not forgotten that we could not observe such a mass directly, nor has he abandoned his position that it is regularities in our perceptions that give rise to attributions of identity.

By Hume's own lights having the experience of perceiving an unchanging object would only prove that one was experiencing an invariant and uninterrupted sequence of perceptions. Equally if we did observe a small change in something, we would have formed

---

<sup>311</sup> Notwithstanding Hirsch's claims that Hume *does* derive the notion of identity from the senses. See Hirsch (1983).

<sup>312</sup> *Treatise* 255:35-256:2 (I, IV, VI, VIII)

a new perception, and be able to perceive this diversity (though we might not make the mental effort to do so). Hume observes “yet as we seldom think so accurately, we scruple not to pronounce a mass of matter the same, where we find so trivial an alteration.”<sup>513</sup> Hume’s meaning here is potentially ambiguous. The accurate thinking Hume refers to is not just the judgement that change in something strictly destroys its identity. It also refers to the mental effort required to act on such observations. This keys in with Hume’s proclaimed system in which resemblance has a dual action. Hume notes that we may well ‘find so trivial an alteration’. However, we rarely make the mental effort to focus on such alterations specifically, nor consider their general implication. Similarly he continues “The passage of the thought from the object before the change to the object after it, is so smooth and easy, that we scarce perceive the transition.”<sup>514</sup> Again ‘scarce’ is ambiguous. We *barely* perceive the change, i.e. it is on the fringes of perception, and we *rarely* i.e. infrequently make the effort to observe the difference. Hence Hume can make his customary conclusion and insist we “are apt to imagine, that ‘tis nothing but a contin’d survey of the same object.”<sup>515</sup> Hume’s purpose has been served. He has demonstrated how the imagination works in very similar fashion when perceiving an unchanging series of perceptions and when perceiving a slightly changing series of perceptions.

It might be objected that this still does not explain why Hume chooses this example. Partly, it is merely slavish adoption of part of Locke’s argument.<sup>516</sup> The main reason, though, is to provide a touchstone for Hume’s project. If Hume is to argue that we come to think of the identity of a river in the same way we think of the identity of a rock he will need to illustrate every mid-point between steadfast inertia and perpetual flux and demonstrate that this is (at least to the mind) some sort of continuum. Hume’s own theory of time has ruled out the observation of any Perfect Identity. He is, therefore, obliged to offer a thought-experiment to establish his bedrock. He offers a manner in which we may conceive what we cannot observe, a perfectly unchangeable object. This is why just as Locke had talked of the parts of a mass being “never so differently jumbled”<sup>517</sup> Hume allows for “whatever motion or change of parts we may observe”.<sup>518</sup> The mass’ perfect lack of change is guaranteed by the identity of its constituents, not the arrangements of its parts. Relying on the arrangement of the parts would be dangerously close to Scholasticism. It is merely the

---

<sup>513</sup> *Treatise* 256:2-4 (I, IV, VI, VIII)

<sup>514</sup> *Treatise* 256:4-6 (I, IV, VI, VIII)

<sup>515</sup> *Treatise* 256:6-8 (I, IV, VI, VIII)

<sup>516</sup> See *Essay* 329:8-9 (Bk. II, Ch. XXVII, §2) and 330:14-20 (Bk. II, Ch. XXVII, §3) See Hall (1974) for further details of Hume’s use of Locke.

<sup>517</sup> *Essay* 330:18 (Bk. II, Ch. XXVII, §3)

<sup>518</sup> *Treatise* 255:34 (I, IV, VI, VIII)

aggregate of parts, however arranged, that constitutes a mass for Locke and Hume. In this way Hume demonstrates how easily we confuse this notion of a steadfast object with something which does change its constituent parts, even if only very slightly.

Not unsurprisingly this argument occurs at the beginning of a sequence of conditions Hume suggests enable us to attribute identity to things which are variant and interrupted. Hence he suggests that if a change is in proportion it is easier to overlook.<sup>519</sup> This is more so if the change occurs “gradually and insensibly”.<sup>520</sup> What is more in the case of artefacts (such as a ship)<sup>521</sup> and living things (“animals and vegetables”)<sup>522</sup> we can assume their parts to have a “common end”<sup>523</sup> and “sympathy”.<sup>524</sup> In each case Hume urges the effect these have on the imagination. They variously result in “the uninterrupted progress of the thought”<sup>525</sup>, its “easy passage”<sup>526</sup> and “easy transition”.<sup>527</sup> As a consequence the mind “at no particular time perceives any interruption in its actions.”<sup>528</sup> This means the mind is not obliged to form a new perception and with it the notion of diversity.<sup>529</sup>

Having shown how these factors prevent us perceiving diversity, Hume discusses occasions when despite perceiving a diversity we still persist in attributing identity. This occurs when “tho’ we commonly be able to distinguish pretty exactly betwixt numerical and specific identity, yet it sometimes happens, that we confound them, and in our thinking and reasoning employ one for the other.”<sup>530</sup>

Hume’s examples here are less than convincing. He suggests that we confuse separate sounds in a sequence (e.g. a fog horn’s blasts) as numerically the same even though “there is nothing numerically the same, but the cause, which produc’d them.”<sup>531</sup> This simply seems wrong. Hume also says we treat artefacts as numerically identical even through entire

---

<sup>519</sup> *Treatise* 256:9-24 (I, IV, VI, IX) It is in this passage that Hume’s controversial use of ‘Imperfect’ occurs.

<sup>520</sup> *Treatise* 256:28, 256:25-35 (I, IV, VI, X)

<sup>521</sup> *Treatise* 257:9-15 (I, IV, VI, XI) Hume clearly has the Ship of Theseus Problem in mind here.

<sup>522</sup> *Treatise* 257:20 (I, IV, VI, XII)

<sup>523</sup> *Treatise* 257:5-15 (I, IV, VI, XI)

<sup>524</sup> *Treatise* 257:16-31 (I, IV, VI, XII) Hume remarks particularly on the strength of this relation *Treatise* 257:23 (I, IV, VI, XII)

<sup>525</sup> *Treatise* 256:22-23 (I, IV, VI, IX)

<sup>526</sup> *Treatise* 256:31 (I, IV, VI, X)

<sup>527</sup> *Treatise* 257:14 (I, IV, VI, XI)

<sup>528</sup> *Treatise* 256:32-22 (I, IV, VI, X)

<sup>529</sup> Hume notes the importance of this in his original definition of identity. See *Treatise* 201:29-34 (I, IV, II) esp. 201:33-34.

<sup>530</sup> *Treatise* 257:33-258:1 (I, IV, VI, XIII) Until this point Hume had resisted suggesting that we confound numeric and specific identity, despite the fact this would be obviously useful in his early discussion of identity. If Hume had suggested that all identity statements arose as a result of confounding numeric and specific identity he would not have been able to explain those occasions on which we do distinguish the two. Hume must avoid saying that specific identity *causes* us to attribute numeric identity. This would not allow for the exercise of reason to prevent us from making the easy associations the imagination would indulge in.

changes of components (e.g. a church re-built with different material in a different style).<sup>532</sup> Interestingly here Hume comments this is enough to make us “denominate”<sup>533</sup> the buildings the same and we do so “without breach of the propriety of language.”<sup>534</sup> Hume speaks of “the same method of reasoning, [...], which has so successfully explain’d the identity of plants, and animals, and ships, and houses, and all the compounded and changeable productions of either art or nature.”<sup>535</sup> This very much suggests that Hume does not distinguish between the identity of natural and non-natural objects.<sup>536</sup> Yet, the identity we attribute to the church is merely a denomination. This does not seem as strong as the kind of fiction that normally accompanies an identity claim and so renders it “not merely a dispute of words.”<sup>537</sup> Similarly Hume continues his discussion of cases where we confuse numeric and specific identity by discussing a river.<sup>538</sup> Hume remarks that “the change of parts [should] be not sudden or entire, in order to preserve the identity”.<sup>539</sup> However, this is not necessary “where the objects are in their *nature* changeable and inconsistent”.<sup>540</sup> This appeal to what is “natural and essential”<sup>541</sup> in objects to explain why we attribute identity to them suggests a covert reliance on a distinction between natural and non-natural identities. Hume’s failure to explicitly discuss this distinction is a weakness in his account (especially compared to Locke’s).

However, it is a minor point, and Hume substantiates his major point: that all identities consist of sequences of perceptions united by the imagination. He does this by extending his original account of identity and without any need to soften his principles or introduce a different notion of identity. The only task left is for Hume “to explain the nature of *personal identity*, which has become so great a question in philosophy, especially of late years in *England*, where all the abstruser sciences are study’d with a peculiar ardour and application.”<sup>542</sup>

---

<sup>531</sup> *Treatise* 258:4-5 (I, IV, VI, XIII)

<sup>532</sup> *Treatise* 258:6-13 (I, IV, VI, XIII)

<sup>533</sup> *Treatise* 258:13 (I, IV, VI, XIII) compare *Essay* 348:21-15 (Bk. II, Ch. XXVII, §29).

<sup>534</sup> *Treatise* 258:6-7 (I, IV, VI, XIII)

<sup>535</sup> *Treatise* 259:5-9 (I, IV, VI, XV)

<sup>536</sup> Though, of course it could merely mean that one method of reasoning stands behind two separate explanations, one for natural objects and one for artefacts.

<sup>537</sup> *Treatise* 255:6-7 (I, IV, VI, VII)

<sup>538</sup> *Treatise* 258:25-29 (I, IV, VI, XIV)

<sup>539</sup> *Treatise* 258:21-23 (I, IV, VI, XIV)

<sup>540</sup> *Treatise* 258:23-24 (I, IV, VI, XIV) emphasis added.

<sup>541</sup> *Treatise* 258:29 (I, IV, VI, XIV)

<sup>542</sup> *Treatise* 259:1-4 (I, IV, VI, XV)

## VII

### HUME'S EXPLANATION OF PERSONAL IDENTITY

As has already been stated Hume distinguishes between the problem of the mind's identity and personal identity. However, he does see the two as being intimately related. This is an obvious consequence of holding a version of the Compositional Account. If a person is composed of a mind and a body establishing the identity of the mind is clearly a crucial issue. Nonetheless, Hume does *not* identify the mind and the person. It is, therefore, necessary to go to some lengths to point up the subtleties of Hume's account. Without this effort it is easy to come to the conclusion that Hume does identify the mind and the person. This is at great cost to the cogency of Hume's account. Of course, as Hume's account is normally regarded as an interesting failure many writers have not troubled themselves to present it in its strongest form.

Hume begins by urging that personal identity will be explained in a manner continuous with the identity of bodies. The previous passages had prepared the ground for this by explaining how the same theory that explained the identity of stable material objects could explain the identity of more changeable objects, as diverse as living things, tunes, rivers, ships and churches. Hence Hume writes:

The identity, which we ascribe to the mind of man, is only a fictitious one, and of a like kind with that which we ascribe to vegetables and animal bodies. It cannot, therefore, have a different origin, but must proceed from a like operation of the imagination upon the objects.<sup>43</sup>

Hume's discussion here, confessedly, concerns the mind. This in no way implies, though, that Hume identifies the mind and the person. All it suggests is that Hume feels he will have to explain the identity of the mind before he can move on to explain personal identity. There seems little doubt that the ordering of the discussion has influenced many readers into thinking Hume is about to identify mind and person. However, two things count strongly against this. First, it is possible to develop a cogent account of Hume in which he doesn't identify the two. That is to say, the identification of mind and person is no way intrinsic to Hume's account. Secondly, and maybe more conclusively, if it is Hume's intention to

---

<sup>43</sup> *Treatise* 259:9-13 (I, IV, VI, XV)

explain personal identity as the identity of the mind one would expect him to declare as much.<sup>344</sup> This he does nowhere.<sup>345</sup>

For the moment, though, mental identity is Hume's topic. Explaining it is the denouement in Hume's account of identity. Hume must attempt to explain a great paradox (the very one that comes to haunt him in the *Appendix*). The mind is made of many perceptions. Hume is clear that all we perceive are perceptions. Yet we *also* experience the fiction of the mind being perfectly identical and simple. Hume remarks on this difficulty:

'Tis evident, that the identity, which we attribute to the human mind, however perfect we may imagine it to be, is not able to run the several different perceptions into one, and make them lose their characters of distinction and difference, which are essential to them.<sup>346</sup>

Yet

[N]otwithstanding this distinction and separability, we suppose the whole train of perceptions united by identity<sup>347</sup>

This 'supposition' of identity means "a question naturally arises concerning this relation of identity".<sup>348</sup> Fascinatingly Hume chooses to pose this question twice: once in first-person terms, once in third-person terms.

[W]hether it be something that really binds our several perceptions together, or only associates their ideas in the imagination. That is, in other words, whether in pronouncing concerning the identity of a person, we observe some real bond among his perceptions, or only feel one among the ideas we form of them.<sup>349</sup>

---

<sup>344</sup> He nowhere, for example, says "I am nothing but a 'personal' bundle of perceptions" as Stroud (1977) p. 132 suggests. This is to run together Hume's theory of personal identity and his bundle theory of *mind*. Each of us owns a personal bundle of perceptions (our minds) but we also have perfectly real bodies as well. So we are not merely our minds, and therefore not merely our personal bundles.

<sup>345</sup> He does on one occasion use the phrase "mind or thinking person" at *Treatise* 260:24-25 (I, IV, VI, XVII). However, this is entirely consistent with the Compositional Account. The phrase occurs at a crucial point in Hume's discussion as he starts to shift from his account of mental identity to personal identity. The phrase, therefore, is a bridging one, and means "compositional person *qua* mental being". There is nothing about 'thinking person' which necessarily excludes the corporeal. It might be objected that if Hume doesn't mean 'mind' by the phrase 'thinking person' he must have a definition for 'unthinking person'. Contrarily, if 'mind' and 'thinking person' are equivalent 'unthinking person' becomes obligingly meaningless. However, Hume *does* have a meaning for this: the unthinking person is the body. In the next sentence but one Hume uses the far more comfortable phrase "mind or thinking principle" *Treatise* 260:30 (I, IV, VI, XVIII). Therefore, this phrase in no way proves that Hume identifies mind and person. For more see discussion below.

<sup>346</sup> *Treatise* 259:17-21 (I, IV, VI, XVI)

<sup>347</sup> *Treatise* 259:25-27 (I, IV, VI, XVI)

<sup>348</sup> *Treatise* 259:27-28 (I, IV, VI, XVI)

<sup>349</sup> *Treatise* 259:28-34 (I, IV, VI, XVI)

This repetition is not an idle amplification of a point.<sup>530</sup> It is the beginning of a systematic treatment of personal identity in both first-person and third-person terms. This will be returned to shortly. For the moment it is necessary to absolutely clarify what Hume thinks the rest of the naturally arising question concerns.

It is clear Hume feels the question is whether we observe a real connection in personal identity or not. It is also abundantly clear that Hume answers that we do not:

This question we might easily decide, if we wou'd recollect what has been already prov'd at large, that the understanding never observes any real connexion among objects<sup>531</sup>... from thence it evidently follows, that identity is nothing really belonging to these different perceptions, and uniting them together; but is merely a quality, which we attribute to them, because of the union of their ideas in the imagination, when we reflect upon them<sup>532</sup>... it follows, that *our notions of personal identity*, proceed entirely from the smooth and uninterrupted progress of the thought along a train of connected ideas<sup>533</sup>

Hume only wishes to explain our *notion* of personal identity. The true and uniting principle of personal identity is not his subject.

Realising this is the case renders obsolete a whole school of Humean criticism. Hume's theory cannot be inadequate because he fails to explain how persons are *actually* bundled together.<sup>534</sup> Hume is not indulging in this metaphysical task. Hume is merely explaining how we come to believe in persons when they are not present in our perceptions in the way we believe them to be. There is no difference between his task here and his task in 'Of Scepticism With Regard to the Senses'. Just as he asserts there that it is vain to ask whether or not body exists, here it is vain to ask whether there are persons or not.<sup>535</sup> As with external objects, Hume's question is one of how we come to hold the belief. At this point Hume

---

<sup>530</sup> Something Hume can hardly ever be accused of, unlike Locke, who one rather wishes had removed as much from the *Essay* as he added.

<sup>531</sup> *Treatise* 259:34-260:1 (I, IV, VI, XVI)

<sup>532</sup> *Treatise* 260:3-7 (I, IV, VI, XVI)

<sup>533</sup> *Treatise* 260:17-20 (I, IV, VI, XVI) emphasis added.

<sup>534</sup> Stroud is typical of this approach asking "Is it simply the resemblance in a sequence of perceptions that leads us to think of them as constituting or belonging to one mind?" Stroud (1977), p. 124 and "Even if a causal chain would 'tie' the perceptions together in our minds in a way that was missing from a bundle formed by mere resemblance, it is not clear that that would be enough to lead us to think of it as one mind." Stroud (1977), p. 125. The condition Stroud is insisting on is too *strong*. Hume is only looking at our notion of personal identity not what actually constitutes persons. Clearly our notion of personal identity can be grounded on weaker imaginative principles. Of course, Hume in the Appendix comes to worry that he cannot discover stronger conditions for personal identity. Stroud's complaints would have more weight here if Hume is committed to thinking that the principles that form our imaginative notion of personal identity are the same ones that actually constitute persons. But further argument is needed to show Hume thinks this, especially as this would be at odds with his usual scepticism.

<sup>535</sup> Or, more properly, in vain to ask whether their be mind or not, and as persons are composed of a mind and a body, in vain to ask if there are persons or not.

sees no problem with treating personal identity as he has treated all other identities, as a fiction of the imagination. In the *Appendix* Hume states:

Philosophers begin to be reconcil'd to the principle, *that we have no idea of external substance, distinct from the ideas of particular qualities*. This must pave the way for a like principle with regard to the mind, *that we have no notion of it, distinct from the particular perceptions*.<sup>556</sup>

It is in this light that one must understand Hume's account of personal identity. When this account is clear Hume's confusions in the *Appendix* become more obvious. Therefore, for the moment Hume's second thoughts will be put aside. Hume's account of our notion of personal identity will remain our present subject. Shortly the "three relations"<sup>557</sup> of the imagination and the use Hume puts them to will be examined. The role they play, it will be argued, can only be understood in terms of the split between first-person and third-person concerns Hume has already hinted at.<sup>558</sup> Most commentators have ignored the role of the third person in Hume's account, or given it short shrift,<sup>559</sup> preferring to concentrate on the first-person implications.<sup>560</sup> The balance, however, will be redressed when one realises how scathing Hume is about the usefulness of first-person observations in acquiring our notion of personal identity. This will be discussed next.

---

<sup>556</sup> *Treatise* 635:10-14 (App. X)

<sup>557</sup> *Treatise* 260:9 (I, IV, VI, XVI)

<sup>558</sup> *Treatise* 259:28-34 (I, IV, VI, XVI)

<sup>559</sup> Stroud, for example, notes "Hume's penchant for talking about other people, not himself" and insists the problems are best understood when translated into first-person ones! See Stroud (1977) p. 129-130

<sup>560</sup> This is indicative of the contemporary trend to regard third-person problems of personal identity as somehow metaphysically dubious or secondary. This results from a confusion between the problem of self-knowledge and the problem of personal identity. Related as the two are it is not obvious that the solution of the former is the necessary route to the solution of the latter.

## VIII

### THE FAILURE OF INTROSPECTION

Hume's asks "From what impression cou'd this idea [of the self] be deriv'd? [...] 'tis a question, which must necessarily be answer'd, if we wou'd have the idea of self pass for clear and intelligible."<sup>61</sup> Knowledge of the Metaphysicians' Self should be derived from perceptions. As Hume reminds the reader "It must be some one impression, that gives rise to every real idea."<sup>62</sup> Hume has already explained how we can arrive at a notion of Perfect Identity by use of reason. As we know every idea must arise from an impression, we can infer what sort of impression would be needed to perceive something perfect and simple.

Hume, then, is looking for a single impression that satisfies the suggested notion of a self, i.e. a unitary, simple entity. This is not to be found. Instead Hume says, "For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception." However, Hume takes the claim further than merely claiming that he finds particular perceptions when he looks in upon his self for "I never can catch *myself* at any time without a perception, and never can observe any thing but the perception."<sup>63</sup> It is certainly the case that Hume believes that the self being made up of a manifold of perceptions means that any idea we have of ourselves cannot be of the Metaphysicians' Self, saying "[S]ensations succeed each other, and never all exist at the same time. It cannot, therefore, be from any of these impressions, [...], that the idea of self is deriv'd; and consequently there is no such idea."<sup>64</sup>

Hume's claim, then, is not just that he can not catch the Self without a perception, but that he can *only* catch the perceptions. The perceptions obscure, necessarily, any attempt to access the Metaphysicians' Self. One *cannot* catch the self with a perception, the way one *can* catch a thief with his swag. It is clear that Hume feels that the notion of the Metaphysicians' Self demands this possibility. Self is "that to which our several impressions and ideas are supposed to have a *reference*."<sup>65</sup> But as Hume claims "'tis intelligible and consistent to say, that objects exist distinct and independent, without any

---

<sup>61</sup> *Treatise* 251:18-23 (I, IV, VI, II)

<sup>62</sup> *Treatise* 251:23-24 (I, IV, VI, II)

<sup>63</sup> *Treatise* 252:15-17 (I, IV, VI, III)

<sup>64</sup> *Treatise* 252:1-5 (I, IV, VI, II)

<sup>65</sup> *Treatise* 251:25-26 (I, IV, VI, III) my italics

common *simple* substance or subject of inhesion.”<sup>366</sup> This restates the attack on substance Hume has maintained in the chapters prior to ‘Of Personal Identity’.

Hume is not alone in wanting to avoid believing in bare particulars. But this does not make it necessary to be a sceptic about the self. Hume goes further claiming the relation of perceptions to any supposed self would be *irrelevant* for “When my perceptions are remov’d for any time, as by sound sleep; so long am I insensible of *myself*, and may truly be said not to exist.”<sup>367</sup> Hume’s point here is double edged. Even if one was a soul or further entity, when one has no perceptions one is not *aware* of this further fact. The Metaphysicians’ Self cannot be constituted by something of which we are not aware.<sup>368</sup> It is something “we are every moment intimately conscious of”.<sup>369</sup> Also, though, as we are just our perceptions, when we have no perceptions, then we truly cease to exist. Conceived of as a purely psychological entity, Hume insists that the self must have an intermittent existence.

It is curious, though, that when Hume himself remarks on the intimate consciousness we have of ourselves he never fails to stress that we are *always* so conscious. This appears to result in a stark contradiction for Hume. If all we are intimately conscious of is perceptions then it is hard to understand what is left to be intimately conscious of when all one’s perceptions are removed. Waxman solves this dilemma by suggesting that Hume does *not* believe the mind has an interrupted existence.<sup>370</sup> Waxman thinks the mistake arises by “confusing discontinuous existence with discontinuous memory”.<sup>371</sup> Hume does think our identity extends beyond what we can remember, however, this does not seem to answer his problem. His point is not that we do not always remember what we think when we sleep. His point is that sound sleep *removes* our perceptions.

However, there are two ways Hume can be saved from this apparent contradiction. One is to limit the scope of ‘always’ in the phrase ‘always intimately conscious’. On this reading Hume is merely insisting that when we are conscious we are always intimately conscious to ourselves. His comment is about the experience of consciousness, not about the metaphysics of mind. As such it is merely a re-expression of the Transparency Thesis. Understood this way Hume is not obliged to claim that the mind always thinks. This can be

---

<sup>366</sup> *Treatise* 634:19-22 (App. V)

<sup>367</sup> *Treatise* 252:17-19 (I, IV, VI, III)

<sup>368</sup> Or at least Hume finds this ridiculous. Compare with Locke’s arguments that thinking that always remained unknown, even to itself, is nonsensical; *Essay* 109:31-116:15 (Bk II, Ch. I, §11-19)

<sup>369</sup> *Treatise* 251:4-5 (I, IV, VI, I)

<sup>370</sup> Waxman (1994) p. 323n19

combined with the second possibility, though it need not be. This is fortunate because I suspect it *doesn't* capture Hume's intention. Hume really does think we are intimately conscious of ourselves at all times.

Hume insists that a deep sleep removes our perceptions but insists that it would take death to remove *all* our perceptions and, further, "nor do I conceive what is farther requisite to make me a perfect non-entity".<sup>571</sup> Clearly there is nothing controversial in Hume's claim that sleep and death are not the same thing. However, the distinction suggests that sleep, even deep sleep, does *not* remove all of our perceptions. It would appear, therefore, that Hume veers towards believing that we always think in some capacity. This could put Hume in opposition to Locke, who goes to great lengths to show that the *soul* does not always think.<sup>572</sup> However, this is only so if Hume is urging that this intimate consciousness is purely *intellectual*. Both Locke and Hume allow that some parts of 'thinking' occur partially in the body. Death would remove the continuity of this sort of activity, whereas sleep would not. Hume could be appealing to a continuity of this kind.

Hume's phrasing of the discussion does provide support for this reading. He states, "And were all my perceptions remov'd by death, and cou'd I neither think, nor feel, nor see, nor love, nor hate *after the dissolution of my body*, I shou'd be entirely annihilated."<sup>574</sup> The question of course arises as to what form this perceiving takes which goes on even in a deep sleep. Obviously it is not thinking on a fully-conscious level.<sup>573</sup> However, Hume could allow for continuity of perceptions at a very low-level. He for example allows that a creature might have only one perception.<sup>576</sup> On this principle all Hume needs is for us to have an ongoing succession of very minor perceptions. It seems reasonable to claim we do have constant sensory perceptions even when asleep. How else could loud noises or bright lights come to wake us if sleep excluded being aware of them? There is a problem for Hume because we are not seemingly transparently conscious of these perceptions. However, Hume can counter that all this means is we do not *remember* many of the

---

<sup>571</sup> Waxman (1994) p. 323n19

<sup>572</sup> *Treatise* 252:22-23 (I, IV, VI, III) Note Hume's use of 'perfect' here. It is, of course, impossible to be an 'imperfect non-entity'. The role of 'perfect' as a qualification here is purely literary. It has no role other than to parody those philosophers who believe in soul, substance or some other substrata. Hume characterises such underproppings as quite unintelligible. This literary or qualifying usage of 'perfect' should be compared with Hume's various usage of terms (attributed, imperfect, improper, etc.) in combination with identity.

<sup>573</sup> See Chapter II, §V in this thesis.

<sup>574</sup> *Treatise* 252:19-22 (I, IV, VI, III) Emphasis added.

<sup>575</sup> Though, Hume *doesn't* commit to defining the distinction between sleeping and waking in these terms. Sleeping could be a purely physical state in which we might find ourselves fully conscious (a hypnotic state maybe). However, Hume says nothing further about the nature of sleep (as far as I'm aware) so speculation on this point is not very helpful, fascinating as the general topic is (especially considering Locke's comments, and the importance of the distinction to Descartes).

thoughts we have when sleeping. Hume's insistence that we can extend our identity beyond what we can remember, therefore, takes care of periods while we sleep along with waking periods we don't remember.

This doesn't entirely settle what Hume means by being always intimately conscious. For present purposes, though, it serves to demonstrate further the failure of introspection on Hume's view. Hume claims our introspective self is intermittent. This does not imply that the person is intermittent, because Hume does not think introspection can discover the person. This should not surprise when we recall Hume's own distinction between personal identity as it regards the imagination and personal identity as it regards the passions. The imagination demands full consciousness, the passions do not. Hume, therefore, can claim that the person is always intimately conscious to us because the person contains both the conscious thoughts of the imagination, reason and the passions. The former of these are intermittent and contrary. The passions, in contrast, are constant. The passions, of course, depend upon the body. Hume is involved in no contradiction here if he makes use of the Compositional Account. In short, Hume's analysis of the role of the imagination is as clear an argument *against* a psychological account of personal identity as one could hope to find.

Given the failure of introspection, Hume has to look elsewhere for the source of our notion of personal identity. This is why he introduces, and lays so much emphasis upon, third-person examples. Our notion of person is a synthesis of the identity we observe in others and the identity we observe in such an unsatisfactory and limited fashion within ourselves.

---

<sup>576</sup> *Treatise* 634:28-31 (App. VII)

## IX

### THREE RELATIONS

Hume suggests there are three relations that convince us that the self is identical, perfect and simple and not merely a succession of perceptions. These are contiguity, resemblance and causation, though Hume has little to say about contiguity.<sup>577</sup> The positivist and constructivist view of Hume presumes these three to be responsible for *creating* personal identity at a metaphysical level. This interpretation arises from the positivist doctrines foisted upon Hume by latter day empiricists. They regard persons as logical constructions, and interpret Hume as offering mechanisms by which this construction takes place (adding memory to the list). This is, as has already been argued, not Hume's project at all.

Hume elects to put contiguity to one side. Although he provides no reason for this in the chapter 'Of Personal Identity' itself, his reasoning is given clearly in the previous chapter 'Of The Immateriality of the Soul'. There Hume has argued that "some beings exist without any place",<sup>578</sup> examples of such beings being tastes, smells and thoughts. Hume concedes that "in this view of things we cannot refuse to condemn the materialists, who conjoin all thought with extension".<sup>579</sup> However, this does not commit Hume to idealism, but to a dualism,<sup>580</sup> for some perceptions, namely those of sight and touch can be properly regarded as extended. This is because sight and touch contain the idea of extension and "To say the idea of extension agrees to any thing, is to say it is extended."<sup>581</sup>

Contiguity can only play a role in the relating of perceptions if they can be said to exist in space. As Hume has excluded many perceptions from being located in space contiguity will not play a significant role. Further, even if a person had a sensory array limited to touch and sight it would not be clear how this would confuse them into mistaking a succession of perceptions for a simple self. Observing things close together is no more likely to produce the smooth flow of perceptions that tricks the mind than is observing distant points. In as much as Contiguity does have a role here it is as a secondary species of resemblance. As such it has no identifiable independent role.

---

<sup>577</sup> *Treatise* 260:22-261:32 (I, IV, VI, XVII-XIX) Stroud (1977) p. 260n1 chastises Hume for not better explaining his decision. Stroud's criticism is not without point, but ultimately Hume's rejection of contiguity makes little difference to his overall theory.

<sup>578</sup> *Treatise* 239:6-7 (I, IV, V)

<sup>579</sup> *Treatise* 239:19-20 (I, IV, V)

<sup>580</sup> Or a dualism of kinds. For extensive discussion of Hume's possible use of neutral monism see Flage (1982)

<sup>581</sup> *Treatise* 240:1-2 (I, IV, V)

Contiguity in time, of course, can play no role for Hume. Our experience of time is *produced* by the succession of perceptions so any attempt to claim that the smooth flow of perceptions through time causes us to regard the self as simple would be doomed to circularity. All in all Hume is right to disregard the role contiguity might play.

Causation, resemblance and memory are intertwined in Hume's account, so it is necessary to unravel the role each is meant to play. Hume does regard causation as providing the metaphysical underpinning for the succession of perceptions that makes up the bundle of perceptions:

we may observe, that the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other.<sup>382</sup>

For the moment it clear that the mind is a succession of perceptions, one caused by another. However, we have no true knowledge of this connection. Hume wishes to find how it is that we experience this chain of causation, and mistake it for a simple self.

The answer for Hume is clear "As memory alone acquaints us with the continuance and extent of this [causal] succession of perceptions, 'tis to be consider'd, upon that account chiefly, as the source of personal identity."<sup>383</sup> This could be read as Hume claiming that connections of memory produce personal identity, but the textual evidence is clear that Hume does not give memory this role. Instead he says "memory does not so much *produce* as *discover* personal identity, by shewing us the relation of cause and effect among our different perceptions."<sup>384</sup> Hume is in fact conducting an attack on those who would claim that memory "produces entirely our personal identity".<sup>385</sup> These authors cannot explain how we can extend our identity "beyond our memory".<sup>386</sup>

However, there is one comment that might lead one to think Hume does give some role in the production of personal identity to memory. Hume writes "the memory not only discovers the identity, but also contributes to its production, by producing the relation of resemblance among perceptions."<sup>387</sup> This comment occurs before the claim that memory discovers rather than produces *personal identity*. The apparent confusion only arises if one

---

<sup>382</sup> *Treatise* 261:9-13 (I, IV, VI, XIX)

<sup>383</sup> *Treatise* 261:33-35 (I, IV, VI, XX)

<sup>384</sup> *Treatise* 262:16-18 (I, IV, VI, XX)

<sup>385</sup> *Treatise* 262:19-20 (I, IV, VI, XX)

<sup>386</sup> *Treatise* 262:21 (I, IV, VI, XX)

fails to realise that at this preliminary stage Hume is only discussing how memory goes to produce resemblances between perceptions and consequently the false idea of a perfectly identical self. Hence Hume's comment is limited to "this particular"<sup>587</sup>, i.e. the production of the false impression of identity. Memory does not produce *personal* identity, only identity. This also explains why Hume comments directly after that "The case is the same whether we consider ourselves or others".<sup>589</sup> Memory being equally responsible for first and third person bundling rests easier with memory producing the impression of identity than it does with memory producing (even if partially) personal identity. If memory produced personal identity there would be no point in commenting on this symmetry. It would be a trivial truth because all persons are surely similar in this respect.<sup>590</sup> However, this symmetry is worth commenting on if the *impression* of personal identity is formed the same way for first and third-person cases. It is just this that Hume is commenting on.

Hume's discussion of memory begins at the same point the distinction between third and first-person cases begins. This, in itself, marks a crucial turning point, for it is here that Hume leaves behind discussing merely the identity of the mind, and begins to discuss personal identity. Memory allows us to acquire the notion of causation. The notion of causation allows us to develop *two* further notions, those of the mind and the person. Now, it is clear that Hume thinks it is equally true that our notions of the mind and the person have to be acquired in the same fashion. Neither one is found in the senses, or directly by reason. Each is present only in the imagination. By careful direction of the reason the philosopher can expose how we acquire these notions and that neither is native or the result of a simple impression. However, this doesn't mean that Hume thinks that the mind and the person are one and the same. Bringing out the distinctions Hume makes by separating first and third person discussions will demonstrate how Hume distinguishes between the mind and person.

---

<sup>587</sup> *Treatise* 261:4-7 (I, IV, VI, XVIII)

<sup>588</sup> *Treatise* 261:4 (I, IV, VI, XVIII)

<sup>589</sup> *Treatise* 261:7-8 (I, IV, VI, XVIII)

<sup>590</sup> Further, Hume would not want to emphasize that our own and other minds are analogous. This is a presumption of Hume's argument, hence his earlier comment that what follows he confidently asserts to be true of "the rest of mankind" *Treatise* 252:32 (I, IV, VI, IV).

## X

### THE TRUE IDEA OF THE HUMAN MIND

When Hume declares "The case is the same whether we consider ourselves or others"<sup>591</sup> I take this to mean that we perceive ourselves to be identical on the same principles that we observe others to be identical. Of course, this observation of others would *presume* some sort of Compositional Account of personal identity, because we can only observe others' bodies.<sup>592</sup> Hume is perfectly aware that the minds of others are private. For this very reason he begins his first discussion of the role of resemblance by asking us to

[S]uppose we cou'd see clearly into the breast of another, and observe that succession of perceptions, which constitutes his mind or thinking principle.<sup>593</sup>

The faculty which most assists in producing resemblances and therefore the "smooth and uninterrupted progress of thought"<sup>594</sup> is, of course, memory which Hume remarks "'tis to be consider'd, [...] chiefly, as the source of personal identity."<sup>595</sup>

The reading presented here is justified because it first provides an extremely strong version of Hume's theory of personal identity.<sup>596</sup> Further, it allows one to make good sense of Hume's second thoughts in the Appendix. However, this reading does claim Hume *presumed*, to some extent, the Compositional Account. This presumption needs to be justified. This is especially so as other explanations of Hume's choice of terminology are available and other authors have either interpreted these phrases differently, or found it viable to ignore them altogether.

---

<sup>591</sup> *Treatise* 261:7-8 (I, IV, VI, XVIII)

<sup>592</sup> I therefore find Green's following comments utterly mistaken "Hume assumed that personal identity over time is equivalent to the identity of a person's mind over time. He paid no attention at all to the possibility that personal identity over time might consist in the identity of a human body over time." Green (1999) p. 104. I also find myself in specific disagreement with Baier, who comments on "One's body, which Book One's discussion of personal identity ignored" (Baier 1991, p. 130), and that "Book One was also virtually silent about our awareness of fellow persons" Baier (1991), p. 133. In general, though, Baier's account and my own are in sympathy. See Baier (1991), Chapter 6 'Persons and the Wheel of Their Passions'.

<sup>593</sup> *Treatise* p. 260: 28-30 (I, IV, VI, XVIII) Waxman claims "Hume's account of personal identity involves a succession of uniform relations... external objects are at no point presupposed." Waxman (1994) p. 238. Other persons are surely external objects, so Waxman's claim must be wrong. Of course, it might well be part of Hume's later difficulties that he did not wish to presume the existence of external objects. However, his free discussion of other persons suggests he does make such a presumption.

<sup>594</sup> *Treatise* p. 260:19 (I, IV, VI, XVI)

<sup>595</sup> *Treatise* p. 261:34-5 (I, IV, VI, XX)

<sup>596</sup> The measure of strength here is the degree an interpretation relies on unembellished presentation of the text. Therefore, a strong interpretation credits a minimal number of errors and inconsistencies to an author and only attributes opinions obviously explicit in the text.

The Compositional Account has the advantage of being a common-sense theory. Hume wishes to explain how vulgar opinions are arrived at. Therefore, one would expect his account to employ a common-sense theory. Hume, on this consideration, can hardly be thought to defend a Bundle Theory. There is nothing about a Bundle Theory which one is inclined to regard as naïve or common-sense. Hume does hold that the emergence of our notions of identity via the imagination is a complex process. However, the products of the imagination are common-sense objects. In his initial account of identity Hume defends the vulgar belief in enduring external bodies. The same seems likely to be true of his account of personal identity. This is more likely to be a version of the Compositional Account as opposed to a version of a Constructivist Bundle Theory. What is more, there is considerable textual evidence to support this view. Hume repeatedly talks about persons in a manner that presumes they are merely human beings. This evidence will be examined shortly when Hume's second thoughts are discussed.

For present purposes, though, this common-sense defence of the presumption of the Compositional Account stands counter to the most obvious competing explanation for Hume's use of both first-person and third-person terminology. One could merely assume that Hume's choice of terminology is a mere quirk of expression or literary device.<sup>597</sup> The support for this is thin. Hume employs both the terms 'myself' and 'ourselves' without complication. The plural merely serves to assert that what Hume finds in his own experience is, in fact, true of everyone else.<sup>598</sup> If Hume's use of third-person examples were merely a literary device it would surely be intended to serve just this purpose. Yet, as Hume is happy to assert generality using the first-person plural it seems unlikely he would introduce a second device. This seems even more likely as the first/third person shift is an unnecessarily complicated and confusing way to achieve this end. For this reason Hume's habit of phrasing problems in the third-person needs to be taken at face value.

Clearly the present reading is committed to accusing commentators who do not remark on this feature of Hume's account of making an oversight. It therefore remains to look at those commentators who do make some use of the distinction and explore the other positions in logical space that might explain Hume's decision.

Hume could intend a solipsistic idealist reading.<sup>599</sup> Hume himself is a bundle of perceptions bound together by the connexions between his thoughts, and every other person (or

---

<sup>597</sup> This is what I take Stroud to do. See Stroud (1977) p. 129-130.

<sup>598</sup> *Treatise* 252:31-32 (I, IV, VI, IV)

<sup>599</sup> Stroud does seem to go some way to taking just such a possibility seriously; Stroud (1977) p. 137ff

appearance of person) is also a bundle of perceptions bound together in Hume's mind. The only connexions in the world with any efficacy are those found in Hume's own mind. All apparent 'sub-bundles' are entirely illusory. There is only one solipsistic bundle (and this itself is a fiction). Observing regularities in our experience causes us to form (though erroneously) the idea of identity we then apply (again erroneously) to ourselves. This is surely just the sort of extravagant scepticism Hume frequently ridicules. Such a sceptical position is irrefutable within its own scope, but its extremity surely militates against it being Hume's own view. If it were Hume's view, he would surely be more explicit about it. Therefore Hume's terminology should be read at face value.

Hume's employment of the third-person perspective gives a central role to the person as an observed object. Both Bricke<sup>600</sup> and latterly Ward<sup>601</sup> take this implication of Hume's expression very literally. On this interpretation there are two sequences of perceptions that occur: one the sequence of perceptions in an observer and second the sequence of perceptions observed in 'the breast of another'. The regularity of perceptions in the observed person would therefore prompt a like sequence of perceptions in the observer. In pursuing this line both Bricke and Ward generate accounts of unusual complexity and difficulty. Primarily they find it unclear how the observer's bundle can be united as it itself remains unobserved while the criterion of union appears to be observed succession.

This is only a difficulty for Hume on a Bundle Theory reading. It is the attempt to preserve the psychological purity of the bundle reading that drives Bricke and Ward to such awkward lengths when discussing Hume's assertion. If Hume had intended such an analysis comparing the union of observed and unobserved sequences he would surely have been explicit about it. In the analogous discussion of object identity he is keen to urge the usefulness of the distinction between observed successions of perceptions and the interrupted objects. Bricke and Ward are right that Hume is acutely aware of the difficulty of observing the succession that makes up a person. However, the difficulty is not the one they highlight.<sup>602</sup>

The purpose of the discussion of the succession of perceptions in the breast of another should therefore be read at face-value. This is possible if one presumes (with Hume) a Compositional Account of personal identity. First, Hume's request to "*suppose we cou'd*

---

<sup>600</sup> Bricke (1977)

<sup>601</sup> Ward (1985)

<sup>602</sup> Indeed, their discussion ultimately lapses into the extremes of the Strict Idealist Reading sketched above.

see clearly into the breast of another”<sup>603</sup>introduces a thought-experiment.<sup>604</sup> Hume has to ask us to suppose we could see another’s perceptions for the simple and obvious reason that he acknowledges that another’s thoughts are private. Hume’s language consistently supports this reading. The succession we observe “constitutes *his* mind or thinking principle”.<sup>605</sup> If the person were constituted by just the mind it would be strange to talk about the mind being owned by anything (other than, trivially, itself). Hume, of course, doesn’t mean this. Instead he is asking us to consider the part of a human being that thinks. This is why he uses the phrase “thinking principle.”<sup>606</sup>This phrase he uses only on this occasion, its sole purpose being to emphasis that the mind is possessed by the person and does not alone constitute the person. It is not itself a technical term but a one-off elaboration of the meaning of ‘mind’ in this context.

The mind of a person would be endlessly varied *if* that mind lacked the faculty of memory. Hume’s difficulty with mental identity is that the mind spontaneously exhibits no kind of regularity for the imagination to work upon.<sup>607</sup> Hume appeals to a simple copy theory of memory to explain the resemblances that do arise for the imagination to unite into identity.

For what is memory but a faculty, by which we raise up the images of past perceptions? And as an image necessarily resembles its object, must not the frequent placing of these resembling perceptions in the chain of thought, convey the imagination more easily from one link to another, and make the whole seem like the continuance of one object?<sup>608</sup>

Hume then makes his famous declaration that “memory not only discovers the identity, but also contributes to its production”<sup>609</sup> and urges that “The case is the same whether we consider ourselves or others.”<sup>610</sup>

One needs to ask what equivalence Hume is urging here. The identity that memory discovers is importantly *not* personal identity but the identity of the mind. The equivalence Hume is discussing is how we come to believe each person has an identical mind, when only one mind is available to our inspection. Hume would obviously be keen to properly explain the role of memory in personal identity, given the importance of memory in

---

<sup>603</sup> *Treatise* 260:28-29 (I, IV, VI, XVIII) emphasis added

<sup>604</sup> As earlier with the mass of matter *Treatise* 255:30 (I, IV, VI, VIII). See §VI in this chapter.

<sup>605</sup> *Treatise* 260:30 (I, IV, VI, XVIII) emphasis added

<sup>606</sup> *Treatise* 260:30 (I, IV, VI, XVIII)

<sup>607</sup> *Treatise* 252:33-253:4 (I, IV, VI, IV)

<sup>608</sup> *Treatise* 260:34-261:4 (I, IV, VI, XVIII)

<sup>609</sup> *Treatise* 261:5-6 (I, IV, VI, XVIII)

<sup>610</sup> *Treatise* 261:7-8 (I, IV, VI, XVIII)

contemporary discussions.<sup>611</sup> His explanation is very neat. He gives a constituting but non-causal role to memory in the production of mental identity. Memory produces resemblances that the imagination then combines. Therefore Hume shows why we might believe memory constituted personal identity while demonstrating that it only contributes to the production of identity.

Hume's thought experiment shows that we can only realise this about the mind by considering it in third-person terms. The body is crucial in this. If Hume merely asked us to imagine the mind of another he would be doomed to circularity. His definition of the mind is, therefore, *grounded* by the Compositional Account. It is only in this way that Hume can introduce the causal definition of mental identity that will allow him to introduce his account of personal identity.

Hume states:

[T]he true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other. Our impressions give rise to their correspondent ideas; and these ideas in their turn produce other impressions. One thought chases another, and draws after it a third, by which it is expell'd in its turn.<sup>612</sup>

In none of this is there any mention of the person. Commentators have by and large chosen to ignore this fact. However, the reason seems quite simple: Hume is offering us the 'true idea of the human mind'.

The passage following is frequently presented as Hume's definition of personal identity.

In this respect, I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts. And as the same individual republic may not only change its members, but also its laws and constitutions; in like manner the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity. Whatever changes he endures, his several parts are still connected by the relation of causation.<sup>613</sup>

---

<sup>611</sup> *Treatise* 262:15-21 (I, IV, VI, XX)

<sup>612</sup> *Treatise* 261:9-17 (I, IV, VI, XIX)

<sup>613</sup> *Treatise* 261:17-28 (I, IV, VI, XLX)

This passage is a frustrating conundrum for many reasons. Prime among these is Hume's lapse into metaphor at such a crucial moment in his discussion (a lapse prefigured in his comparison of the mind to a theatre<sup>614</sup> earlier). This lack of transparency cannot easily be remedied, although it can be explained.

Hume's mention of the 'soul' seems paradoxical. The soul is berated as a fiction on a par with self and substance.<sup>615</sup> However, there is a precedent for Hume's apparently approving use of the term. Earlier, while commenting on the variability of our perceptions, Hume writes "nor is there any single power of the soul, which remains unalterably the same, perhaps for one moment."<sup>616</sup> Hume maintains that the exact constitution, materials and powers that make up the mind are all equally unknown. Hume can consistently berate and employ the notion of the soul if what he objects to about the soul is its supposed explanatory role. This explains why Hume employs metaphor. By his own lights Hume cannot say what the mind or soul actually are. He can only explicate them, therefore, metaphorically. Only metaphor can avoid metaphysical commitment on this delicate matter. It is in this decidedly Lockean spirit that Hume develops a causal account of mental identity and then projects this to provide a causal account of personal identity. These two mentions of the soul are also typical of the deflationary way Hume occasionally employs his opponents' terms. Hume indirectly equates the soul and the person here, the effect being to deflate the notion of the soul into the Humean notion of person. Likewise the earlier mention of the powers of the soul deflated these mysterious entities into Humean perceptions.

However, one is not yet out of the interpretative woods. Comparison and metaphor are not definition. Therefore Hume's assertion that a person's identity may survive certain changes in a like manner with a republic does not imply an exhaustive definition of person. Two factors point towards this metaphor not being the full extent of Hume's definition.

Firstly Hume reminds us of the earlier distinction he had drawn between "personal identity, as it regards our thought or imagination, and as it regards our passions or the concern we take in ourselves."<sup>617</sup> He writes:

And in this view our identity with regard to the passions serves to corroborate that with regard to the imagination, by the making our distant perceptions

---

<sup>614</sup> *Treatise* 253:4-7 (I, IV, VI, IV)

<sup>615</sup> *Treatise* 254:27 (I, IV, VI, VI)

<sup>616</sup> *Treatise* 253:2-4 (I, IV, VI, IV)

<sup>617</sup> *Treatise* 253:19-21 (I, IV, VI, V)

influence each other, and by giving us a present concern for our past or future pains or pleasures.<sup>618</sup>

Secondly the republic metaphor does not suggest that the extension of identity beyond our memory to “times, and circumstances, and actions, which we have entirely forgot, but suppose in general to have existed”<sup>619</sup> is a purely psychological matter.

The passions belong to the body and to the whole person. Personal identity as it concerns the passions presumes the Compositional Account. As Baier notes, “To become recognizable, persons must become incarnate, and in Book Two they are treated as ordinary persons of flesh and blood, whose self-concern includes concern about their offspring, their address in dancing, their strength and vigor, their communicable diseases, as well as about their reputation, their wit and virtues.”<sup>620</sup> Hume draws the reader’s attention to this division once again to remind us that the account of personal identity he is presenting is *systematically incomplete*. His interest is with how persons are known in thought, not how persons actually are. Hume fully presumes the metaphysical person to be a human being composed of a mind and a body. The account of personal identity as regards the imagination, though, is one view<sup>621</sup> of the composed person *qua* ‘thinking being’. Hume is mounting an attack on introspection as a source of personal identity, and putting in its place an account that will premise the discussions of Book II.

The metaphor of the republic is by Hume’s own admission incomplete. He remarks that ‘I cannot compare the soul more properly to any thing than to a republic or commonwealth’, suggests that although the comparison is the best available it is not perfect. Given Hume’s propensity to comment on the completeness, perfection and entirety of relations his failure to do so on this occasion further suggests his ultimate dissatisfaction with the metaphor. The republic metaphor should also be compared with Hume’s earlier metaphor of the mind as theatre.<sup>622</sup> There Hume urged that “The comparison of the theatre must not mislead us.”<sup>623</sup> The theatre metaphor suggests we might know the location and composition of the mind, and Hume is quick to remind us of our perfect ignorance in this matter.<sup>624</sup>

---

<sup>618</sup> *Treatise* 261:28-32 (I, IV, VI, XIX)

<sup>619</sup> *Treatise* 262:6-7 (I, IV, VI, XX)

<sup>620</sup> Baier (1991) p. 131

<sup>621</sup> Recall Hume noted identity itself is prone to the view we take of it in his original account. *Treatise* 201:21 (I, IV, II)

<sup>622</sup> *Treatise* 253:4-14 (I, IV, VI, IV)

<sup>623</sup> *Treatise* 253:10 (I, IV, VI, IV)

<sup>624</sup> *Treatise* 253:12-14 (I, IV, VI, IV)

Likewise the republic metaphor is suggestive of the composition of persons. Hume's emphasis in the metaphor is clearly on organisation of a person's "parts" as he terms them.<sup>625</sup> He describes "reciprocal ties"<sup>626</sup>, that "propagate"<sup>627</sup> and "give rise"<sup>628</sup> to each other. The republic achieves this by "laws and constitutions"<sup>629</sup> the person by "character and disposition."<sup>630</sup> This description is conspicuously Lockean in its emphasis on continuity and deliberate agnosticism about the exact nature of the 'parts' involved. Hume's only addition is to insist of a person that "Whatever changes he endures, his several parts are still connected by the relation of causation."<sup>631</sup>

There is nothing in all of this to suggest a purely psychological constitution of persons. Hume again comments it is memory that without which "we shou'd never have any notion of causation, nor consequently of that chain of causes and effects, which constitute our self or person."<sup>632</sup> Nowhere in the *Treatise* does Hume suggest that only psychological entities have causal efficacy, so there is nothing about Hume's definition that excludes the physical from his understanding of person.

In the passage that discusses the extension of identity beyond memory Hume emphasises not just the supposition of forgotten "thoughts"<sup>633</sup> but repeatedly of forgotten "actions".<sup>634</sup> The natural way to understand action here encompasses corporeality. Hume's use of the Lockean distinction between self and person in this passage reinforces this impression:

Or will he affirm, because he has entirely forgot the incidents of these days, that the present self is not the same person with the self of that time; and by that means overturn all the most establish'd notions of personal identity?<sup>635</sup>

The Lockean momentary self is constituted by whatever substances coincide at that moment. These are united over time into one person. In this way Locke (and Hume) can have us completely self-aware at a time while allowing for diachronic amnesia without compromising personal identity. Like Locke, Hume has explained how we arrive at the

---

<sup>625</sup> The republic's 'parts' *Treatise* 261:22 are compared with the person's parts *Treatise* 261:27 (I, IV, VI, XIX)

<sup>626</sup> *Treatise* 261:19 (I, IV, VI, XIX)

<sup>627</sup> *Treatise* 261:21 (I, IV, VI, XIX)

<sup>628</sup> *Treatise* 261:20 (I, IV, VI, XIX)

<sup>629</sup> *Treatise* 261:23-24 (I, IV, VI, XIX)

<sup>630</sup> *Treatise* 261:25 (I, IV, VI, XIX)

<sup>631</sup> *Treatise* 261:26-28 (I, IV, VI, XIX)

<sup>632</sup> *Treatise* 261:36-262:2 (I, IV, VI, XX)

<sup>633</sup> *Treatise* 262:10 (I, IV, VI, XX)

<sup>634</sup> *Treatise* 262:6, 262:8, 262:10 (I, IV, VI, XX)

<sup>635</sup> *Treatise* 262:11-15 (I, IV, VI, XX)

notion of believing that the experience we have currently of ourselves is identical with experiences in the past, and those yet to happen in the future.

In this manner Hume feels he avoids overturning 'the most establish'd notions of personal identity'. If this is Hume's aim it seems utterly implausible that his aim is to establish a revisionary metaphysical account of personal identity such as the bundle theory. Nor is it likely the 'establish'd notions' Hume alludes to are those of philosophy. Philosophy he finds vague and elusive. Instead they are the principles of common-sense and the imagination. Hume's explicit aim is to provide an account of personal identity consistent with common belief. This common belief is of a person being perfectly identical and perfectly simple. The imagination is naturally disposed to create the illusion of the introspective Metaphysicians' Self. This naturally biases us to believe we are constituted by something ethereal and mental, such as a soul or self.<sup>636</sup> This is merely an accident of the imagination itself being a mental process. However, Hume shows that this fallacy is dependent on the observation of other corporeal beings, and itself systematically presumes a causal chain dependent on the physical. Hume himself remarks we must explain "that identity, which we attribute to plants and animals; there being a great analogy betwixt it, and the identity of self or person."<sup>637</sup> The natural fiction of identity we all experience, then, is properly truly causal at basis and based on both psychological and physical regularities. In other words, Hume takes the most established notion of personal identity to be the Compositional Account.

If one merely relied on the chapter 'Of Personal Identity' to establish the grounding role of the Compositional Account the evidence would always appear oblique. Hume's words allow a great number of interpretations to be foisted upon them. This is why there are almost as many opinions of Hume as there are commentators. However, the role of the Compositional Account can be further secured in a manner other readings cannot. Firstly, only by presuming the Compositional Account can one explain Hume's doubts in the Appendix in a suitably straightforward and satisfactory manner. Secondly, only the Compositional Account chimes in easily with Hume's aims in Books II & III. Thirdly, and most interestingly, only by adopting this strategy can Hume's exclusion of personal identity from his later writings be adequately explained. The next two sections will flesh this out. The first will demonstrate the inadequacy of existing explanations and establish what

---

<sup>636</sup> Though, belief in the soul is not inevitable. Belief in an identical self, however, *is*. Hume urges both the importance and degree of these variant fictions at *Treatise* 254:11-255:5 (I, IV, VI, VI) & 262:32-36 (I, IV, VI, VI)

<sup>637</sup> *Treatise* 253:23-25 (I, IV, VI, V)

should constitute a useful strategy for solving this conundrum in Humean scholarship. The second will demonstrate the role of the Compositional Account in Hume's second thoughts in the Appendix and trace these themes in the *Enquiry*.

## XI

### HUME'S MISSING THIRD PRINCIPLE

For Hume the search for a satisfactory account of personal identity proved an impossible one. He laments in the *Appendix* that "I cannot discover any theory which gives me satisfaction on this head."<sup>638</sup> A considerable amount of philosophical effort has gone into trying to explain why Hume casts doubt over his own theory. Cause of so much of this effort has been Hume's allegedly obscure statement of his own doubts.

It would seem that in attempting to understand a text good practice would demand one accept an obvious reading over an obscure one. The very nature of the *Appendix*, though, tempts the abuse of this simple practice. Hume's initial account is itself difficult. Many readers find it the source of a multitude of absurdities, suggestive errors, paradoxes and red herrings. When Hume expresses his own discontent the temptation to project one's own misgivings onto the text has proved overwhelming. Hume tempts this because the *Appendix* presumes a careful reading of the original text. This is a presumption Hume sadly cannot afford. Regrettably a tradition has emerged that Hume has failed to make himself clear. This further encourages flights of fancy. The *Appendix* is taken as a declaration of open season on any part of Hume's original account.

The chapter 'Of Personal Identity' has frequently been misread because its role as a part of Book I is not fully appreciated (not to mention the relationship with Books II & III). Equally the *Appendix* passages, short as they are, have received a very selective reading.

In the *Appendix* Hume puts forward two principles that he claims he is unable to make consistent:

*that all our distinct perceptions are distinct existences*

and

*that the mind never perceives any real connexion among distinct existences.*<sup>639</sup>

Now it is clear (as many have been at pains to point out)<sup>640</sup> that the two principles are not inconsistent with each other. It must, therefore, be some further principle that Hume finds these inconsistent with.

---

<sup>638</sup> *Treatise* 636:2-3 (App., XI)

<sup>639</sup> *Treatise* 636:6-8 (App., XII)

Hume does not then list this third principle with equal clarity. It is this state of affairs that promulgates endless Humean archaeology and has produced more than a little Humean alchemy. This poses a simple question, which is almost never asked. Hume goes to the effort of adding an Appendix to the *Treatise* that contains the candid confession of the inadequacies of his previous theories. This would be a remarkable act by an insignificant philosopher. It is a monumental one for a major philosopher, especially one as supposedly concerned with public appearances as Hume. Given this effort of calling attention to a difficulty, it seems hardly tenable to believe Hume would find himself unable to state the problem. Hume complains he can find no 'theory' to satisfy him. He does not complain that he cannot articulate his problems. Therefore, one should prefer to believe that Hume *did* clearly state his problem. Considerable and overwhelming evidence is needed to justify any contrary conclusion.<sup>641</sup>

Despite the heavy weather that has been made of this Hume is, in fact, very clear as to what his problems are. In the paragraphs previous to the two principles Hume does not feel "it in my power to renounce"<sup>642</sup> he outlines his difficulties.

The opening paragraphs of the *Appendix* restate Hume's contention that all perceptions are separable. Hume principally, though, draws attention to the parallels in his philosophy between the identity of bodies and the identity of mind:

I HAD entertain'd some hopes, that however deficient our theory of the intellectual world might be, it wou'd be free from those contradictions, and absurdities, which seem to attend every explication, that human reason can give of the material world.<sup>643</sup>

Philosophers begin to be reconcil'd to the principle, *that we have no idea of external substance, distinct from the idea of particular qualities*. This must pave the way for a like principle with regard to the mind, *that we have no notion of it, distinct from the particular perceptions*.<sup>644</sup>

---

<sup>640</sup> Stroud (1977) p. 128, Scarre (1983) p. 218, Garrett (1997) p. 167.

<sup>641</sup> I find Garrett's solution, despite its inventiveness, unsuccessful merely on grounds of its complexity and degree of removal from Hume's statements in the Appendix. See Garrett (1997) p. 180-185. Garrett argues that his account deals with each element of Hume's description of the problem and describes what should have been a real worry for Hume. However, I feel Garrett misses some of Hume's clear statements of his problems, and the worry he describes to in fact be one Hume would not have worried about. Briefly, Garrett asserts that Hume has a problem with possible co-ownership of perceptions. In fact the Humean notion of intimacy absolutely guarantees against this possibility.

<sup>642</sup> *Treatise* 636:5 (App., XII)

<sup>643</sup> *Treatise* 633:10-14 (App. I)

<sup>644</sup> *Treatise* 635:10-14 (App., X)

It is these two statements that are the crux of Hume's confusion. The scepticism Hume can tolerate about knowledge of bodies he cannot tolerate about persons. The argument progresses as follows.

Hume's initial account is defective on two fronts; it does not "explain the principle of connexion, which binds them together",<sup>645</sup> nor what "makes us attribute to them a real simplicity and identity".<sup>646</sup>

The main aim of the original chapter 'Of Personal Identity' was to explain the propensity to believe the mind was a simple substance. The Appendix laments the lack of explicit positive theory about the nature of the successions of perceptions Hume felt were wrongly described as simple.

The effect of his account, Hume suggests, is to allow that "thought alone finds personal identity".<sup>647</sup> Hume notes that "[h]owever extraordinary this conclusion may seem, it need not surprise us"<sup>648</sup> because "[m]ost philosophers seem inclin'd to think, that personal identity *arises* from consciousness".<sup>649</sup> Hume too accepts that the *idea* of personal identity arises in consciousness. However, he finds it difficult to accept that personal identity is *produced* by consciousness, or *constituted* by consciousness.

The crux of the problem then is a simple one. Hume can give an account of personal identity that works for others and for ourselves when we think of ourselves as others see us. This public person can be discovered. However, this means that a thinking being does not know itself. This appears to be a sceptical leap too far. Hume can accept that bodies, being only sporadically and partially present to us, are constructed by the imagination. Hume can also accept that there is a sense in which personal identity is like this. However, he cannot shake the ghost of the Metaphysicians' Self, that the self should be intimately known in introspection. This is the paradox of The Humean Insight. Hume rejects the Metaphysicians' Substantial Self because it is unknowable and mysterious like all substances. However, Hume's empirical urge that the person should be easily comprehensible is not satisfied by his own theory. Hume continues to search within himself for an answer, despite his own best arguments that he will find nothing this way. Introspection is a natural mode of thinking, practised by the vulgar, but it is one Hume finds

---

<sup>645</sup> *Treatise* 635:17-18 (App., XI)

<sup>646</sup> *Treatise* 635:18-19 (App., XI)

<sup>647</sup> *Treatise* 635:27 (App., XI)

<sup>648</sup> *Treatise* 635:30-31 (App., XI)

<sup>649</sup> *Treatise* 635:31-33 (App., XI)

himself unable to defend as he might wish. His pleading the privilege of the sceptic has for once the tone of exasperation, even desperation.

This reading has the virtue of being simply and obviously present in the text. It is also pitched at the right level. Hume is clearly aware where his problem lies, but is unable to provide a solution to it. An interpretation of Hume is entitled to intuit what Hume's solution (or problem) might have been. This does permit the attribution of extra unstated beliefs to Hume. However, these should be minimal in number, and justified by the text. The Humean Insight is one such extra belief. However, it is a minimal addition, and is widely justified by the text, as it synthesises Hume's concerns about introspection. It is also, crucially, compatible with the complaints Hume does express. It ignores none of them, and does not foist extra ones upon Hume. No other treatment of Hume's Appendix does so little violence to Hume's text while providing such a powerful insight into his philosophy.

Clearly my interpretation is in sympathy with the one offered by Kemp-Smith. Kemp-Smith makes particular use of Hume's treatment of the self in Book II, especially this passage:

'Tis evident, that the idea, or rather impression of ourselves is always intimately present with us, and that our consciousness gives us so lively a conception of our own person, that 'tis not possible to imagine, that any thing can in particular go beyond it.<sup>650</sup>

Kemp-Smith suggests that the Appendix is Hume's confession of the tension between this intimate presence and the lack of any one impression of the self Hume insisted upon in Book I. Where my interpretation differs from Kemp-Smith is in the detailed reliance on the Compositional Account. However, objections to Kemp-Smith's solution might also threaten my account. Garrett objects that "there simply is no inconsistency between Book I and II to be the object of Hume's concern. He is quite entitled to write in Book II of both an idea and an impression of 'ourselves'."<sup>651</sup> Garrett is correct to note that Hume does allow there is an impression of the self. Hume's dispute is merely about what we gain from this impression and how it arises. Garrett himself points out Hume gives other examples of impressions that violate the Copy Principle (time, space, the missing shade of blue). Garrett is mistaken, though, to assert that this is an objection to the spirit of Kemp-Smith's claim. Hume does sense a tension between the intimacy we have of ourselves, and the manner of

---

<sup>650</sup> *Treatise* 317:26-30 (II, I, XI)

our awareness of ourselves as spelt out in Book I. That is to say, Hume recognises the tension between his analysis of the failure of introspection and intimate consciousness. Introspection, not an impression of the self, is the real crux of Hume's difficulty. The following will build on this, in the spirit of Kemp-Smith, despite Garrett's misgivings.<sup>651</sup>

Hume emphasises that his dissatisfaction stems from the relative failure of his account to explain the intellectual world compared with the material world. In part this might amount to an analogous failure to explain the mind as an object. Material objects are external and observable, whereas the mind must observe itself. Once again this is connected to the failure of introspection. This problem is resolved in my interpretation by taking the mind to be part of something material and observable, that is a human being. Hume can be argued to settle for this solution, although it does not entirely allay his fears about the failure of introspection.

There is, however, a school of interpretation that rests the entirety of Hume's difficulties on the failure of the self to be properly an object to itself.<sup>652</sup> If the mind is a bundle, it is argued, it cannot be an active self. A bundle contains only perceptions and cannot itself believe, will, act, remember and so forth. Instead it only contains beliefs, volitions, memories and so forth as members. Clearly this is no difficulty for the Compositional Account as it is the human being that remembers, wills and acts. However, for those who mistake Hume as holding a purely psychological theory, such an answer is not available. This approach is right to point out that Hume shows that the self cannot be expansively present to itself because all that can be known at any one time are certain perceptions. This, however, is Hume's very point. The denial of introspection is a central plank of Hume's analysis, and one he does not abandon. It can be argued that Hume should have held a different theory of introspection, and this is why his initial account failed. However, it cannot be argued that his problem in the Appendix is how a bundle introspects.

An allied school emphasises Hume's disappointment at not finding the true relation that binds perceptions. This interpretation is particularly popular with those who are impressed

---

<sup>651</sup> Garrett (1997) p. 168

<sup>652</sup> Garrett also objects that on Kemp-Smith's analysis Hume should reject Book I's account, but in fact continues to express his satisfaction with it. However, this does not carry much weight at all. Hume's confession makes *more* sense because he maintains his belief in Book I's account of the self. Otherwise Hume would have been able to write a simple recantation of Book I in the Appendix, and this he does not do. Further Hume can be satisfied that Book I's account is sound *as far as it goes*. Hume's stated problem is that the system which explained the material world does not explain the intellectual world to his satisfaction. Book II exposes a shortfall in Book I's ambitions. Hume's problem is not that the Books conflict (as Garrett suggests Kemp-Smith must maintain) but that they do not adequately support each other. There is nothing about the latter problem that demands Hume must reject one account in the favour of the other.

with the constructivist tones in Hume's writing. In its weak form Hume is taken to express dismay as to how bundles are held together.<sup>654</sup> In its strong form Hume is taken to flounder by explaining where one bundle begins and another ends.<sup>655</sup> This line of enquiry contains a basic insight, but is far from being Hume's main worry. The true nature of causal bonds Hume contentedly maintains is mysterious. He is equally sanguine in accepting that we do not know the materials of the mind. Hume does not expect to explain these things. Therefore any solution which rests on Hume allegedly expressing dissatisfaction with resemblance and causation as bonds is quite mistaken.<sup>656</sup> Hume never rejects this analysis. However, Hume does have qualms about our mode of knowledge of ourselves given these limits. He feels it must have some special guarantee, hence his urging its intimacy. The solution to the Appendix lies with understanding the nature of this intimacy and what Hume feels it achieves. This intimacy is not, however, explained at all by examining the structure of the bundle of the mind. There is nothing in the Appendix to suggest that Hume does reject his basic conception of the mind as a bundle of perceptions. Although he queries how far his account goes in explaining matters he does not reject the analysis of the bundle, and indeed relies on it in his explanation.

None of these prevailing trends goes to the heart of Hume's worries. In the next section the preferred interpretation sketched above will be fleshed out further, and hopefully will go to the heart of Hume's worries.

---

<sup>653</sup> Robison (1974), Stroud (1977), Nathanson (1976).

<sup>654</sup> Beauchamp (1979)

<sup>655</sup> Stroud (1977)

<sup>656</sup> Basson (1958), Patten (1976b), Stroud (1977), Pears (1993)

## XII

### HUME AND THE COMPOSITIONAL ACCOUNT

Confessedly the direct evidence for Hume's use of the Compositional Account in the chapter 'Of Personal Identity' is oblique. Primarily it rests on claiming Hume held our notion of a person to arise only because we assume that our third-person observation of persons must yield a notion consistent with our first-person experience. We observe other human beings, Humean and Lockean thinking beings, to be identical in the same manner as other objects. This identity we impute to ourselves as like thinking beings. It is by this analogy that we confidently extend our identity beyond experience and memory. One can fairly ask if Hume gestures towards such opinions in other parts of his writings.

The Compositional Account does in fact seem to be a presumption of the entirety of the *Treatise*. Hume does not indulge in a Cartesian project of scepticism. He does not confirm and re-establish the value of each of our senses and faculties as he progresses. Instead he assumes their presence. For example he stresses the role of the senses in acquiring new ideas and impressions: "To give a child an idea of scarlet or orange, of bitter or sweet, I present the objects"<sup>657</sup> and observes the impossibility of acquiring certain ideas when "one is born blind or deaf".<sup>658</sup>

Later in the *Treatise* Hume goes much further towards embracing the Compositional Account. He relies on the physiology of the brain to explain the mistakes of the imagination.

When I receiv'd the relations of *resemblance*, *contiguity*, and *causation*, as principles of union among ideas, without examining into their causes, 'twas more in prosecution of my first maxim, that we must in the end rest contented with experience, than for want of something specious and plausible, which I might have display'd on that subject. 'Twou'd have been easy to have made an imaginary dissection of the brain, and have shewn, why upon our conception of any idea, the animal spirits run into all the contiguous traces, and rouze up the other ideas, that are related to it. But tho' I have neglected any advantage, which I might have drawn from this topic in explaining the relations of ideas, I am afraid I must here have recourse to it, in order to account for the mistakes that arise from these relations. I shall therefore observe, that as the mind is endow'd with a power of exciting any idea it pleases; whenever it dispatches the spirits into that region of the brain, in which the idea is plac'd; these spirits always excite the idea, when they run precisely into the proper traces, and

---

<sup>657</sup> *Treatise* 5:5-6 (I, I, I)

<sup>658</sup> *Treatise* 5:24 (I, I, I)

rummage that cell, which belongs to the idea. But as their motion is seldom direct and naturally turns a little to the one side of the other; for this reason the animal spirits, falling into the contiguous traces, present other related ideas in lieu of that, which the mind desir'd at first to survey. This change we are not always sensible of; but continuing still the same train of thought, make use of the related idea, which is presented to us, and employ it in our reasoning, as if it were the same with what we demanded. This is the cause of many mistakes and sophisms in philosophy; as will naturally be imagin'd, and as it wou'd be easy to shew, if there was occasion.<sup>659</sup>

This passage makes it quite clear that Hume connects the very actions of the mind with the body. As already noted Hume follows Locke in trying to develop a notion of the mind that cannot be undermined by our lack of knowledge about the actual materials of the mind. Locke arrived at his notion of extended consciousness. Hume declares that our only true (i.e. incorruptible) notion of the mind is a bundle of perceptions. The true notion of personal identity is of such a mind as part of a thinking being, i.e. a human being. Hume continues his scepticism in the *Enquiry* where he remarks:

The same difficulty occurs in contemplating the operations of mind on body – where we observe the motion of the latter to follow upon the volition of the former, but are not able to observe or conceive the tie which binds together the motion and volition, or the energy by which the mind produces the effect.<sup>660</sup>

Hume, of course, accepts that the mind and body do interact. He is merely sceptical about our hopes of ever understanding how this comes about.

It is here that Hume's twin concerns of introspection and intimate consciousness meet. Hume insists, quite correctly, that we are aware of the contents of our own minds. What is more, this awareness is of a special order. Hume accepts freely that first-person experience is quite different to third-person experience. It is part of human nature to experience a peculiar relationship with only our own perceptions. This is what Hume hopes to capture by the notion of intimacy. Intimacy is neither deliberate or under our control. It is a basic fact of being a thinking being. Introspection, in contrast, is in our power. We notice introspection more than the ongoing intimacy of ourselves because introspection is under our control. However, introspection is not stronger than intimacy, but weaker. Without mental effort introspection can discover nothing. In periods of sleep introspection is quite inert. Despite its promise of penetrating understanding introspection cannot go beyond the

---

<sup>659</sup> *Treatise* 60:21-61:15 (I, II, V)

<sup>660</sup> *Enquiry* 74:2-7 (VII, II)

simple fact of intimacy. Hume cannot help but conclude that introspection is impotent in this way. However, he never quite accepts this outcome.

If Hume's worry was with the failure of introspection given the intimate consciousness we enjoy of ourselves one might expect these themes to reappear in the *Enquiry*. Notoriously the section on personal identity has no direct successor in the *Enquiry*. However, the themes and attacks are present in the manner one would expect if Hume had resolved his worries into a less thorough acceptance of the Compositional Account.

The theme of intimate consciousness is reasserted. However, the person of Book I of the *Treatise* has now obviously become the Compositional person:

It may be said, that we are every moment conscious of internal power; while we feel, that, by simple command of our will, we can move the organs of our body, or direct the faculties of our mind. An act of volition produces motion in our limbs, or raises a new idea in our imagination. This influence of the will we know by consciousness. Hence we acquire the idea of power or energy; and are certain, that we are ourselves and all other intelligent beings are possessed of power. This idea, then, is an idea of reflection, since it arises from reflecting on the operations of our own mind, and on the command which is exercised by will, both over the organs of the body and the faculty of the soul.<sup>661</sup>

However Hume remains equally clear that "the energy, by which the will performs so extraordinary an operation... must forever escape our most diligent enquiry."<sup>662</sup> Hume remains equally clear that introspection cannot reveal anything further either. The discussion of introspection, in the guise of consciousness, argues:

[I]s there any principle in all nature more mysterious than the union of soul with body...? [I]f by consciousness we perceived any power or energy in the will, we must know this power; we must know its connexion with the effect we must know the secret union of soul and body, and the nature of both these substances; by which the one is able to operate, in so many instances, upon the other.<sup>663</sup>

These things, of course, all remain secrets. The extent of consciousness runs no further than the exercise of the will and the contents of our thoughts. Hume continues this attack by noting other occasions when the interaction of mind and body produce effects we are all familiar with but in the dark as regards ultimate causes. We do not know how the soul

---

<sup>661</sup> *Enquiry* 64:13-24 (VII, I)

<sup>662</sup> *Enquiry* 65:5-8 (VII, I)

<sup>663</sup> *Enquiry* 65:9-22 (VII, I)

produces new ideas.<sup>664</sup> We experience the interaction of soul and body as being limited, a limit we learn only by experience.<sup>665</sup> Our 'self-command' varies with health and practice.<sup>666</sup>

Hume also is quite sure that humans are continuous with the animal kingdom. Thinking is something animals do. Hence his remarks in the Appendix:

We can conceive a thinking being to have either many or few perceptions. Suppose the mind to be reduc'd even below the life of an oyster. Suppose it to have only one perception, as of thirst or hunger. Consider it in that situation. Do you conceive any thing but merely that perception? Have you any notion of *self* or *substance*? If not, the addition of other perceptions can never give you that notion.<sup>667</sup>

Hume's point is simple. Animals have perceptions in greater or lesser numbers. There is nothing about a creature (even a human) having many perceptions that adds anything extra to that creature's being.

It is no surprise to find Hume remark in the *Enquiry* of the similarity of our thinking parts with that of animals.

[T]he experimental reasoning itself, which we possess in common with beasts, and on which the whole conduct of life depends, is nothing but a species of instinct or mechanical power, that acts in us unknown to ourselves.<sup>668</sup>

Hume also remarks on the analogy between like creatures, noting how "readily and universally do we acknowledge a uniformity in human motives and actions as well as in operations of body."<sup>669</sup> Thus Hume establishes the principle by which we can predict claims about personal identity. The basis for such claims is our body of knowledge about Human Beings. Hume writes of man that "as we are otherwise acquainted with the nature of the animal, we can draw a hundred inferences concerning what maybe expected from him; and these inferences will all be founded in experience and observation."<sup>670</sup>

It is in these remarks that Hume latterly dissipates his concerns about personal identity. The advantage of this reading is that it explains what became of Hume's worries while illustrating from where they emerged.

---

<sup>664</sup> *Enquiry* 68:3ff (VII, I)

<sup>665</sup> *Enquiry* 68:15-25 (VII, I)

<sup>666</sup> *Enquiry* 68:26-69:2 (VII, I)

<sup>667</sup> *Treatise* 634:28-34 (App. VII)

<sup>668</sup> *Enquiry* 108:10-14 (IX)

<sup>669</sup> *Enquiry* 84:28-30 (VIII, I)

<sup>670</sup> *Enquiry* 144:4-8 (XI)

The Humean Insight brings together these concerns. Hume feels that personal identity should be well-grounded and easily known. Hume notes the intimacy each of us has with ourselves. However he cannot progress beyond asserting this as a primitive fact. Introspection would appear to be the faculty that would reveal more about personal identity. Hume, however, finds introspection lacking. Therefore he is left defending the most obvious account of personal identity, the Compositional Account. Hume, in providing an account of Human Nature unsurprisingly presumes the existence of human beings as the basis of his analysis.

Noonan remarks of Hume that “The only problem that exists [for Hume] is the *genetic* one of specifying the psychological causes of the universal but mistaken belief in the existence of enduring persons.”<sup>671</sup> Noonan’s observation is typical. It seems obvious that anyone adopting such an approach must be making an *assumption* about what persons are. Any other enquiry would be pointlessly circular. It seems equally obvious that one must conclude that Hume presumed a version of the Compositional Account.

The Humean Insight finds ghostly and gnomonic expression in one curious remark in the *Enquiry*.

It is remarkable concerning the operations of the mind, that, though most intimately present to us, yet, whenever they become the object of reflexion, they seem involved in obscurity; nor can the eye readily find those lines and boundaries, which discriminate and distinguish them. The objects are too fine to remain long in the same aspect and situation; and must be apprehended in an instant, *by a superior penetration, derived from nature, and improved by habit and reflexion.*<sup>672</sup>

Hume once again remarks on intimacy. However, he cannot abandon introspection. The exertions of the *Treatise* had persuaded Hume that introspection did not provide a reliable or pelucid route to self-knowledge. Consequently, it receives no systematic treatment in the *Enquiry*. However, Hume cannot bring himself to totally abandon introspection and the power of reflexion. Hence his hopes for a superior penetration into thought, furnished by nature and perfected by philosophy. It is this cherished hope that is the secret of Hume’s confusions over personal identity.

---

<sup>671</sup> Noonan (1989) p. 78

<sup>672</sup> *Enquiry* 13:7-15 (I, I) emphasis added.

## CONCLUDING POSTSCRIPT

The previous chapters have established the historical support for the Compositional Account. As already noted, there is an affinity between the Compositional Account and the present-day theory of personal identity known as Animalism. Both theories identify persons and human animals. However, Animalism prefers a biological understanding of the human animal. The Compositional Account, on the other hand, analyses the notion of a human being as a composition of a mind and a body. As a consequence the Compositional Account is a broader church, capable of uniting metaphysics as diverse as those of Descartes, Locke and Hume. This postscript will explore the links between Animalism and the Compositional Account and also outline the strengths and value of both theories in contemporary debate about personal identity. The aim of this postscript is not to be definitive, but programmatic, and suggest what the Compositional Account brings to the personal identity debate.

## I

### DIFFICULTIES IN FRAMING THE PROBLEM

The problem of personal identity is the putative philosophical problem of what it is to be and remain the same person over time. On the presumption that each of us thinks we are a person it is also, therefore, a problem of what it is to be the sort of thing that *we* are. This is central to the difficulty of pursuing the philosophical problem of personal identity. It is only on a presumption of how to answer the question '*what are we?*' that debate can proceed. Yet, of course, this is to deeply prejudice the debate.

Two immediate possibilities confront us. One is to assert that persons represent a primitive and irreducible category. There is nothing philosophically dubious about this position. However, such quietism should only be adopted if all other possibilities have been exhausted. The other alternative is to adopt the most obvious answer to the question and take this as a starting point for our enquiry. The most obvious answer to the question '*what are we?*' is to answer that we are human animals. This may itself turn out to be a primitive fact. However, the notion is clearly open to philosophical analysis. We assume ourselves to be persons, and we also assume ourselves to be human animals. Our enquiry then turns to reasons for denying an identification of the two categories.

Two basic objections arise. First, one may object that the identification is not necessary. This objection itself has two forms. The identification may be too narrow. That is, there may be persons that are not human animals. Secondly, the identification may be too broad. That is, there may be human animals that do not qualify as persons. The second objection is that the identification is not sufficient. Being a human animal is not in itself sufficient to substantiate a claim of personhood.

All these objections arise because it appears possible to find differences between our pre-philosophical notions of persons and human animals.<sup>673</sup> It is instructive to try to establish why this pre-philosophical cleave occurs, and further why we regard it as important enough to warrant philosophical grounding.

---

<sup>673</sup> Some people object to the idea of being called an animal altogether. These people normally feel that being human by definition is not to be an animal (and *vice versa*). However, by animal all that is meant is something biological. Very few people object to the claim that we are not in some sense biological, even if this is only a matter of contingent embodiment. By Human I mean a member of the species currently represented by *Homo Sapiens*.

As we see ourselves as paradigmatically persons the answers to the questions as to what it is to be human and what it is to be a person must be compatible. Otherwise humans are excluded from being persons. This presents a dilemma. On the one hand we wish to escape giving an account of personhood that is nothing more than a description of what it is to be a human person. There is something philosophically jejune about giving a merely descriptive account of surface qualities. Nor would it be particularly satisfactory to give an account that allowed non-human persons honorary status as persons in as much as they could be shown to be 'like us'<sup>674</sup>. This would fail to make 'person' the right sort of taxonomic distinction to have any useful role<sup>675</sup>. Personhood must be defined in suitably neutral terms. On the other hand, such neutrality risks excluding some (or even all) humans from being counted as persons. At the very least this would seem self-defeating and absurd, at the worst pernicious. There is further attendant risk of admitting entities that one would not wish to regard as persons at all. Such liberalism threatens both absurdity and triviality.

The philosophical procedure adopted in response to these legitimate pressures has been to abstract *part* of what is seen as being typical of human existence and take this as being essentially constitutive of personhood. Hence all humans will have this capacity (at least potentially), and other non-human entities can (or conceivably might) have it as well.

Human persons have mental and physical qualities. Theories of personal identity have been developed in terms that assume this mind-body distinction. Any doubts readers might have entertained about this presumption should have been allayed in this thesis. The Compositional Account itself presumes the mind-body distinction. The previous chapters have argued for its historical influence. The point of interest here is to explain why the psychological aspects of human experience have been emphasised in recent theories of personal identity to the exclusion (already commented on in the introduction) of corporeal existence.

The reason is seems simple and obvious enough even if its consequences have been neither. Gross corporeality is ubiquitous and abundant in our earthbound experience.<sup>676</sup> Mental activity, in comparison, is scarce and exotic. In our experience it is only living entities exhibit mental characteristics. Further only human beings exhibit high-order mental states. The sheer commonness of physicality makes it appear a very unpromising place to start looking for a definition of personal identity. The uniqueness of sophisticated mental states

---

<sup>674</sup> For a discussion of this approach see Smith (1989)

<sup>675</sup> For arguments I take to be definitive on this issue see Lowe (1991b)

<sup>676</sup> Of course, in the greater part of the universe matter is scarce. Planets are uncommon local concentrations.

to human animals, our paradigm persons, makes it seem the ideal grounding for such a project.<sup>677</sup> What is more, if one encountered something that was not human but still exhibited high order mental states then one would be inclined to count such a thing as a person.<sup>678</sup> Whereas there are common examples of objects that appear to be human animals but which we may well not want to call persons, e.g. the severely brain-damaged and corpses.

Clearly, though, just choosing *any* characteristic unique to human persons would not provide a suitable criterion for personal identity. As we have seen both Descartes and Locke brusquely, and correctly, derided the philosophical value of human persons being the only rational animals and the only featherless bipeds. Something more than being a unique feature of human persons recommends high-order mentality as the qualifying criterion of personhood. High-order mental states are thoughts about such things as ethics, past events, meanings, chess, art, science and philosophy. We do not just value the fact that human persons have high-order mental states because they are unique to us. We value them because they are integral to the very *experience* of human existence.

Yet it is not clear that what we *value* about ourselves is any better guide to what constitutes personal identity than what is unique. Human persons all need hearts to survive, but many people are conspicuously cavalier about the condition of their hearts (especially those devoted to more cerebral lifestyles). Surely whims about what we value about ourselves should not decide what constitutes personhood. What we value seems determined by our experience of the human condition not what is essential to being human. If we had a more intimate connection to our biology (say we could introspect states of physical health in a manner which normally requires an x-ray) we might well value that aspect of our existence more highly.<sup>679</sup>

Consequently the personal identity debate has a profound dissonance at its heart. We experience ourselves as corporeal beings, yet our embodiment seems philosophically mundane. Instead philosophical efforts have been concentrated on our psychological

---

<sup>677</sup> A relatively recent and much celebrated example of just this is Frankfurt (1971)

<sup>678</sup> At least this intuition is appealed to by supporters of the psychological criterion. It is not obvious that everyone shares this intuition. Even if everyone did, treating a rational parrot as a suitable partner for conversation and an object of respect would not mean that one is treating the parrot as a person.

<sup>679</sup> The impulse to insist that conditions of personhood are not malleable is, nonetheless, peculiarly philosophical. Outside of philosophy, in disciplines such as anthropology and sociology the notion of person is taken to be some sort of social or logical construct. On this view what is valued can be the basis for personhood. One cannot deny that different cultural perceptions of personal identity exist, e.g. those cultures that believe in a personal afterlife and those that do not. Clearly to believe one could survive bodily death will

capacities. What is more a legion of examples suggest that our mental and physical beings are distinct entities. One can easily be shown to continue without other. In recent debate this has come to be expressed as the distinction between a person and the human animal each coincides with. This is the New Dualism.<sup>680</sup> The relationship one has to one's material part is like the relationship between a statue and the copper that makes it up. The statue can come to be composed of different pieces of copper, or even a different material altogether. Likewise a person might come to have a different body. Indeed persons seem even more robust than statues. It has been suggested that persons can survive as information records in computers, as disembodied spiritual entities, or as non-biologically embodied beings. If one is not persuaded by these fanciful suggestions of persons' longevity then one need only consider cases where a person seems to not survive for as long as their coincident body. The senile, those in vegetative states and the dead all seem to have bodies. Many find these cases persuasive in demonstrating that the person and the human animal are distinct.

What makes this dissonance so frustrating is that there seems to be no way of resolving it that does not prejudge our answer to the question at hand. Once the distinction has been accepted between person and body writers are forced to choose between the two answers. Thought-experiments and 'intuition-pumps' vie for our attention. However, the primitive status of our pre-philosophical notions of both person and human animal forces presumptions onto the arguments. The very way we frame and describe the problem cases presumes the theory one is trying to prove.<sup>681</sup> The only way to overcome this impasse is to dissolve the dualism upon which it is predicated.

---

significantly alter how one feels. The philosopher, however, maintains that these different views of personhood must, in fact, be different views of *common and shared* phenomena.

<sup>680</sup> The term is adopted from Burke (1997b). Similar arguments are found in van Inwagen (1997) and Olson (1997a).

<sup>681</sup> This is one reason why one might argue against the worth of thought experiments. See Wilkes (1988) for length discussion of this issue. For cautionary comments about thought experiments I find myself in great sympathy with see van Inwagen (1997).

## II

### DISSOLVING THE NEW DUALISM

The central debate for personal identity, therefore, turns on the legitimacy of the New Dualism. Harold Noonan in a recent provocative paper has sketched five possible solutions.<sup>682</sup>

1. Reject the existence of human beings
2. Reject any possibility of a divergence between personal identity and animal identity
3. Reject the claim that animals are ever thinking intelligent beings
4. Accept that the concept of a person is not a sortal concept at all
5. Accept that the utterer of 'I' need not be identical with the referent of 'I'

Noonan rightly rejects (1). (3) is associated with Shoemaker and there are good reasons for rejecting this as well. One is obliged to adopt this curious position if one wishes to insist that being a thinking intelligent being is sufficient for being a person. If a person thinks with the brain of a spatially coincident human animal then surely the animal can also be said to think. In this case it too is surely a person. This is known as the paradox of the two thinkers.<sup>683</sup> This paradox does have to be avoided, but (3) is both drastic and absurd. It is therefore not to be preferred as a solution.

5 embraces the New Dualism. Noonan bills this as the Neo-Lockean solution and it is the answer he prefers to Animalism (as Noonan construes it). Noonan's version of Animalism is expressed in (2). Although Noonan's definition captures the central drive of Animalism there are many ways of cashing out the details of this assertion. These will be explored below. Noonan rejects 4 because it "involves the radical claim that the topic of personal identity is strictly speaking non-existent".<sup>684</sup> In fact, 4 is worth considering despite its radical appearance. Personal identity is treated, following Locke and Hume, as a special problem. This is not without good reason as the problem cases are uniquely complex and

---

<sup>682</sup> Noonan (1998) p. 317

<sup>683</sup> See van Inwagen (1990), Garrett (1998), Burke (1997b)

intriguing. Further, given that philosophers are all persons there is a peculiarly intimate fascination with settling these issues. However, if the New Dualism is now the battleground for personal identity denying that one half of the dualism exists is a legitimate way of dissolving it. To this end the following sections will explore the links between different brands of Animalism, the New Dualism, and the possibility of eliminating the philosophical problem of personal identity.

---

<sup>694</sup> Noonan (1998) p. 318

### III

#### TWO TYPES OF ANIMALISM

Noonan asserts that Animalism must reject any possibility of a divergence between personal identity and animal identity. However, it is possible to develop versions of Animalism which are both weaker and stronger than the version Noonan provides.

First, though, it is worth considering why an embodiment thesis is not a version of Animalism. Although Neo-Lockeans believe the person is distinct from the human animal the prevalence of materialism in contemporary philosophy means that Neo-Lockeans invariably also believe a person must be instantiated by some material object. Disembodied existence is typically regarded as implying the existence of 'Cartesian Egos' (notwithstanding that such entities are a metaphysical caricature as has been discussed in Chapter 1). It is consequently discredited and rejected out of hand.<sup>685</sup>

Neo-Lockeans diverge on how strict these conditions of embodiment must be. Our typical example of embodiment will of course be a human animal. Neo-Lockeans can appeal to few empirical examples of non-human personal embodiment.<sup>686</sup> Instead the literature proliferates with thought experiments which, of course, presume the possibility of non-human embodiment they are meant to prove. It may be easy to imagine metamorphosing into a giant insect, as Kafka did and the motion picture *The Fly* so vividly illustrates. However, despite the thoroughness of these artistic thought experiments, and the clarifications and ramifications philosophers feel they have added, such unnatural metamorphoses do not occur. What is more it seems clear the act of the imagination is only made possible by the premise that the person is not identical with its original human body.

However, a Neo-Lockean who was unconvinced by these and more lavish arguments about Brain Zap Machines<sup>687</sup> and Psychological Blueprints<sup>688</sup> could insist that persons must be embodied by human animals. This is clearly in sympathy with the assertion of Animalists

---

<sup>685</sup> Strawson (1959) is famously permissive about the possibility of disembodied existence.

<sup>686</sup> Locke's rational parrot is meant to be one such example. However see Chapter 2, §IV above for a contrary interpretation.

<sup>687</sup> Shoemaker & Swinburne (1984) p. 87

<sup>688</sup> This is a presumption of the 'teletransporter machine' as exploited in Parfit (1984). The source of this thought experiment is the Science-Fiction world of *Star Trek*. Parfit claims we must distinguish between thought experiments which are deeply impossible and those which are only technically impossible. For comprehensive and conclusive arguments that the teletransporter is deeply impossible see Krauss (1996). I join van Inwagen in his sentiments that so-called 'logical possibility' is frequently spurious and unrevealing. See van Inwagen (1997) p. 308

like Paul Snowdon, who comments that “Animalism does not simply claim that we are animals; it also says that we are animals of a certain *human* kind.”<sup>689</sup> This restricted version of the embodiment thesis might appear to be a form of Animalism. It is not, though, because the identity is not primarily animal identity. The Neo-Lockean can still permit, on this account, that the person might only emerge after the animal begins to exist, and may dissolve before the animal does. Nor can the embodiment thesis be turned into Animalism by asserting that whenever a human animal is present a person is present, and *vice versa*. Such a constant coincidence would merely be a sort of pre-established harmony. Personal identity would still be *constituted* by something other than the human animal, even though the human animal might always represent reliable *evidence* for the presence of a person. Embodiment cannot close the gap between the Neo-Lockean and the Animalist because embodiment is a weaker condition than constitution.<sup>690</sup>

The weakest possible form of Animalism would regard it as a contingent fact about our world that only human animals are persons. This position itself has degrees. This contingency may have a merely temporal limit. In due course other species may emerge as persons as well. The contingency might be merely spatial. It might be a contingent fact of the evolution of Earth’s flora and fauna that only humans are persons. On some distant planet a very different type of animal, or even plant, might exhibit personhood.<sup>691</sup> In a stronger version it is only in another possible world that non-human persons exist. All these positions are of little interest because they collapse into the New Dualism.<sup>692</sup> Any concession that persons and human animals might be distinct is enough to justify the New Dualism.

Noonan is therefore correct in insisting the Animalist must reject these weaker versions and reject all possible divergences between persons and human animals. Animalists have developed a battery of arguments to this end. Olson best captures the thrust the Animalist must take in the paper-title ‘Is Psychology Relevant to Personal Identity?’<sup>693</sup> When we talk of what we value most in life we often talk of our ‘hopes and dreams’. Hopes and dreams

---

<sup>689</sup> Snowdon (1998) p. 73. As will become apparent Snowdon’s insistence may be unwarranted.

<sup>690</sup> Wiggins, as I understand him, does think embodiment is like constitution, which is why he is often taken to be a kind of Animalist. For my reasons for denying Wiggins is an Animalist see the Introduction, especially note 11. For discussion more sympathetic to Wiggins see Snowdon (1996).

<sup>691</sup> This claim is not what David Lewis dubs the ‘telescope’ (mis-construal) of possible worlds in Lewis (1986). This planet might be something literally viewable by a telescope in the way Mars and Venus are. Its discovery would be an *empirical* matter.

<sup>692</sup> Their only appeal is as a heuristic device to establish that the absence of real examples non-human persons means that being a human is, within well stated constraints, a guarantee of being a person. This is why Weak Animalism needs some consideration. However, it remains of little interest to the more metaphysically motivated philosopher.

are psychological phenomena. The Neo-Lockean bases the appeal of their thought-experiments and intuition pumps by appealing to this (quite natural) concern for our psychology. The Animalist must work to undercut the strength of these intuitions.

The Animalist agrees that life without a fully-functioning mind would be unbearable. The Animalist agrees that prenatal and neonatal children have very few psychological characteristics. Animalists fear senility and decay of the mind as much as anyone (as they are philosophers and professional thinkers maybe more). Yet none of this obliges the Animalist to accept the Neo-Lockean's arguments.

First, the Animalist points out that although each of us was once a foetus we became adults. A Psychological Criterion of personal identity must deny that a foetus is person because it lacks sufficient psychological activity. The adult is a person, and so cannot be identical with the foetus. The Psychological Criterion has the bizarre outcome that starting to think is not a process of development for a neonate, but its cessation. What is more, every adult alive was never in fact a baby. Each of us was never carried in our mother's womb. The Animalist asserts that this consequence is absurd. If the Neo-Lockean feels obliged to pay this philosophical price for some other gain elsewhere, the Animalist insists the price is far too high.<sup>694</sup>

The Animalist can also point out that only on their presupposition can we make sense of our fear of senility. One should fear becoming cerebrally enfeebled only if one will become the afflicted future person. On the Animalist's account whatever psychological deterioration afflicts one's body one continues to exist. This explains our fear. On the Psychological Criterion the future cerebrally enfeebled human is not a person. We do not have to fear becoming such a person because we will always cease to exist before this occurs. Of course, one can still fear the point where one becomes insufficiently psychologically endowed to exist as a person. This would be one's death. Our intuitions are torn by this example because the Psychological Criterion offers solace, yet only Animalism explains a common fear.

However, the Animalist can appeal to a version of the Two Thinkers' Paradox to demonstrate the absurdity of the Neo-Lockean position. If one can die, then one is a living thing. Surely, though, the remaining human animal is also a living thing. Each of us in life is, according to the Neo-Lockean, therefore not one living thing, but two. The Neo-

---

<sup>693</sup> Olson (1994). The arguments of this paper are revisited in the chapter of the same name in Olson (1997a).

<sup>694</sup> This argument is found in Olson (1997b) and revisited in the chapter of the same name in Olson (1997a).

Lockean may claim that cessation of psychological capacity is not death. Animals die, persons cease to exist. However, this implies that no-one dies. This hardly seems satisfactory, as it also implies no-one lives. The Neo-Lockean may want to embrace the Two Lives' Paradox, but it is to be avoided for the same reasons as the Two Thinkers' Paradox. It should be rejected as a strategy for reasons parallel to those that led above to the rejection of (3).

Finally, the Animalist can insist that only animals provide the sort of robust metaphysical grounding that persons need. A collection of examples have grown up to supposedly demonstrate that more than one thinker may occupy the same body. Split Brain<sup>695</sup> cases and cases of Multiple Personality<sup>696</sup> supposedly demonstrate that mental unity does not coincide with physical unity. For an entity to be identical it must be possible to identify it both *diachronically* and *synchronically*. Intimately related to this are conditions of unity. A unity condition should determine whether a possible part of something is or is not actually a part of it. For example, an ice cube placed in warm water will start to melt. A unity condition for the cube will provide some principle to determine whether a particular water molecule is or is not still part of the cube.<sup>697</sup> Likewise a Neo-Lockean account of personal identity should be able to provide unity conditions for being a person. The difference between Split-Brain and Multiple Personality cases is only one of putative underlying causes, one physical the other psychological. For this reason only Multiple Personality will be investigated. The same considerations provide the Animalist's answer to Split-Brain cases.

It is clear that Multiple Personality is defined as a problem only because it offends the standard situation where one human being exhibits one personality. The Psychological Criterion, in turn, uses this presumption of a single personality to justify the New Dualism, equating personality to person. According to the Psychological Criterion a person is a unified maximal psychological entity. Hence if a personality is a person those who hold the Psychological Criterion assume that personalities are unified maximal psychological entities. While being a maximal unified psychological entity is a good definition for a metaphysical category like 'person' it is not clear it is a good definition for a non-metaphysical, folk-psychological term like 'personality'. Separating these two will both

---

<sup>695</sup> These were first philosophically exploited in the context of personal identity in Nagel (1971).

<sup>696</sup> Two full length studies of these cases are Hacking (1995) and Braude (1995). See also discussion in Wilkes (1988).

<sup>697</sup> Alternatively if the conditions are indeterminate the bounds of this indeterminacy should be expressed.

explain why cases of Multiple Personality are so described, and show that Multiple Personality presumes the identity of a human animal.

In cases of multiple personality personalities are not defined as incommunicable consciousnesses as in the way Locke did in his related but quite distinct thought experiment. A personality is more akin to a loosely coherent set of beliefs and desires. It is obvious that many people, not just Multiples, have subsets of beliefs and desires that are internally coherent, but mutually antagonistic. One's set of beliefs about being a parent might conflict with one's set of beliefs about a career, for example. In cases of Multiple Personality this antagonism is starker and personalities form that have almost no common elements. These personalities are not incommunicable, indeed personalities are frequently aware of each others beliefs and despise or approve of each other accordingly.

In the normal case a person exhibits a set of beliefs which can be summed into a maximal psychological entity. Although the Animalist asserts that this definition is wrong, for the moment it will be granted that this definition could easily tempt. Also in the normal run of things each person will only exhibit one broad personality. The conflict in beliefs each of us has will not prompt our personality to fragment. This state of affairs is upset in the case of the multiple. The beliefs are fragmented enough to justify delimiting different personalities. This fragmentation of personality seems to provide problems for the maximal psychological entity that the person is supposed to be. The Psychological Criterion tempts one to say that there is not only more than one personality, but also more than one person.

The problem emerges because the Psychological Criterion cannot separate personality and person properly. The Animalist can do this. The Animalist can accept that personalities are coherent subsets of beliefs, and that each person has a maximal set of psychological events and these all belong to one animal. The Animalist accepts the intuition that all of our beliefs belong to us, without making this constitute personal identity. The Psychological Criterion does not do this. Psychological continuity and connectedness are not merely a typical manifestation of personal identity, but have to constitute personal identity as well. This insistence causes the confusion between personality and person. The Psychological Criterion ends up committed to saying that each apparent set of beliefs (which should merely constitute a personality) is actually a person. The Animalist can appeal to a non-psychological connection to ground our mental lives. This allows the Animalist to properly distinguish between the person and personalities. The Animalist can draw support from the fact that this supports our intuition that each of us can exhibit different personalities without threat to personal identity, and that Multiples are not another species of being, but

fragmented and mentally distressed persons. None of this is available to the Psychological Criterion. It will also be clear from this that cases of Multiple Personality clearly presume and rely on an Animalist account to underwrite their presumption that unified personalities are the norm. This is clear both from the practice of treating the human being as the patient, not each individual personality as an individual patient. It is also philosophically clear that the edifice of personalities presumes a metaphysically sound account of personal identity that only Animalism can provide.

## IV

### DEFLATIONARY ANIMALISM & THE COMPOSITIONAL ACCOUNT

It is a presumption of both sides of the debate that Animalism asserts an identity between persons and *human* animals. This preserves the belief that there is something peculiar about humans that justifies their claim to be persons. This anthropocentrism is inevitable given both the way the problem is usually framed and conceived. Yet, it is also philosophically undesirable. Neo-Lockeans do not find themselves burdened with such chauvinism and regard this as a strength of their position. Therefore, if Animalism can ameliorate this objection it should do so.

The Animalist states that personal identity is nothing more than the identity of a human animal. Humans are just another mammalian animal and their identity conditions are surely the same as for other mammals. This implies that there is no difference between the conditions for squirrel identity and for human identity. Transitivity then demands the Animalist accept there is no difference between squirrel identity and personal identity. As squirrels are *not* persons this conclusion would seem *prima facie* absurd. However, this absurdity can be avoided by embracing the radical agenda Noonan rejected as the consequence of 4. That is, we should deny the existence of persons, and in doing so, eliminate the problem of personal identity.

This elimination does not imply that all talk of persons must also be eliminated. It does, however, imply the elimination of persons as a metaphysical category. Instead, strictly speaking, we should only talk of animal identity. Talk of persons is only shorthand used to isolate certain kinds of behaviour typically exhibited by certain kinds of animal. This could be reduced to talk of certain kinds of behaviour. This clearly avoids anthropocentrism because personal identity is no longer explicitly linked to humans.

However, the Animalist must move carefully or accidentally surrender to the Neo-Lockean. The Neo-Lockean will suggest that all the Animalist has produced is a list of typical kinds of behaviour with the notion of an animal at the core. Unless the Animalist can provide reasons why these lists of typical behaviours must only belong to animals the New Dualism can simply reassert itself.

The Animalist can block this argument. While accepting that these lists can be constructed, the Animalist can insist that no list of behavioural qualities can ever constitute a metaphysical category. Persons cannot be rebuilt. The only metaphysical category that can

sustain the weight of personal identity is the category of animal. So-called 'personal' behaviour is ontologically dependent on the existence of animals. Hence the Neo-Lockean is unable to establish the distinction they require between person and animal.

It is in these uncharted areas of the personal identity debate that the Compositional Account finds contemporary application. Animalism expressed in its biological form is metaphysically robust but fails to capture common intuitions. The Compositional Account offers what one might call a folk metaphysical position. Persons are the kind of entities that are composed of minds and bodies. Folk metaphysics' inherent parsimony is antagonistic to theories of coincidence (and hence paradoxes of doubling like the Two Thinkers' Paradox). Therefore the best metaphysical category to identify with these composites is that of animal. The notion of a mind and a living body are clearly as fundamental as our notion of a person. This agrees with the Compositional Account's folk metaphysical sympathies. It provides a deflation of personal identity that is primitive yet open to analysis. Further the Compositional Account can avoid the charge of being anthropocentric or even privileging animals. *Any* entity that met the conditions of being a self-sufficient, living body-in union with a complex enough psychological structure to warrant the title of a mind could be regarded as person-like. This seems correct. These would be the conditions we would look for in alien beings if we were to treat them the way we treat other members of our own species.

Neo-Lockeans like Noonan are wrong to balk at the radical implications of abandoning person as an abiding sortal term. The future challenge in the personal identity debate is to explore demystify these previous radical positions. The Compositional Account can play a very interesting role in exploring these. What is more, the traditional philosophical heroes (and villains) of the personal identity debate, Descartes, Locke and Hume, have already provided a rich framework of ideas and arguments to adopt, adapt and rediscover.

## BIBLIOGRAPHY

- Adams, R. (1979). 'Primitive Thisness and Primitive Identity.' *Journal of Philosophy* Vol. 76
- Alexander, Peter. (1980). 'Locke on Substance-in-General – Part I' *Ratio* Vol. 22
- Alexander, Peter. (1981). 'Locke on Substance-in-General – Part II' *Ratio* Vol. 23
- Allison, Henry E. (1966). 'Locke's Theory of Personal Identity: A Re-examination.' *Journal of the History of Ideas* Vol. 27. Reprinted in Tipton (1977) with revisions.
- Allison, Henry E. and Jolley, Nicholas (1981). 'Locke's Pyrrhic Victory.' *Journal of the History of Ideas* Vol. 42
- Alston, William P. & Bennett, Jonathan. (1984). 'Identity and Cardinality: Geach and Frege.' *Philosophical Review* Vol. 93
- Alston, William P. & Bennett, Jonathan. (1988). 'Locke on People and Substances.' *Philosophical Review* Vol. 97
- Ameriks, Karl. (1976). "Personal Identity and Memory Transfer." *Southern Journal of Philosophy* Vol. 14
- Anscombe, G.E.M. (1953). 'The Principle of Individuation.' *Proceedings of the Aristotelian Society* Vol. 27
- Anscombe, G.E.M. (1975). 'The First Person' Reprinted in Yourgrau (1990)
- Arcaya, Jose M. (1989). 'Memory and Temporality: A Phenomenological Alternative.' *Philosophical Psychology* Vol. 2
- Ashley, Lawrence & Stack, Michael. (1974). 'Hume's Theory of the Self and Its Identity.' *Dialogue* Vol. 13
- Atherton, Margeret. (1983). 'Locke's Theory of Personal Identity.' *Midwest Studies in Philosophy* Vol. 8
- Ayer, A.J. (1956). *The Problem of Knowledge*. Harmondsworth: Penguin Books.
- Ayer, A.J. (1964). *The Concept of a Person and Other Essays*. London: Macmillan.
- Ayer, A.J. (1975). 'Identity and Reference.' *Philosophia* Vol. 5
- Ayers, Michael A. (1981). *Locke*. London: Routledge
- Ayers, Michael R. (1974). 'Individuals Without Sortals.' *Canadian Journal of Philosophy* Vol. 4
- Ayers, Michael R. (1975). 'The Ideas of Power and Substance in Locke's Philosophy.' *Philosophical Quarterly* Vol. 25. Reprinted in Tipton (1977) with revisions.
- Ayers, Michael R. (1994). 'The Foundations of Knowledge and the Logic of Substance: The Structure of Locke's General Philosophy.' In Rogers (1994).
- Baier, Annette. (1991). *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, Massachusetts: Harvard University Press.
- Baier, Annette. (1979). 'Hume on Heaps and Bundles.' *American Philosophical Quarterly* Vol. 16
- Baillie, J. (1993). 'Recent Work on Personal Identity.' *Philosophical Books* Vol. 34
- Baker, Gordon P. & Morris, Katherine J. (1997). *Descartes' Dualism*. London: Routledge.
- Basson, A.H. (1958). *David Hume*. London: Pelican Books
- Baxter, Donald M. (1987). 'A Defence of Hume on Identity Through Time.' *Hume Studies* Vol. 13
- Baxter, Donald M. (1988). 'Identity in the Loose & Popular Sense.' *Mind* Vol. 97
- Beauchamp, Tom L. (1979). 'Self-Inconsistency or Mere Self-Perplexity?' *Hume Studies* Vol. 5
- Beck, Robert N. (1969). 'Some Remarks on Logic and the *Cogito*.' in Magnus & Wilbur (1969)
- Behan, D.P. (1985). 'Hume's Labyrinth.' *History of Philosophical Quarterly* Vol. 2
- Bergsma, Daniel, ed. (1967). *Conjoined Twins: Birth Defects Original Article Series Vol. III*. New York: The National Foundation.
- Bermudez, Jose Luis, ed. (1998). *The Body And The Self*. Cambridge, MA: MIT Press
- Bradley, M.C. (1969). Critical Notice of Wiggins' *Identity and Spatio-Temporal Continuity*. *Australian Journal of Philosophy* Vol. 47
- Braude, Stephen E. (1995). *First Person Plural: Multiple Personality and the Philosophy of Mind* (Revised Edition). Lanham, Maryland: Rowman & Littlefield.
- Brennan, Andrew. (1969). 'Persons and Their Brains.' *Analysis* Vol. 30
- Brennan, Andrew. (1982). 'Personal Identity and Personal Survival.' *Analysis* Vol. 41
- Brennan, Andrew. (1984). 'Survival.' *Synthesis* Vol. 59
- Brennan, Andrew. (1984). Critical study of Eli Hirsch *The Concept of Identity*. *Nous* Vol. 18
- Brennan, Andrew. (1985). 'Amnesia and Psychological Continuity.' in Copp & McIntosh (1985)
- Brennan, Andrew. (1988). *Conditions of Identity: A Study in Identity and Survival*. Oxford: Clarendon Press.
- Bricke, John (1977). 'Hume on Self-Identity, Memory and Causality.' in Morice (1977).
- Brody, Baruch A. (1980). *Identity and Essence*. Princeton: Princeton University Press.
- Bryan, Elizabeth. (1992). *Twins and Higher Multiple Births: A Guide to their Nature and Nurture*. London: Edward Arnold.
- Burge, Tyler. (1977). 'A Theory of Aggregates.' *Nous* Vol. 11
- Burge, Tyler. (1988). 'Individuation and Self-Knowledge.' *Journal of Philosophy* Vol. 85

- Burke, Michael B. (1980). 'Cohabitation, Stuff, and Intermittent Existence.' *Mind* Vol. 89
- Burke, Michael B. (1985). 'Spatial Analogues of "Annihilation and Re-Creation."' *Analysis* Vol. 45
- Burke, Michael B. (1992). 'Copper Statues and Pieces of Copper: A Challenge to the Standard Account.' *Analysis* Vol. 52
- Burke, Michael B. (1994a). 'Preserving the Principle of One Object to a Place: A Novel Account of the Relations Among Objects, Sorts, Sortals, and Persistence Conditions.' *Philosophy and Phenomenological Research* Vol. 54
- Burke, Michael B. (1994b). 'Dion and Theon: An Essentialist Solution to an Ancient Puzzle.' *Journal of Philosophy* Vol. 91
- Burke, Michael B. (1997a). 'Coinciding Objects: Reply to Lowe and Denkel.' *Analysis* Vol. 57
- Burke, Michael B. (1997b). 'Persons and Bodies: How To Avoid the New Dualism.' *American Philosophical Quarterly* Vol. 34
- Bursen, H.A. (1978). *Dismantling the Memory Machine* (London: D.Reidel Publishing Company)
- Butler, Joseph. (1736). 'Of Personal Identity.' in Perry (1975a)
- Butler, R.J., ed. (1965). *Analytical Philosophy: Second Series*. Oxford: Basil Blackwell.
- Campbell, John. (1992). 'The First Person: The Reductionist View of the Self.' In Charles and Lennon (1992).
- Capaldi, N. (1995). 'The Historical and Philosophical Significance of Hume's Theory of the Self.' Reprinted in Tweyman (1995)
- Capaldi, Nicholas. (1975). *David Hume: The Newtonian Philosopher*. Boston: Twayne.
- Capaldi, Nicholas. (1991). 'The Hume Literature of the 1980's.' *American Philosophical Quarterly* Vol. 28
- Capaldi, Nicholas. (1992). 'The Dogmatic Slumber of Hume Scholarship.' *Hume Studies* Vol. 18
- Care, Norman S., & Grimm, Robert H. eds. (1969). *Perception and Personal Identity: Proceedings of the 1967 Oberlin Colloquium in Philosophy*. Cleveland: The Press of Case Western Reserve University.
- Carruthers, Peter (1986). *Introducing Persons: Theories and Arguments in the Philosophy of Mind*. London: Routledge.
- Carter, W.R. (1982). 'On Contingent Identity and Temporal Worms.' *Philosophical Studies* Vol. 41.
- Carter, William. (1983). 'In Defence of undetached Parts.' *Pacific Philosophical Quarterly* Vol. 64
- Cassam, Quassim. (1992). 'Reduction and First-Person Thinking.' In Charles and Lennon (1992).
- Cassam, Quassim. (1993). 'Parfit on Persons.' *Proceedings of the Aristotelian Society* Vol. 93
- Chappell, V.C. (1970). 'Stuff and Things.' *Proceedings of the Aristotelian Society* Vol. 71
- Chappell, V.C. (1973). 'Matter.' *Journal of Philosophy* Vol. 70
- Chappell, Vere. (1989). 'Locke and Relative Identity.' *History of Philosophy Quarterly* Vol. 6
- Chappell, Vere. (1990). 'Locke on the Ontology of Matter, Living Things and Persons' *Philosophical Studies* Vol. 60
- Charles & Lennon, eds. (1992). *Reduction, Explanation and Realism*. Oxford: Oxford University Press.
- Chisholm, Roderick M. (1969). 'Reply to Shoemaker's Comments.' in Care & Grimm (1969).
- Chisholm, Roderick M. (1969). 'The Loose and Popular and the Strict and Philosophical Senses of Identity.' in Care & Grimm (1969).
- Chisholm, Roderick M. (1976). *Person and Object: A Metaphysical Study*. London: George Allen & Unwin.
- Church, R.W. (1935). *Hume's Theory of The Understanding*. London: George, Allen & Unwin.
- Cockburn, David, ed. (1991). *Human Beings*. Cambridge: Cambridge University Press.
- Collins, James. (1971). *Descartes' Philosophy of Nature*. Oxford: Basil Blackwell.
- Coons, P. M. (1986). 'The Prevalence of Multiple Personality Disorder.' *International Society for the Study of Multiple Personality and Dissociation Newsletter* Vol. 4.
- Copp, R. & McIntosh, J.J., eds. (1985). *New Essays in the Philosophy of Mind: Canadian Journal of Philosophy Supplement* Vol. 11
- Cottingham, John, ed. (1992). *The Cambridge Companion to Descartes*. Cambridge: Cambridge University Press.
- Cottingham, John, ed. (1994a). *Reason, Will, and Sensation: Studies in Descartes's Metaphysics*. Oxford: Clarendon Press.
- Cottingham, John. (1985). 'Cartesian Trialism.' *Mind* Vol. 94
- Cottingham, John. (1986). *Descartes*. Oxford: Basil Blackwell.
- Cottingham, John. (1994b). 'Plus una vice agendum: Cartesian Metaphysics Three and a Half Centuries On' in Cottingham (1994a)
- Cresswell, M.J. (1992). 'The Ontological Status of Matter in Aristotle.' *Theoria* Vol. 58
- De Sousa, Ronald. (1976). 'Rational Homunculi.' in Rorty (1976a)
- Denkel, A. (1995). 'Theon's Tale: Does a Cambridge Change Result in a Substantial Change?' *Analysis* Vol. 55
- Dennett, Daniel. (1976). 'Conditions of Personhood.' in Rorty (1976a)
- Descartes, Rene. (1911). *The Philosophical Works of Descartes Volumes I & II* (ed. Haldane & Ross).

- Cambridge: Cambridge University Press.
- Descartes, Rene. (1984). *The Philosophical Writings of Descartes Volume II* (ed. Cottingham, Stoothoff & Murdoch). Cambridge: Cambridge University Press.
- Descartes, Rene. (1985). *The Philosophical Writings of Descartes Volume I* (ed. Cottingham, Stoothoff & Murdoch). Cambridge: Cambridge University Press.
- Descartes, Rene. (1991). *The Philosophical Writings of Descartes Volume III* (ed. Cottingham, Stoothoff, Murdoch & Kenny). Cambridge: Cambridge University Press.
- Doepke, Frederick. (1982). 'Spatially Coinciding Objects.' *Ratio* Vol. 24
- Doepke, Frederick. (1986). 'In Defence of Locke's Principle.' *Mind* Vol. 95
- Doepke, Frederick. (1987). 'The Structure of Persons and Artefacts.' *Ratio* Vol. 29
- Doney, Willis, ed. (1970). *Descartes: A Collection of Critical Essays*. London: Macmillan.
- Dummett, Michael. (1981). *Frege: Philosophy of Language*. London: Duckworth.
- Edwards, Paul, ed. (1967). *The Encyclopedia of Philosophy*. New York: Macmillan
- Eilan, Naomi. (1994). 'The First Person Perspective.' *Proceedings of the Aristotelian Society* Vol. 95.
- Evans Gareth. (1985). *Collected Papers*. Oxford: Clarendon Press.
- Evans, Gareth. (1982). *The Varieties of Reference*. Oxford: Clarendon Press.
- Fang, Wan-Chuan. (1984). 'Hume on Identity.' *Hume Studies* Vol. 10
- Fine, Alan. (1986). 'Transplantation in the Central Nervous System.' Originally in *Scientific American*, collected in Llinás (1988).
- Flage, Daniel E. (1990). *David Hume's Theory of Mind*. London: Routledge.
- Flage, Daniel. (1982). 'Hume's Dualism.' *Nous* Vol. 16
- Flew, Anthony. (1951). 'Locke on the Problem of Personal Identity.' *Philosophy* Vol. 26.
- Flew, Anthony. (1977). 'Personal Identity: Offences of the Hughes Defence.' *Philosophy* Vol. 52.
- Flew, Anthony. (1986). *David Hume: Philosopher of Moral Science*. Oxford: Basil Blackwell.
- Flew, Antony. (1979). 'Locke and Personal Identity - Again.' *Locke Newsletter* No.10
- Fogelin, Robert J. (1985). *Hume's Skepticism in the Treatise of Human Nature*. London: Routledge.
- Foster, John. (1979). 'In Self-Defence.' in Macdonald (1979)
- Frankfurt, Harry. (1971). 'Freedom of the Will and the Concept of a Person.' *Journal of Philosophy* Vol. 68
- Frankfurt, Harry. (1976). 'Identification and Externality.' in Rorty (1976a)
- Frasca-Spada, Marina. (1998). *Space and The Self in Hume's Treatise*. Cambridge: Cambridge University Press.
- Frege, G. (1953). *The Foundations of Arithmetic (1884)* (Trans. J.L. Austin). Oxford: Basil Blackwell.
- Garber, Daniel & Ayers, Micheal, eds. (1998). *The Cambridge History of Seventeenth Century Philosophy*. Cambridge: Cambridge University Press
- Garrett, Brian. (1991). 'Personal Identity and Reductionism.' *Philosophy and Phenomenological Research* Vol. 51.
- Garrett, Brian. (1998). *Personal Identity & Self-Consciousness*. London: Routledge.
- Garrett, Don. (1997). *Cognition and Commitment in Hume's Philosophy*. Oxford: Oxford University Press.
- Geach, P.T. (1968). 'Identity.' *Review of Metaphysics* Vol. 21
- Geach, Peter T. (1962). *Reference and Generality: An Examination of Some Medieval and Modern Theories*. Ithaca: Cornell University Press.
- Gedda, Luici. (1961). *Twins in History & Science*. Springfield, Illinois: Charles C. Thomas.
- Gill, Christopher, ed. (1989). *The Person and the Human Mind in Ancient Philosophy*. Oxford: Oxford University Press
- Glover, Jonathan. (1988). *I: The Philosophy and Psychology of Personal Identity*. Harmondsworth: Penguin Books.
- Gomulicki, Bronislaw. (1953). *The Development and Present Status of the Trace Theory of Memory*. Cambridge: Cambridge University Press
- Green, Michael J. (1999). 'The Idea of a Momentary Self and Hume's Theory of Personal Identity.' *British Journal for the History of Philosophy* Vol. 7
- Greig, J.Y.T. (1931). *David Hume*. London: Jonathan Cape.
- Grice, H.P. (1941). 'Personal Identity.' *Mind* Vol. 50. Reprinted in Perry (1975a)
- Grosholz, Emily. (1991). *Cartesian Method and the Problem of Reduction*. Oxford: Oxford University Press
- Hacking, Ian. (1995). *Rewriting The Soul: Multiple Personality and the Sciences of Memory*. Princeton, New Jersey: Princeton University Press.
- Hall, Roland. (1974). 'Hume's Use of Locke on Identity.' *Locke Newsletter* Vol. 5
- Harris, Henry, ed. (1995). *Identity: Essays Based on Herbert Spencer Lectures Given in the University of Oxford*. Oxford: Clarendon Press.
- Harris, Ian. (1994). *The Mind of John Locke: A Study of Political Theory in its Intellectual Setting*. Cambridge: Cambridge University Press
- Heil, John. (1978). 'Traces of Things Past.' *Philosophy of Science* Vol. 45
- Helm, Paul. (1981). 'Did Locke Capitulate to Molyneux?' *Journal of the History of Ideas* Vol. 42

- Helm, Paul. (1977). 'John Locke's Puzzle Cases About Personal Identity.' *Locke Newsletter* No. 8
- Helm, Paul. (1979). 'Locke's Conditions for Personal Identity.' *Locke Newsletter* No. 10
- Hintikka, Jarkko. (1962). 'Cogito, Ergo Sum: Inference or Performance?' *The Philosophical Review*, Vol. 71. Reprinted in Doney (1970).
- Hirsch, Eli. (1982). *The Concept of Identity*. Oxford: Oxford University Press.
- Hirsch, Eli. (1983). 'Hume's Distinction Between Genuine and Fictitious Identity.' *Midwest Studies in Philosophy* Vol. 8
- Hoffman, Joshua & Rosenkrantz, Gary S. (1997). *Substance: It's Nature & Existence*. London: Routledge.
- Hoffman, Joshua. (1980). 'Locke on Whether a Thing Can Have Two Beginnings of Existence.' *Ratio* Vol. 22
- Hooker, M. (1978a). *Descartes: Critical and Interpretative Essays*. London: The John Hopkins University Press.
- Hooker, M., ed. (1978b). 'Descartes' Denial of Mind-Body Identity.' In Hooker (1978a)
- Horton, P. & Miller, D. (1972). 'The Etiology of Multiple Personality.' *Comparative Psychology* Vol. 13.
- Hughes, Martin W. (1975). 'Personal Identity: A Defence of Locke.' *Philosophy* Vol. 50.
- Hume, David. (1978). *A Treatise of Human Nature* (ed. Selby-Bigge and Nidditch). Oxford: Clarendon Press.
- Hume, David. (1975). *Enquiries Concerning Human Understanding & Concerning the Principles of Morals* (ed. Nidditch). Oxford: Clarendon Press.
- Ishiguro, Hidé. (1980). 'The Primitiveness of the Concept of a Person.' in Van Straaten (1980).
- James, William. (1890). *Principles of Psychology*. (New York: Holt)
- Jessop, T.E. (1974). 'The Misunderstood Hume.' in Todd (1974).
- Johnston, Mark. (1987). 'Is There A Problem About Persistence?' *Aristotelian Society Supplement* Vol. 61
- Johnston, Mark. (1992). 'Constitution is Not Identity.' *Mind* Vol. 101
- Johnston, Mark. (1992). 'Reasons and Reductionism.' *Philosophical Review* Vol. 101
- Kemp-Smith, Norman. (1941). *The Philosophy of David Hume*. London: Macmillan.
- Kim, Jaegwon. (1997). 'The Mind-Body Problem: Taking Stock After Forty Years.' *Philosophical Perspectives* Vol. 11.
- King-Farlow, John. (1960). 'Substance, Substratum, and Personal Identity'. Review of *Metaphysics* Vol. 13
- Klibansky, R. & Mossner, E.C., eds. (1954). *New Letters of David Hume*. Oxford: Oxford University Press.
- Kornegay, R.J. (1985). 'Hume on Identity and Imperfect Identity.' *Dialogue* Vol. 24
- Krauss, Lawrence M. (1996). *The Physics of Star Trek*. London: Harper Collins.
- Kripke, Saul A. (1972). *Naming and Necessity*. Oxford: Basil Blackwell.
- Kripke, Saul A. (1977). 'Identity and Necessity.' in Schwartz (1977)
- Kurtzman, Howard S. (1983). 'Modern Conceptions of Memory.' *Philosophy and Phenomenological Research* Vol. 44
- Laird, John. (1932). *Hume's Philosophy of Human Nature*. London: Methuen.
- Law, Edmund. (1801). 'A Defence of Mr. Locke's Opinion Concerning Personal Identity.' In Vol. III of the 1801 edition of Locke's Works.
- Laycock, Henry. (1972). 'Some Questions of Ontology.' *Philosophical Review* Vol. 81
- Lewis, David. (1976). 'Survival and Identity.' in Rorty (1976a)
- Lewis, David. (1983a). *Philosophical Papers: Volume I*. Oxford: Oxford University Press.
- Lewis, David. (1983b). 'Postscript to "Survival and Identity."' in Lewis (1983a)
- Lewis, David. (1986). *On The Plurality of Worlds*. Oxford: Basil Blackwell.
- Lewis, Delmas. (1983). 'Dualism and the Causal Theory of Memory.' *Philosophy and Phenomenological Research* Vol. 44
- Lewis, H.D. (1976). *Contemporary British Philosophy: Forth Series*. London: George Allen and Unwin.
- Llinás, Rodolfo R., ed. (1988). *The Biology of The Brain: From Neurones to Networks*. New York: W.H. Freeman & Co.
- Locke, Don. (1971). *Memory*. (London: Macmillan)
- Locke, John. (1975). *An Essay Concerning Human Understanding* (ed. Peter H. Nidditch). Oxford: Oxford University Press.
- Locke, John. (1979). *The Correspondence of John Locke* (ed. De Beer). Oxford: Clarendon Press.
- Locke, John. (1988). *Essays on the Laws of Nature* (ed. W. Von Leyden). Oxford: Clarendon Press.
- Locke, John. (1990). *Drafts for the Essay Concerning Human Understanding & Other Philosophical Writings, Vol. 1* (eds. Peter H. Nidditch & G.A.J. Rogers). Oxford: Clarendon Press.
- Lovibond, S. E., & Williams, S. G., eds. (1996). *Essays for David Wiggins: Identity, Truth & Value (Aristotelian Society Series Vol. 16)*. Oxford: Basil Blackwell.
- Lowe, E. J. (1994). 'Ontological Dependency.' *Philosophical Papers* Vol. 23.

- Lowe, E. J. (1996). *Subjects of Experience*. Cambridge: Cambridge University Press.
- Lowe, E.J. (1982a). 'On Being A Cat.' *Analysis* Vol. 42
- Lowe, E.J. (1982b). 'On The Alleged Necessity of True Identity Statements.' *Mind* Vol. 91
- Lowe, E.J. (1982c). 'Reply To Geach.' *Analysis* Vol. 42
- Lowe, E.J. (1982d). 'The Paradox of the 1,001 Cats.' *Analysis* Vol. 42
- Lowe, E.J. (1983a). 'Instantiation, Identity and Constitution.' *Philosophical Studies* Vol. 44
- Lowe, E.J. (1983b). 'On The Identity Of Artefacts.' *Journal of Philosophy* Vol. 80
- Lowe, E.J. (1986). 'Sortal Terms and Absolute Identity.' *Australian Journal of Philosophy* Vol. 64
- Lowe, E.J. (1987). 'Lewis On Perdurantism Versus Endurance.' *Analysis* Vol. 47
- Lowe, E.J. (1988a). 'Substance, Identity And Time.' *Aristotelian Society Supplement* Vol. 62
- Lowe, E.J. (1988b). 'The Problems Of Intrinsic Change: Rejoinder To Lewis.' *Analysis* Vol. 48
- Lowe, E.J. (1989a). 'Impredicative Identity Criteria and Davidson's Criterion of Event Identity.' *Analysis* Vol. 49
- Lowe, E.J. (1989b). 'What Is A Criterion Of Identity?' *Philosophical Quarterly* Vol. 39
- Lowe, E.J. (1989c). *Kinds of Being: A Study of Individuation, Identity and the Logic of Sortal Terms*. Oxford: Basil Blackwell.
- Lowe, E.J. (1991a). 'One-Level Versus Two-Level Identity Criteria.' *Analysis* Vol. 51
- Lowe, E.J. (1991b). 'Real Selves: Persons as a Substantial Kind.' *Philosophy* Supp. Vol. 29
- Lowe, E.J. (1991c). 'Substance and Selfhood.' *Philosophy* Vol. 66
- Lowe, E.J. (1992). 'The Problem of Psychophysical Causation.' *Australian Journal of Philosophy* Vol. 70
- Lowe, E.J. (1993). 'Self, Reference and Self-Reference.' *Philosophy* Vol. 68
- Lowe, E.J. (1994). 'Primitive Substances.' *Philosophy and Phenomenological Research* Vol. 54.
- Lowe, E.J. (1995). 'Coinciding Objects: In Defence of the "Standard Account."' *Analysis* Vol. 55
- Lowe, E.J. (1995). *Locke on Human Understanding*. London: Routledge.
- Macdonald, G.F., ed. (1979). *Perception and Identity: Essays Presented to A.J. Ayer with his Replies to them*. London: Macmillan.
- Mackie, John L. (1976). *Problems from Locke*. Oxford: Clarendon Press.
- Mackie, John L. (1980). 'The Transcendental 'I'.' in Van Straaten (1980) and also Mackie (1985a).
- Mackie, John L. (1985a). *Persons and Values: Selected Papers Volume II*. Oxford: Clarendon Press.
- Mackie, John L. (1985b). 'Multiple Personality.' in Mackie (1985a)
- Madell, Geoffrey. (1981). *The Identity of the Self*. Edinburgh: Edinburgh University Press.
- Magnus B., and Wilbur, J. (eds.). (1969). *Cartesian Essays: A Collection of Critical Studies*. The Hague: Martinus Nijhoff.
- Markie, Peter. (1992). 'The Cogito and it's Importance.' in Cottingham (1992)
- Markie, Peter. (1994). 'Descartes' Concepts of Substance' in Cottingham (1994a)
- Martin, C.B. & Deutscher, Max. (1966). 'Remembering.' *Philosophical Review* Vol. 75
- Mattern, Ruth M. (1978). 'Re-interpreting Descartes on the Notion of the Union of the Mind and Body.' *Journal of the History of Philosophy* Vol. 16
- Mattern, Ruth. (1980). 'Moral Science and the Concept of Persons in Locke.' *Philosophical Review* Vol. 89.
- Matthews, Eric. (1977). 'Descartes and Locke on the Concept of a Person.' *Locke Newsletter* Vol. 8
- McCall, Catherine. (1990). *Concepts of Person: An Analysis of Concepts of Person, Self & Human Being*. Aldershot: Avebury.
- McCann, Edwin. (1987). 'Locke on Identity: Matter, Life, and Consciousness.' *Archiv Für Geschichte Der Philosophie* Vol. 69
- McGinn, Colin. (1982). *The Character of Mind*. Oxford: Oxford University Press.
- McIntyre, Jane L. (1989). 'Personal Identity and the Passions.' *Journal of the History of Philosophy* Vol. 27
- McIntyre, Jane L. (1993). 'Hume's Underground Self.' Reprinted in Tweyman (1995)
- Mendus, Susan. (1978). 'Personal Identity and Moral Responsibility.' *Locke Newsletter* No.9
- Mendus, Susan. (1980). 'Personal Identity: The Two Analogies in Hume.' *Philosophical Quarterly* Vol. 30
- Mijuskovic, Ben. (1971). 'Hume and Shaftesbury on the Self.' *Philosophical Quarterly* Vol. 21
- Montefiore, A., ed. (1973). *Philosophy and Personal Relations: An Anglo-French Study*. London: Routledge.
- Morice, G. P., ed. (1977). *David Hume: Bicentenary Papers*. Edinburgh: Edinburgh University Press.
- Morton, Adam. (1989). 'Why There is no Concept of a Person.' in Gill (1989)
- Nagel, Thomas. (1971). 'Brain Bisection and the Unity of Consciousness.' *Synthese* Vol. 22. Reprinted in Nagel (1979) and Perry (1975a)
- Nagel, Thomas. (1974) 'What is it like to be a bat?' *Philosophical Review*, Vol. 83. Reprinted in Nagel (1979)
- Nagel, Thomas. (1979). *Mortal Questions*. Cambridge: Cambridge University Press.
- Nason, John W. (1942). 'Leibniz and the Logical Argument for Individual Substances.' *Mind* Vol. 51. Reprinted in Woolhouse (1981)

- Nathanson, Stephen. (1976). 'Hume's Second Thoughts on the Self.' *Hume Studies* Vol. 2
- Neujahr, P.J. (1978). 'Hume on Identity.' *Hume Studies* Vol. 4
- Noonan, Harold W. (1981). 'Wiggins' Second Thoughts on Identity.' *Philosophical Quarterly* Vol. 31
- Noonan, Harold W. (1978a). 'Locke on Personal Identity.' *Philosophy* Vol. 53.
- Noonan, Harold W. (1978b). 'Can One Thing Become Two?' *Philosophical Studies* Vol. 30
- Noonan, Harold W. (1982). 'Williams on "The Self and The Future."' *Analysis* Vol. 42.
- Noonan, Harold W. (1986). 'Reply to Simons on Coincidence.' *Mind* Vol. 95
- Noonan, Harold W. (1989). *The Problems of Philosophy Their Past and Present: Personal Identity*. London: Routledge.
- Noonan, Harold W. (1998). 'Animalism "Versus" Lockeanism: A Current Controversy.' *Philosophical Quarterly* Vol. 48
- Norton, David Fate. (1982). *David Hume: Common-Sense Moralist, Sceptical Metaphysician*. Princeton: Princeton University Press.
- Noxon, J. (1969). 'Senses of Identity in Hume's *Treatise*.' *Dialogue* Vol. 8
- Nozick, Robert. (1981). *Philosophical Investigations*. Cambridge, Massachusetts: Belknap Press.
- O'Brien, Lucy F. (1995a). 'Evans on Self-Identification.' *Noûs* Vol. 29.
- O'Brien, Lucy F. (1995b). 'The Problem of Self-Identification.' *Proceedings of the Aristotelian Society* Vol. 95.
- Oaklander, L. Nathan. (1987). 'Parfit, Circularity, and the Unity of Consciousness.' *Mind* Vol. 96.
- Odegard, Douglas. (1970). 'Locke and Mind-Body Dualism.' *Philosophy* Vol. 45.
- Olson, Eric T. (1994). 'Is Psychology Relevant to Personal Identity?' *Australian Journal of Philosophy* Vol. 72
- Olson, Eric T. (1995a). 'Human People or Human Animals.' *Philosophical Studies* Vol. 80
- Olson, Eric T. (1995b). 'Why I Have No Hands.' *Theoria* Vol. 61
- Olson, Eric T. (1996). 'Composition and Coincidence.' *Pacific Philosophical Quarterly* Vol. 77
- Olson, Eric T. (1997a). *The Human Animal*. Oxford: Oxford University Press
- Olson, Eric T. (1997b). 'Was I Ever a Fetus?' *Philosophy and Phenomenological Research* Vol. 57
- Olson, Eric T. (1997c). 'Dion's Foot.' *Journal of Philosophy* Vol. 94
- Olson, Eric T. (1997d). 'Relativism and Persistence.' *Philosophical Studies* Vol. 88
- Palma, A.B. (1964). 'Memory and Personal Identity.' *Australian Journal of Philosophy* Vol. 42
- Parfit, Derek. (1973). 'Later Selves and Moral Principles.' in Montefiore (1973)
- Parfit, Derek. (1976). 'Lewis, Perry, and What Matters.' in Rorty (1976a)
- Parfit, Derek. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, Derek. (1995). 'The Unimportance of Identity.' In Harris (1995).
- Patten, Steven C. (1976a). 'Change, Identity and Hume.' *Dialogue* Vol. 15
- Patten, Steven C. (1976b). 'Hume's Bundles, Self-Consciousness, and Kant.' *Hume Studies* Vol. 2
- Pears, David. (1993). 'Hume on Personal Identity.' *Hume Studies* Vol. 19
- Penelhum, Terence. (1955). 'Hume on Personal Identity.' *Philosophical Review* Vol. 64
- Penelhum, Terence. (1959). 'Personal Identity, Memory, Survival.' *Journal of Philosophy* Vol. 56
- Penelhum, Terence. (1967). 'Personal Identity.' In Edwards (1967).
- Penelhum, Terence. (1975). 'Hume's Theory of Self Revisited.' *Dialogue* Vol. 14. Reprinted in Tweyman (1995)
- Penelhum, Terence. (1976). 'Self-Identity and Self-Regard.' In Rorty (1976a)
- Perry, John, ed. (1975a). *Personal Identity*. Berkeley: University of California Press.
- Perry, John. (1970). 'The Same F.' *Philosophical Review* Vol. 79
- Perry, John. (1975b). 'Personal Identity, Memory, and the Problem of Circularity.' in Perry (1975a)
- Perry, John. (1975c). 'The Problem of Personal Identity.' in Perry (1975a)
- Perry, John. (1976a). Review of *Problems of the Self*. *Journal of Philosophy* Vol. 73
- Perry, John. (1976b). 'The Importance of Being Identical.' in Rorty (1976a)
- Perry, John. (1978a). 'Relative Identity and Number.' *Canadian Journal of Philosophy* Vol. 8
- Perry, John. (1978b). *A Dialogue on Personal Identity and Immortality*. Indianapolis: Hackett Publishing Company.
- Pike, N. (1967). 'Hume's Bundle Theory of the Self: A Limited Defence.' *American Philosophical Quarterly* Vol. 4. Reprinted in Tweyman (1995).
- Pollock, John. (1989). *How To Build a Person: A Prolegomenon*. Cambridge, MA: MIT Press
- Puccetti, Roland. (1973). 'Remembering the Past of Another.' *Canadian Journal of Philosophy* Vol. 2
- Puccetti, Roland. (1977). 'Memory and Self: A Neuropathological Approach.' *Philosophy* Vol. 52.
- Putnam, F.W. (1989). *Diagnosis and Treatment of Multiple Personality Disorder*. New York: The Guildford Press.
- Quine, W.V.O. (1948). 'On What There Is.' in Quine (1953)
- Quine, W.V.O. (1953). *From a Logical Point of View*. New York: Harper and Row.
- Quine, W.V.O. (1958a). *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Quine, W.V.O. (1958b). 'Speaking of Objects.' In Quine (1958a)

- Quinton, Anthony. (1962). 'The Soul.' *Journal of Philosophy* Vol. 59. Reprinted in Perry (1975a)
- Quinton, Anthony. (1973). *The Nature of Things*. London: Routledge.
- Rabb, Douglas J., (1977). 'Reflection, Reflexion, and Introspection.' *Locke Newsletter* No. 8
- Reid, Thomas. (1785a). 'Of Identity.' in Perry (1975a).
- Reid, Thomas. (1785b). 'Of Mr. Locke's Account of Our Personal Identity.' in Perry (1975a).
- Rey, Georges. (1976). 'Survival.' in Rorty (1975a)
- Robinson, Daniel N. & Beauchamp, Tom L. (1978). 'Personal Identity: Reid's Answer to Hume.' *The Monist* Vol. 61
- Robison, Wade L. (1974). 'Hume on Personal Identity.' *Journal of the History of Philosophy* Vol. 12. Reprinted in Tweyman (1995)
- Rogers, G.A.J., ed. (1994). *Locke's Philosophy: Content and Context*. Oxford: Clarendon Press.
- Rorty, Amélie Oksenberg, ed. (1976a). *The Identities of Persons*. Berkeley: University of California Press.
- Rorty, Amélie Oksenberg, ed. (1986a). *Essays on Descartes' Meditations*. Berkeley: University of California Press.
- Rorty, Amélie Oksenberg. (1986b). 'The Structure of Descartes' *Meditations*.' In Rorty (1986a).
- Rorty, Amélie Oksenberg. (1986c). 'Cartesian Passions and the Union of Mind and Body.' In Rorty (1986c).
- Rorty, Amélie Oksenberg. (1976b). 'A Literary Postscript: Characters, Persons, Selves, Individuals.' In Rorty (1976a).
- Rosen, Deborah. (1975). 'An Argument for the Logical Notion of a Memory Trace.' *Philosophy of Science* Vol. 42
- Rosenkrantz, Gary S. & Hoffman, Joshua. (1997). *Substance: It's Nature & Existence*. London: Routledge.
- Rovane, Carol. (1987). 'The Epistemology of First-Person Reference.' *Journal of Philosophy* Vol. 84
- Rovane, Carol. (1990). 'Branching Self-Consciousness.' *Philosophical Review* Vol. 99
- Rovane, Carol. (1993). 'Self-Reference: The Radicalization of Locke.' *Journal of Philosophy* Vol. 90
- Rovane, Carol. (1999). *The Bounds of Agency: An Essay of Revisionary Metaphysics*. Princeton, NY: Princeton University Press
- Rozemond, Marleen. (1998). *Descartes's Dualism*. Cambridge, Massachusetts: Harvard University Press.
- Ryle, Gilbert. (1949). *The Concept of Mind*. London: Penguin.
- Scarre, Geoffrey. (1983). 'What Was Hume's Worry About Personal Identity?' *Analysis* Vol. 43
- Schechtman, Marya. (1990). 'Personhood and Personal Identity.' *Journal of Philosophy* Vol. 87
- Schechtman, Marya. (1994). 'The Same and the Same: Two Views of Psychological Continuity.' *American Philosophical Quarterly* Vol. 31
- Schechtman, Marya. (1994). 'The Truth About Memory.' *Philosophical Psychology* Vol. 7
- Schouls, Peter A. (1980). *The Imposition of Method: A Study of Descartes and Locke*. Oxford: Clarendon Press.
- Schwartz, Stephen P., ed. (1977). *Naming, Necessity, and Natural Kinds*. Ithaca: Cornell University Press.
- Searle, John R. (1979). 'What Is an Intentional State?' *Mind* Vol. 88
- Shoemaker, Sydney & Swinburne, Richard. (1984). *Personal Identity*. Oxford: Basil Blackwell.
- Shoemaker, Sydney. (1959). 'Personal Identity and Memory.' *Journal of Philosophy* Vol. 56. Reprinted in Perry (1975a).
- Shoemaker, Sydney. (1963). *Self-Knowledge and Self-Identity*. Ithaca: Cornell University Press.
- Shoemaker, Sydney. (1969). 'Comments on Chisholm's 'The Loose and Popular and the Strict and Philosophical Senses of Identity.' in Care & Grimm (1969).
- Shoemaker, Sydney. (1970). 'Persons and their Pasts.' *American Philosophical Quarterly* Vol. 7, reprinted in Shoemaker (1984a) with revisions.
- Shoemaker, Sydney. (1975). 'Functionalism and Qualia.' *Philosophical Studies* Vol. 27. Reprinted in Shoemaker (1984a).
- Shoemaker, Sydney. (1976). 'Embodiment and Behaviour.' in Rorty (1976a)
- Shoemaker, Sydney. (1981a). 'Are Absent Qualia Possible?' *The Philosophical Review* 90. Reprinted in Shoemaker (1984a).
- Shoemaker, Sydney. (1981b). 'Some Varieties of Functionalism.' *Philosophical Topics* 12. Reprinted in Shoemaker (1984a).
- Shoemaker, Sydney. (1981c). 'The Inverted Spectrum.' *The Journal of Philosophy* 74. Reprinted in Shoemaker (1984a).
- Shoemaker, Sydney. (1984a). *Identity, Cause and Mind*. (Cambridge: Cambridge University Press).
- Shoemaker, Sydney. (1984b). 'Personal Identity: A Materialist's Account.' In Shoemaker & Swinburne (1984).
- Shoemaker, Sydney. (1984c). 'Reply to Swinburne.' In Shoemaker and Swinburne (1984).
- Shoemaker, Sydney. (1986). 'Introspection and the Self.' *Midwest Studies in Philosophy* Vol. 10. Reprinted in Shoemaker (1996a).

- Shoemaker, Sydney. (1988). 'On Knowing One's Own Mind.' *Philosophical Perspectives* Vol. 2. Reprinted in Shoemaker (1996a).
- Shoemaker, Sydney. (1990). 'First Person Access.' *Philosophical Perspectives* Vol. 4. Reprinted in Shoemaker (1996a).
- Shoemaker, Sydney. (1994a). 'Self-Knowledge and "Inner-Sense: Three Lectures."' *Philosophy and Phenomenological Research* Vol. 54. Reprinted in Shoemaker (1996a).
- Shoemaker, Sydney. (1994b). 'The First-Person Perspective.' *Proceedings and Addresses of the American Philosophical Association* Vol. 68. Reprinted in Shoemaker (1996a).
- Shoemaker, Sydney. (1996a). *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Shoemaker, Sydney. (1996b). 'Unity of Consciousness and Consciousness of Unity.' In Shoemaker (1996a).
- Shoemaker, Sydney. (1997). 'Self and Substance.' *Philosophical Perspectives* Vol. 11
- Simons, Peter M. (1986). 'Unkindly Coincidences.' *Mind* Vol. 95
- Simons, Peter. (1985). 'Coincidence of Things of a Kind.' *Mind* Vol. 94
- Simons, Peter. (1987). *Parts: A Study in Ontology*. Oxford: Clarendon Press
- Smart, Brian. (1977). 'How Can Persons be Ascribed M-Predicates?' *Mind* Vol. 86
- Smith, Peter. (1989). 'Human Person.' In Gill (1989)
- Snowdon, Paul F. (1989). 'Persons, Animals and Ourselves.' In Gill (1989)
- Snowdon, Paul F. (1991). 'Personal Identity & Brain Transplants in Human Beings.' In Cockburn (1991)
- Snowdon, Paul F. (1996). 'Persons and Personal Identity.' In Lovibond & Williams (1996)
- Snowdon, Paul F. (1998). 'Persons, Animals & Bodies.' In Bermudez (1998)
- Sorrell, Tom. (1994). 'Descartes's Modernity' in Cottingham (1994a)
- Sosa, Ernest. (1990). 'Surviving Matters.' *Noûs* Vol. 24
- Southgate, Susan. (1994). *Personal Identity: The Simple View*. Durham University Thesis
- Spicker, ed. (1970). *The Philosophy of the Body*. Chicago: Quadrant Books.
- Spiegel, David. (1993a). 'Letter to the Executive Council of ISSMP&D 20<sup>th</sup> May 1993.' *International Society for the Study of Multiple Personality and Dissociation Newsletter* Vol. 11
- Spiegel, David. (1993b). 'Lecture to the Tenth International Conference on Multiple Personality/Dissociate States.' Tape VII-860-93. Alexandria, Va: Audio Transcripts.
- Stone, Jim. (1981). 'Hume on Identity: A Defence.' *Philosophical Studies* Vol. 40
- Strawson, G. (1989). *The Secret Connexion: Causation, Realism, & David Hume*. Oxford: Clarendon Press.
- Strawson, P.F. (1959). *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen.
- Strawson, P.F. (1976). "Entity and Identity." in Lewis (1976)
- Stroud, Barry. (1977). *Hume*. London: Routledge
- Sutton, John. (1998). *Philosophy and Memory Traces*. Cambridge: Cambridge University Press.
- Swinburne, Richard (with Shoemaker, Sydney) *see* Shoemaker & Swinburne (1984).
- Swinburne, Richard. (1984a). 'Personal Identity: The Dualist Theory.' In Shoemaker and Swinburne (1984).
- Swinburne, Richard. (1984b). 'Reply to Shoemaker.' In Shoemaker and Swinburne (1984).
- Taylor, Charles. (1976). 'Responsibility for Self.' in Rorty (1976a)
- Tennant, R.C. (1982). 'The Anglican Response to Locke's Theory of Personal Identity.' *Journal of the History of Ideas* Vol. 43
- Thiel, Udo. (1997). "'Epistemologism" and Early Modern Debates About Individuation and Identity.' *British Journal of the History of Philosophy* Vol. 5
- Thiel, Udo. (1998a). 'Individuation.' In Garber & Ayers (1998)
- Thiel, Udo. (1998b). 'Personal Identity.' In Garber & Ayers (1998)
- Thomas, Janice. (1979). 'On a Supposed Inconsistency in Locke's Account of Personal Identity.' *Locke Newsletter* No.10
- Tipton, I.C., ed. (1977). *Locke on Human Understanding*. Oxford: Oxford University Press
- Todd, William B., ed. (1974). *Hume and the Enlightenment: Essays Presented to Ernest Campbell Mossner*. Edinburgh: Edinburgh University Press.
- Traiger, S. (1985). 'Hume on Finding as Impression of the Self.' *Hume Studies* Vol. 11. Reprinted in Tweyman (1995).
- Troyer, John. (1975). 'Locke on the Names of Substances.' *Locke Newsletter* No. 6
- Tweyman, Stanley, ed. (1995). *Critical Assessment's: Hume Vol. III*. London: Routledge.
- Unger, Peter. (1979). 'I Do Not Exist.' in MacDonald (1979)
- Urmson, J.O. (1967). 'Memory and Imagination.' *Mind* Vol. 76
- Urmson, J.O. (1971). 'Memory and Imagination.' *Mind* Vol. 80
- Van Inwagen, Peter. (1978). 'The Possibility of Resurrection.' *International Journal of Philosophy & Religion* Vol. 9
- Van Inwagen, Peter. (1981). 'The Doctrine of Arbitrary undetached Parts.' *Pacific Philosophical Quarterly* Vol. 62

- Van Inwagen, Peter. (1990). *Material Beings*. Ithaca: Cornell University Press.
- Van Inwagen, Peter. (1997). 'Materialism and the Psychological-Continuity Account of Personal Identity.' *Philosophical Perspectives* Vol. 11.
- Van Straaten, Zak, ed. (1980). *Philosophical Subjects: Essays Presented to P.F. Strawson*. Oxford: Clarendon Press.
- Vesey, Godfrey. (1974). *Personal Identity*. London: Macmillan.
- Von Leydon, W. (1957). 'Hume and "Imperfect Identity."' *Philosophical Quarterly* Vol. 7
- Voss, Stephen. (1994). 'Descartes: The End of Anthropology' in Cottingham (1994a).
- Ward, Andrew. (1985). 'Hume, Demonstratives, and Self-Ascription of Identity.' *Hume Studies* Vol. 11
- Waxman, Wayne. (1994). *Hume's Theory of Consciousness*. Cambridge: Cambridge University Press.
- Wieand, Jeffrey. (1980). 'Locke on Memory.' *Locke Newsletter* No. 11
- Wiggins, David. (1965). 'Identity Statements.' in Butler (1965)
- Wiggins, David. (1967). *Identity and Spatio-Temporal Continuity*. Oxford: Basil Blackwell.
- Wiggins, David. (1976). 'Locke, Butler and the Stream of Consciousness: And Men as Natural Kind.' in Rorty (1976a)
- Wiggins, David. (1979a). 'The Concern to Survive.' *Midwest Studies in Philosophy* Vol. 4
- Wiggins, David. (1979b). 'Ayer on Monism, Pluralism and Essence.' in Macdonald (1979)
- Wiggins, David. (1980). *Sameness and Substance*. Oxford: Basil Blackwell.
- Wilbur MacKenzie, Ann. (1994). 'The Reconfiguration of Sensory Experience' in Cottingham (1994a).
- Wilbur, James B. (1969). 'The *Cogito*, An Ambiguous Performance.' in Magnus & Wilbur (1969)
- Wilkes, Kathleen V. (1984). 'Is Consciousness Important?' *British Journal of Philosophy Science* Vol. 35.
- Wilkes, Kathleen V. (1988). *Real People: Personal Identity Without Thought Experiments*. Oxford: Clarendon.
- Williams, Bernard. (1956). 'Personal Identity and Individuation.' *Proceedings of the Aristotelian Society* Vol. 18. Reprinted in Williams (1973a)
- Williams, Bernard. (1961a). 'Mr. Strawson on Individuals.' *Philosophy* Vol. 36. Reprinted in Williams (1973a)
- Williams, Bernard. (1961b). 'Personal Identity and Bodily Continuity - A Reply.' *Analysis* 21. Reprinted in Williams (1973a)
- Williams, Bernard. (1962). 'The Certainty of the *Cogito*.' *Cahiers de Royaumont, Philosophie No. 4: La Philosophie Analytique*. Reprinted in Doney (1970)
- Williams, Bernard. (1966). 'Imagination and the Self.' British Academy Annual Philosophical Lecture. Reprinted in Williams (1973a)
- Williams, Bernard. (1970a). 'Are Persons Bodies?' in Spicker (1970) and reprinted in Williams (1973a)
- Williams, Bernard. (1970b). 'The Self and the Future.' *Philosophical Review* Vol. 79. Reprinted in Williams (1973a)
- Williams, Bernard. (1973a). *Problems of the Self: Philosophical Papers 1956-1972*. Cambridge: Cambridge University Press.
- Williams, Bernard. (1973b). 'The Makropulos Case: Reflections on the Tedium of Immortality.' In Williams (1973a)
- Williams, Bernard. (1976). 'Persons, Character and Morality.' in Rorty (1976a)
- Williams, Bernard. (1979). 'Another Time, Another Place, Another Person.' in Macdonald (1979)
- Williams, Bernard. (1994). 'Descartes and the Historiography of Philosophy' in Cottingham (1994a).
- Wilson, Margaret Daulet. (1978). *Descartes*. London: Routledge.
- Winkler, Kenneth P. (1991a). 'The New Hume.' *The Philosophical Review* Vol. 50.
- Winkler, Kenneth P. (1991b). 'Locke on Personal Identity.' *Journal of the History of Philosophy* Vol. 29.
- Wollheim, Richard. (1979). 'Memory, Experiential Memory and Personal Identity.' in *Perception and Identity* (ed. Macdonald).
- Wood, David. (1979). 'Hume on Identity and Personal Identity.' *Australian Journal of Philosophy* Vol. 57
- Woods, M.J. (1965). 'Identity and Individuation.' in *Analytical Philosophy: Second Series* (ed. Butler).
- Woolhouse, R.S. (1971). *Locke's Philosophy of Science and Knowledge*. Oxford: Basil Blackwell.
- Woolhouse, R.S. (1983). *Locke*. Minnesota: Minnesota University Press.
- Woolhouse, R.S. (1993). *Descartes, Spinoza, Leibniz: The Concept of Substance in Seventeenth Century Metaphysics*. London: Routledge.
- Wright, John P. (1983). *The Sceptical Realism of David Hume*. Minneapolis: University of Minnesota Press.
- Yolton, John W. (1985). *John Locke: An Introduction*. Oxford: Basil Blackwell.
- Yourgrau, Palle, ed. (1990). *Demonstratives*. Oxford: Oxford University Press.
- Zemach, E.M. (1983). 'Memory: What It Is, and What It Cannot Possibly Be.' *Philosophy and Phenomenological Research* Vol. 44
- Zimmerman, Dean W. (1995). 'Theories of Masses and Problems of Constitution.' *Philosophical Review* Vol. 104

- Zimmerman, Dean W. (1996a). 'Could Extended Objects be made out of Simple Parts? An Argument for Atomless Gunk.' *Philosophy and Phenomenological Research* Vol. 56
- Zimmerman, Dean W. (1996b). 'Indivisible Parts and Extended Objects: Some Philosophical Episodes from Topology's Prehistory.' *Monist* Vol. 79
- Zimmerman, Dean W. (1997). 'Coincident Objects: Could a 'Stuff Ontology' Help?' *Analysis* Vol. 57

