

University of St Andrews



Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

The Persistence of Persons

SUNE HOLM

Submitted for the degree of PhD

Department of Philosophy

University of St. Andrews

June 8th 2006



Supervisor: Professor Sarah Broadie
Date of final examination: August 28th, 2006
Examiners: Dr Katherine Hawley
Professor Robin LePoidevin

I, Sune Holm, hereby certify that this thesis, which is approximately 78000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

June 8th 2006. Signature of candidate

I was admitted as a research student in September 2001 and as a candidate for the degree of PhD in September 2001; the higher study for which this is a record was carried out in the University of St. Andrews between 2001 and 2006.

June 8th 2006 Signature of candidate

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

June 8th 2006 signature of candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit the thesis in application for that degree.

June 8th 2006

signature of supervisor

ACKNOWLEDGMENTS

Writing this thesis I've learned a lot from the many lively and stimulating philosophers I've been privileged to meet in St. Andrews. I'd like to thank a number of people who have commented on written material during my research: John Skorupski, Katherine Hawley, Josh Parsons, Simon Prosser, Daniel Nolan, Alan Millar, Kent Hurtig, Iwao Hirose, Joe Diekemper, and Marcus Schlosser.

I'd also like to thank Rafaelle Rodogno, Patrice Phillie, Cyrus Panjvani, Simon Robertson, Philip Ebert, Jorn Sonderholm, Brian McElwee, Marcus Rossberg, Eline Busck, Nikolaj Jang Pedersen, Ross Cameron, Robbie Williams, Michael Weh, Jamie Dow, Chiara Tabet, Darren McDonald, Enzo Rossi, Jesper Kallestrup, Patrick Greenough, and Claus Hansen who have all given me much valuable feedback and made my time in St. Andrews such a memorable one.

I am also very grateful for the meticulous and inspiring criticism I've received from my principal supervisor, Professor Sarah Broadie, whose comments have been extremely helpful. Finally I would like to thank the Danish Research Council for awarding me a PhD scholarship.

ABSTRACT

In this thesis I discuss what it takes for one and the same person to exist at different times. I begin by presenting how to understand this question as a philosophical question and argue that we should accept a so-called psychological criterion of personal identity.

In the course of the thesis I discuss what I take to be the most pertinent objections facing accepting a psychological criterion. I consider the role of prudential concern in arguments in favour of the psychological criterion and defend it against a recent criticism. I then turn to the complaint that the psychological criterion is viciously circular in that it accounts for personal identity in terms of psychological relations which presuppose personal identity and defend the appeal to the central notion of quasi-memory.

An important part of the thesis concerns how a proponent of the psychological criterion can deal with problems arising from the possibility that what is plausibly thought to be two distinct persons may stand in the relation of psychological continuity to what is plausibly claimed to be a single person. I argue that a proponent of the psychological criterion should accept the thesis that numerical identity is temporary and provide a defence of the temporary identity thesis.

In the final chapters of the thesis I discuss how a proponent of the psychological criterion can accommodate the view that persons are enduring material beings, paradigmatically human organisms and submit that accepting the temporary identity thesis not only provides an attractive way of maintaining the claim that paradigm persons are identical with human organisms, but also sheds light on the role of personal identity in memory knowledge and prudential concern.

Contents

INTRODUCTION	1
1. THE PROBLEM OF PERSONAL IDENTITY	4
1.1 Introduction.....	4
1.2 The Identity Question	4
1.3 The Formal Character of Identity	6
1.4 Persistence, Change , and Indiscernibility	10
1.5 Perdurantism and Endurantism	12
1.6 Personal Identity and Personal Histories.....	21
1.7 The Perdurantist Interpretation	24
1.8 The Endurantist Interpretation	26
1.9 Broad and Narrow Questions about Personal Identity	30
1.10 Conclusion	35
2. THE PSYCHOLOGICAL CRITERION OF PERSONAL IDENTITY	37
2.1 Introduction.....	37
2.2 In Favour of the Psychological Criterion	37
2.3 Psychological Continuity	44
2.4 Circularity, Psychological Continuity, and Quasi-Memory	46
2.5 The Fission Problem.....	48
2.6 The Nature and Persistence of Persons	49
2.7 Imaginary Cases	50
2.8 Conclusion	53
3. PERSONAL AND PRACTICAL IDENTITY	54
3.1 Introduction.....	54
3.2 Persistence and Practical Concern	56
3.3 Parfit's Argument from Fission	58
3.4 The Practical Identity of Persons	60

3.5 A Response: Identity is a Sufficient Condition for What Matters even if not a Necessary Condition	65
3.6 Conclusion	67
4. THE CIRCULARITY OBJECTION.....	68
4.1 Introduction.....	68
4.2 Introducing the Circularity Objection.....	68
4.3 Memory.....	71
4.4 Memory and Immunity to Error through Misidentification.....	78
4.5 Quasi-Memory and the Previous Awareness Condition.....	83
4.6 The Content and Independent Intelligibility of Quasi-Memory.....	85
4.7 Conclusion.....	90
5. FISSION, CLOSEST CONTINUERS, AND MULTIPLE OCCUPANCY	92
5.1 Introduction.....	92
5.2 A Case of Fission.....	94
5.3 The Reduplication Problem.....	97
5.4 Five Ways to Respond to Fission.....	100
5.5 Nozick's Closest Continuer Schema: 'To be something later, is to be its closest continuer'	104
5.6 The Multiple Occupancy Thesis	111
5.7 Conclusion.....	116
6. FISSION AND THE TEMPORARY IDENTITY THESIS.....	117
6.1 Temporary Identity and Lifetimes	117
6.2 Restricting Transitivity and Leibniz's Law	124
6.3 Responding to the T-Leibniz's Law Argument (TLL-Argument)	127
6.4 Can Things Change their Time-Indexed Properties?.....	132
6.5 Replacement and the Only x and y Principle	135
6.6 The Reoccurrence of the Transitivity Argument and Trans-Temporal identity.....	138
6.7 What happens to Brown?	141

6.8 On Perry's Lifetime View	146
6.9 Short Remark on Contingent Identity	151
6.10 Conclusion	153
7. INTRODUCING THE PROBLEM OF MULTIPLE-KIND FISSION	155
8. PERSONALISM AND ANIMALISM	167
8.1 Introduction.....	167
8.2 Personalism and the Problem Cases.....	168
8.3 Animalism and Personalism	173
8.4 Is Personalism Committed to a Metaphysical Impossibility?	183
8.5 Personalism and Sortal Dominance.....	185
8.6 Olson's Locomotor Argument	190
8.7 Response to Olson's Locomotor Argument.....	197
8.8 Does Biological Death entail Ceasing to Exist for any Objects recognised by our Ordinary Ontology?.....	202
8.9 Conclusion.....	204
9. THE CONSTITUTION VIEW	207
9.1 Introduction.....	207
9.2 Coincidentalism and 'Too Many'-Style Objections.....	207
9.3 Shoemaker's View: Animals can't Think	215
9.4 Baker's Idea of Constitution.....	218
9.5 Derivative and Nonderivative Properties.....	221
9.6 Counting by Constitution – Criticism of the Constitution View	225
9.7 Conclusion	230
10. THE TEMPORARY IDENTITY SOLUTION	232
10.1 Introduction.....	232
10.2 Temporary Identity and Multiple-Kind Fission.....	233
10.3 Temporary Identity, Animalism, Personalism, and Constitutionalism	243
10.4 Temporary Identity and the Circularity Problem	250

10.5 Temporary Identity and What Matters in Survival	253
10.6 Conclusion	255
BIBLIOGRAPHY	257

Introduction

'A wonderful fact to reflect upon, that every human creature is constituted to be that profound secret to every other' (Dickens 1859, p. 21.).

In this thesis I defend a psychological approach to the persistence of persons according to which personal identity is constituted by a relation of psychological continuity. In chapter 1 I present the question which the psychological approach is claimed to answer and in chapter 2 I give my argument for finding that approach attractive.

I then consider criticism from Eric Olson (1997) that the arguments in favour of the psychological approach are based on the premise that our practical concerns can serve as a guide to our identity over time. Olson appeals to Parfit's notorious claim that personal identity is not what matters in prudential concern, and thus the pattern of such concerns does not have authority with respect to metaphysical views about persons and their identity. In chapter 3 I argue that Olson's criticism presupposes more than Parfit's notorious argument gives him.

The psychological approach is supposed to provide a non-trivial and informative criterion of personal identity and it has been a worry for its proponents that it is viciously circular in that it accounts for personal identity in terms of psychological relations such as memory relations. The worry is that in order to determine whether a person e.g. remembers some past event, it must be established whether he is identical to the person whose experience he apparently remembers. Proponents of the psychological approach typically appeal to a notion of quasi-memory to counter the circularity objection and in chapter 4 I engage in that discussion defending the appeal to quasi-memory.

In chapter 5 I introduce the important problem raised by the possibility of fission. The formal structure of the relation of psychological continuity doesn't rule out that two distinct persons picked out at one time are both psychologically continuous with a single person picked out at a past time. I consider some popular ways in which proponents of the psychological approach have responded to this kind of case which I ultimately reject. I then move on to suggest that the fission case is best described as a case in which two persons are identical at one time and distinct at another. In chapter 6 I go some way to answer objections to this way of describing fission, and conclude that the temporary identity thesis is a sound one and that a proponent of the psychological approach should adopt it.

Paradigmatic persons such as you and I are naturally said to be *human* persons, as opposed to robotic or angelic persons, who belong to the species *homo sapiens* forming part of biological taxonomy. In chapter 7 I introduce a case of 'multiple-kind' fission which exposes a tension between the psychological approach and the plausible claim that we are human organisms. In multiple-kind fission a human person has a biological continuer which is not his psychological continuer and which is plausibly said to be the same human organism which was once a person. In the subsequent three chapters I discuss ways in which one might respond to the multiple-kind fission scenario.

In chapter 8 I discuss whether one or the other kind of continuity is metaphysically dominant in cases where they come apart and conclude that they're on par. In chapter 9 I consider the coincidental suggestion that we should take the possibility of multiple-kind fission to show that the claim that paradigm persons are human organisms isn't to be understood as the claim that they are strictly speaking identical, but as a claim that they stand

in an intimate relation of coincidence which is not identity. I argue that the coincidentalist position is unsatisfactory.

Finally I present how a temporary identity thesis allows the proponent of the psychological approach to deal with the multiple-kind fission case in a way parallel to the way it deals with the standard fission case, and I end by pointing out how appeal to temporary identity also has interesting consequences for how to understand the claim that personal identity isn't what matters in practical concern discussed in chapter 3, and for how to understand the notion of quasi-memory discussed in chapter 4. I conclude that accepting the temporary identity thesis provides the proponent of the psychological approach to the persistence of persons with an excellent basis for responding to important objections within a materialist and endurantist framework.

1. The Problem of Personal Identity

1.1 INTRODUCTION

In this chapter I introduce the question about what constitutes personal identity over time, the formal character of identity, the problem of change and how one might deal with it within an endurantist and perdurantist framework respectively. I then discuss how endurantism and perdurantism conceive of the problem of identity over time, and assuming an endurantist framework I find that the best way to characterise the problem is as being about what unites person-involving events into a personal history. I end with some critical remarks about Olson's distinction between 'narrow' and 'broad' questions about personal identity.

1.2 THE IDENTITY QUESTION

The question about identity over time sounds simple enough: What does it take for an individual thing existing at one time to be identical with an individual thing existing at some other time? Or, alternatively: What does it take for one and the same individual to exist at different times?¹

In this thesis the focus is what it takes for persons to persist. One reason why the question about the identity of persons is particularly interesting is of course that we are ourselves paradigmatic examples of persons, and thus the question about personal identity is the question about what it takes for *us* to exist over time.

Furthermore, we seem to take special interest in the welfare of the persons existing at other times with whom we are identical. When I save up money

¹ As it turns out, I will reject that individuals existing at different times can stand in a relation of identity across time, while this is not to deny that one and the same thing may exist at different times. However, for the time being I will assume that the two questions are equivalent such that in order for an individual to exist at different times there is a relation of identity across time.

for my pension or quit smoking, the concern I have is prudential in that I make a present sacrifice in order for myself, the person presently sacrificing, to gain a larger benefit in the future. Prudential concern is concern for my future self and my own future welfare, and thus the future person, whom I find it justified to benefit on the basis of prudential considerations, is supposed to be identical to me. It would be imprudent of me to save up money for someone else's pension (which is of course not to say that it would be wrong). So personal identity seem to be a key relation when it comes to prudential concern.

Personal identity also has relevance for moral judgments. As Locke noted, the concept of a person is a 'forensic' concept, and personal identity is important for 'appropriating actions and their merit' (Locke 1996, II, xxvii, 26). Thus a central part of a trial is to ascertain whether the defendant is identical to the person who did some unlawful action in the past. The evidence given during the trial is aimed at establishing whether the defendant is identical with some past person or not. If the identity is established this will merit responsibility and punishment for the unlawful action, if identity isn't established, the defendant is acquitted.

Now it is important to point out, that when a jury is presented with evidence for or against confirming an identity claim, which states that the defendant is indeed identical to the person who committed the unlawful act in question, the evidence is only *evidence* of personal identity. It is not, however, what personal identity will *consist* in. The distinction between what constitutes personal identity and what is good evidence for personal identity is important. The evidence is evidence that certain facts which constitute personal identity do or do not obtain. Still, the philosophical question and debate concerning personal identity has to do with the metaphysical and semantic issue about what facts the evidence is evidence for, i.e. it is about what facts must obtain in order for an identity judgment to be true. Thus

fingerprints and eye witnesses may all form part of the evidence for a personal identity judgment, but neither will be said to constitute personal identity.

We may of course ask about the identity of all sorts of things. In general, when asking identity questions about material objects, we ask whether an individual x of some kind K existing at a time t is the same K as some material being y of the kind K existing at some other time t^* . Moreover, we suspect that we can discover the truth or falsity of such statements by establishing whether certain facts obtain or not. Empirical investigation may provide us with evidence that the truth conditions of the identity statement are satisfied. That is to say, when asking questions about the identity over time of some kind of individual, we expect there to be conditions which are necessary and sufficient for the truth of identity statements about that kind of individuals. I suggest that the general form of the identity question can be expressed thus:

Identity Question What are the necessary and sufficient conditions for the truth of the claim that for Ks x and y , the x which is F at t is the same K as the y which is G at t^* ?

An instantiation of this formula may be this: Is the person who committed the unlawful action at some past time t the same person as the person who is on trial today? An answer to the identity question typically takes the form of a criterion stating necessary and sufficient conditions for the truth of an identity statement about some kind of being.

1.3 THE FORMAL CHARACTER OF IDENTITY

It is notable about my formulation of the identity question that it presents the identity question as a question about when a relation 'the same K as' holds between individual Ks picked out at different times. In order to clarify that

formulation it will be helpful first to give a more formal characterisation of the notion of identity in play in debates about identity over time and then explain how 'the same K as' relates to that notion.

Lewis writes: 'Everything is identical to itself. Nothing is identical to anything else. There is never any problem about what makes something identical to itself; nothing can fail to be. And there is never a problem about what makes two things identical: two things never can be identical' (Lewis 1997, 126-27).

The notion of identity characterised by Lewis is *numerical* identity and it is this notion which is in play in the personal identity debate. To begin with numerical identity must be distinguished from *qualitative* identity. Often people say things like, 'I have the same racket as you', or 'It was so embarrassing, she wore the same dress as I did.' Similarly we talk about 'identical twins' and so on. These locutions are about qualitative identity. When we say such things what we mean is not that two things are *one and the same*, but that they are qualitatively completely similar. It is in this sense that duplicates are the same, even though perfectly identical duplicates will be a plurality. On the other hand, if we say that your racket and my racket are one and the same, i.e. if we take the claim to be about numerical identity, then it means that there is only one racket, and that it is both yours and mine.

Formally numerical identity can be described as an equivalence relation (it's reflexive, symmetrical and transitive). Furthermore, identity entails *indiscernibility*: If x is identical to y , then x and y will have all 'their' properties in common (the locution 'they are identical' is simply a convenient way of expression). Schematically it can be put like this:

Indiscernibility of the Identical $(x) (y) (x = y \rightarrow (F) (Fx \leftrightarrow Fy))$ ²

As such the notion of identity is univocal. Whether we take the variables 'x' and 'y' to be persons or plants doesn't make a difference to what must be the case for it to be true to say of *x* that it is identical to *y*. For *x* and *y* to be one and the same they must share all their properties. The indiscernibility of the identical may appear to conflict with the possibility of change. In many cases we want to claim that *x* picked out at one time is identical with *y* picked out at another even though *x* at *t* lacks a property that *y* has at *t**. I come back to how one might deal with this problem later in the chapter.

What I think must be acknowledged is that on the one hand identity is a relation which we can say *a priori* holds between a thing and itself and never between a thing and something else. As pointed out by Lewis, there's no problem about this notion of identity. However, we also find ourselves engaged with questions about the identity of material objects of various kinds the answers to which seem to be established by empirical investigation into whether certain conditions hold. We discover identities. Problems about identity belong to the domain of empirically based identity judgments, typically with respect to material entities of various kinds. As I have pointed out, the problem which is the focus here is not epistemological and about what counts as good evidence for and how we can establish the truth of such identity judgments, but about what facts the obtaining of which the evidence is supposed to be evidence for.

Now in general it seems we have, for different kinds of objects, a relation, which we take to constitute the identity of objects of that kind, and with respect to objects of that kind this relation is equivalent to numerical identity.

² I deliberately name the principle this way, though it is usually called 'the indiscernibility of identicals.' It is supposed to be uncontroversial that identity doesn't hold between individuals, but between a thing and itself, so 'identicals' is a bit misleading.

The facts constitutive of the identity of my pipe with the pipe that belonged to my great grandfather will presumably be different from those constitutive of my plant's identity with the one I bought three years ago, and different from those constitutive of my identity with some past or future person. In this sense there are different identity criteria for different kinds of objects, personal identity, pipe identity, and plant identity.

However we should not take this to suggest that we mean something different by *identity* when using these expressions. When we say that 'the x which is F at t is the same K as the y which is G at t^* if and only if the individuals picked out at different times stand in some relation R , (where R will be different with respect to different kinds of objects), we are not giving necessary and sufficient conditions for different notions of identity. Rather we say something about what it takes to be of the kind in question. If something is of the pipe kind, it will have different persistence conditions than something which is of the person kind or the plant kind.

We may say, as is reflected in the expressions 'personal identity' and 'pipe identity', that for objects of kind K , there is a relation, *K-identity*, which is equivalent to identity with respect to K s. I.e. when K s picked out at different times stand in this relation, the formal requirements of the concept of identity are satisfied by 'them.' This is a relation which we might *discover* holds between individuals at different times.³ When the defendant is found guilty, part of the process leading to the verdict is to establish his identity with the culprit. Establishing the identity of the defendant and the culprit entails that the formal characteristics of the concept of identity will be satisfied by them. Thus we can confidently ascribe the property of breaking the law to the defendant at some past time, since he is identical with

³ Importantly, identity criteria, as I have pointed out, aren't epistemic, but they do state the facts the discovery of which allows us to conclude that individuals picked out at different times are one and the same. In this sense identity criteria are about what to discover when discovering identities. We don't discover *what* identity is, but *that* it is.

someone who did break the law at that time, and identity entails indiscernibility.

As should be clear from the foregoing there is an intimate connection between an individual's being of a certain kind and its having certain persistence conditions. Whereas the concept of numerical identity is 'kind-neutral' in that different kinds of objects aren't *identical* in different ways, identity criteria or persistence conditions vary according to the kind of object in question. Thus it seems as if being a K and having the persistence conditions described in a criterion of K-identity are two sides of the same metaphysical coin. In this sense I think it is fair to say that the question about identity, when this is understood in the sense relevant for discussion of personal identity, is really more a question about kind-membership.⁴

1.4 PERSISTENCE, CHANGE, AND INDISCERNIBILITY

Persons like other things may change their accidental and temporary properties over time.⁵ Last week I was ill, but now I'm well; the plant in my window is blooming, while last week it wasn't, and my old putter was straight, while now it is bent. Despite these changes in their properties, persons, plants, and putters continue to exist. Like other ordinary material objects they *undergo* changes in many of their properties. It is *the same plant* which is now blooming that was barren last week; it is the same person who is now well and who was ill last week, and it is the same old putter which is now bent that used to be straight. Things, we may plausibly contend, persist through change.

⁴ The idea that the question about identity is better thought of as being about kind-membership (since identity is the univocal notion formally characterised) goes back to e.g. Quine (1950). However, Quine and other perdurantists who accept that objects are spatiotemporally extended worms, are pretty liberal about what kinds of things there are and thus quite liberal about what diachronic relations make for the identity of some kind of thing over time. I come back to the distinction between perdurantism and endurantism later in this chapter.

⁵ I will primarily be concerned with changes in intrinsic properties, not relational changes such as my becoming an uncle.

The debate about identity over time is intimately connected with the truth conditions for tensed property ascriptions. I *was* ill, but currently I *am* well; my putter *was* straight, but *now* it is bent, etc. As became clear when contemplating the trial scenario in which it was established that the defendant was guilty due to his identity with the culprit, the truth of such tensed property ascriptions traces identity over time. In order for it to be true that I was ill, there must be some past person who was ill and with whom I am identical, and for my bent putter to have been straight, there must be a past straight putter with which it is identical. However, given that numerical identity entails indiscernibility, the intimate connection between the truth of tensed property ascriptions the truth of identity statements seem to generate a 'problem of change.'

When we say that things change we seem to claim that one and the same thing can have different properties, but this may at first sight appear to conflict with the indiscernibility principle. If the putter I am playing with today has different properties from the putter I played with yesterday; e.g. yesterday's putter was straight not bent and today's putter is bent not straight. So if identity entails indiscernibility it seems it can't be one and the same putter, which is being picked out at those distinct times. The possibility of change is appears hard to reconcile with the fact that identity entails indiscernibility.

One might argue that when we take ordinary physical objects to be identical over time, we use the 'is' of identity in a 'loose and popular sense' as opposed to 'the strict and philosophical sense.' A putters and persons and plants, like other physical objects, gain and loose parts and properties over time, and our ordinary discourse allows for such changes in the objects. The objects inhabiting our daily lives we don't take to be so fragile as not to survive changes in properties and parts. But strictly speaking, the changes

they *seem to* undergo are not changes in a single object, not changes that any single thing undergoes, but, strictly speaking, involve a succession of distinct and rather fragile entities. This sort of view is defended by Chisholm (1976) and goes back to e.g. Butler (1975, pp. 100-101) and Hume (1978, p. 255).

My aim here is not to assess this position. But what these considerations bring out is that if we take persons to be material individuals, along with putters and plants, there is a central question about how to account for the possibility of change over time whilst acknowledging the indiscernibility principle. How can we have 'identity in diversity' or 'persistence involving change' if our concept of identity entails indiscernibility? We seem to face a dilemma of either having to give up on the presumption that the notion of numerical identity applies to ordinary material objects, or reject the possibility that they may change over time. Neither alternative is attractive.

In order to deal with this point it will be helpful to introduce two general views about how things persist and see how the dilemma might be handled within the metaphysical frameworks they suggest.

1.5 PERDURANTISM AND ENDURANTISM

It is one thing to ask 'the special persistence question' about what it takes for a particular kind of object such as a person or a putter to persist. The answer to that question will vary according to the kind of object in question. But there is also 'a general persistence question', which concerns what it is for an object to persist, whatever kind it may belong to. In this section I will introduce two main views, *perdurantism* and *endurantism*.⁶ The views suggest different answers to the way we should conceive of persistence and change in general.

⁶ A third view on the market which has recently been defended is *stage theory* (e.g. Hawley 2001 and Sider 2001), but for present purposes I will focus on perdurantism and endurantism.

Let me note that the contemporary debate in metaphysics about the general persistence question is a highly intricate matter which contains much more detailed positions and arguments than I can take up here. Presently my focus is on how one might account for the possibility of change on the basis of one's view of persistence in general. I then consider how the two views will allow us to think about the identity question. It is not my aim here to decide the general question about how things persist. I will assume an endurantist framework in the chapters to come, and hence I'm not going to engage in much of a defence against arguments challenging it. Still I think contrasting endurantism with perdurantism and presenting how they attempt to deal with the problem of change sheds light on endurantism and how to understand the question about identity over time.

Perdurantism

So far I have formulated the issue of identity over time as being about the existence of one and the same thing at different times. Perdurantists when confronted with that question suggest, that for a thing, say a person, to exist at some time is for it to have a part at that time. Thus on what I shall refer to as the *perdurance* theory, an object exists at different times by having *temporal parts* at those times. On *endurance theory* an object is *wholly* present whenever it exists and thus such entities as temporal parts are rejected.

There are various reasons why one might be attracted to perdurance theory. The route to considering it here is that it might provide us with an account of how things persist, which allows for the possibility of change in a way that doesn't conflict with the indiscernibility principle. Yet another basis for the view, which also provides an explanatory analogy, is to think of existence over time as analogous to spatial extension. For an object to extend through space (or 'exist through space'), for it to exist in different places (at a time), it is natural to think that it must have parts in those places. E.g. a road extends

through space. Thus the A91 is both in Auchtermuchty and in St. Andrews and a natural explanation of how this can be so is that there's a part of the A91 in Auchtermuchty and another part in St. Andrews.

Perdurantists suggest that we can account for an object's existing at different *times* by invoking a notion of *temporal parts*. A temporal part of a putter I shall refer to as a temporal putter-part, and in general a temporal part of a *KI* I shall refer to as a temporal *K*-part.⁷ Call the putter 'Smokey'. On this analysis Smokey exists at different times by there being temporal putter-parts at those times. Thus on the perdurantist picture Smokey is a temporally extended thing, a 'spatiotemporal worm' or a 'fourdimensional object', its persistence being constituted by there being different temporal putter-parts at different times, which are related (putter-related) such as to compose a single spatiotemporally extended putter. Importantly only maximal 'putter-related' aggregates of temporal putter-parts compose putters.⁸

There are several questions which come to mind when presented with the perdurance theory. A first question which I think one might reasonably ask is about what happens to *the object*, e.g. the putter, the plant, or the person, if they are spread out in time? A first reaction is that the object itself, that which we refer to when we call the putter Smokey or the person Jackson, is never really present at any time. If I take a snapshot of Jackson putting on the 18th hole of the Old Course, I won't so to speak get a picture of the whole person, *Jackson*, but a picture of Jackson's temporal part present at that time. I get a snapshot of Jackson-at-t, where this is no more a picture of the whole of Jackson than is a picture of his left arm (and Jackson's putter and Jackson's

⁷ This is because I want to reserve the term '*K*-stage' for endurantist purposes.

⁸ The maximality constraint is described by Lewis (1976, p. 22): 'Something is a continuant person if and only if it is a maximal R-interrelated aggregate of person-stages. That is: if and only if it is an aggregate of person-stages, each of which is R-related to all the rest (and to itself), and it is a proper part of no other such aggregate.'

left arm are themselves temporal parts of a temporally extended putter and arm).

It is, I think, fair to say, that the perdurantist approach to persistence doesn't fit our ordinary conception of the way *things* persist. While events and processes such as football games, symphonies, and dinners are thought to be extended in time, this is not so with the things which are caught up in such events. The football and the football players are not typically thought to be event-like entities themselves, nor are the musicians and their instruments thought to extend in time either. An immediate worry of mine when faced with the perdurantist picture of persistence is that it seems to undermine genuine persistence as normally conceived, namely as the existence of one and the same object at different times. We start out with an interest in the fact that things picked out at different times may be discovered to be one and the same, but perdurance theory about persistence seems to entail that we can't pick out the same entity at different times, but merely distinct entities, temporal parts, which are not strictly speaking *the same thing*, but which are related in some way or another.

Facing this objection the perdurantist can point out that such a question illegitimately assumes an answer to what it is to persist. It can't be denied at the outset that persistence of objects involves temporal parts. That is what is up for consideration. When the perdurantist suggests that Leo the lion has a temporal lion-part today and a temporal lion-part yesterday and that this is what Leo's persistence comes to, we are provided with an account of how e.g. lions persist. It is not a way of saying that objects such as lions don't persist at all. There is a single lion which existed yesterday and exists today, just as there is a single road, the A 91, which is in Auchtermuchty and in St. Andrews by having a part in Auchtermuchty and a part in St. Andrews. We can't reject perdurantism for revising our view of what it is for something to persist, when that is the very question under consideration.

Perdurantism may seem attractive due to the way it may provide a framework for explaining what it is for an object to change from one time to another. When it comes to an object's changing over time, the perdurantist can invoke the notion of temporal parts. When some kind of individual, say Leo the lion, changes from napping to hunting, this change is explained by him having different temporal parts at those times. It is a complex issue just how perdurantists understand the appeal to temporal parts as an answer to the problem of change.

Take the road analogy. The A91 varies through space in that it is wet in one place and dry at another due to its Auchtermuchty-part being wet and its St. Andrews-part being dry. I.e. the A91 changes its properties through space in virtue of having different spatial parts with incompatible properties. The Auchtermuchty-part is wet and the St. Andrews-part is dry.

One way to understand the perdurantist's solution to the problem of change is to see it as exploiting the road analogy. The perdurantist may be understood as suggesting that, on the basis of this analogy, we can ascribe incompatible properties to a persisting thing at different times without violating the indiscernibility of the identical. We can appeal to the fact that when we say that a persisting object has incompatible properties at different times, what this means is that there are distinct entities, temporal parts, instantiating the properties which the persisting object (atemporally) has at those times. But if the incompatible properties are instantiated by different entities, then we don't get into a conflict with the indiscernibility principle. In this way perdurance theory may be said to provide a way of explaining the possibility of change over time by invoking a metaphysics of temporal parts.⁹

⁹ I think this description of perdurantism is justified by remarks by Lewis (1986, p. 204) and Sider (2001, p. 93) who writes: The temporal parts account of change is that incompatible

One might complain that this way of dealing with the problem of change is no answer at all. We are being told that there are distinct temporal parts which instantiate incompatible properties. But presumably there is also a single persisting entity, the spatiotemporally extended worm, which, in the case of Leo the lion, is napping at one time and hunting at another, and invoking some other entities, temporal parts, and their properties doesn't directly answer any questions about Leo the lion. We want to be told something about *the lion* or *Leo* and the way he changes over time. But on the explanation suggested in terms of temporal parts, we don't seem to get an account of what happens with the persisting thing, Leo the lion.

To this sort of objection I think the perdurantist can answer that we get what we want. The perdurantist has given us a theory of what it is for some object to change from one time to another, just as it has given us an account of what makes for a single object's persistence, and we beg the question against his theory, if we then come and say that that's not what we mean by change.

Another issue central to the perdurantist proposal has to do with ascribing properties to temporal parts such as the properties of hunting and napping. It is unlikely that a short temporal part of Leo can correctly be said to possess these properties. Those properties are presumably for the spatiotemporal worm to have in virtue of facts about temporal parts. And even in case there are temporal parts which are 'long enough' to be hunting or napping, a perdurantist will, I suspect, not be keen to allow them to possess them on pain of there being too many hunters and nappers around, namely both the 'long' temporal parts and the worm of which they are parts.

properties are had by different objects, different temporal parts of the whole.' I raise some problems which faces this way of thinking about perdurantism below.

What this points to is that a perdurantist must not only provide an account of the ontological status of temporal parts, but say something about what sort of properties they can have and about how long they last for. Some find the notion of temporal parts to be incomprehensible, others take the temporal parts metaphysics to be justified due to the machinery it provides for solving 'metaphysical puzzles' and its accordance with modern physics, views on time and so on.

The issues just raised with respect to perdurantism are intricate and an ongoing subject of debate which I will leave aside. My aim here is merely to provide a sketch of the general perdurantist approach to how things persist through change and contrast it with the endurantist approach, and I think the perdurance theory has been sufficiently characterised for that purposes.

To sum up: Perdurantism claims that an object is temporally extended and persists by having different temporal parts at different times; for such a temporally extended object to change over time is for there to be temporal parts of the object which differ in their properties. On this account it is not the numerically same entity which is both F and not F at different times, so indiscernibility isn't violated.

Endurantism

Scepticism about whether perdurantism really does provide an account of genuine persistence and change is perhaps not much of an objection to perdurantism, but it does bring out what may be attractive about another view of persistence often supposed to be 'the common sense view', viz. that objects persist by being 'wholly present' at each time at which they exist. This view is expressed by the *endurance* theory.

The slogan of endurantists is that the *whole* object exists at any time at which it exists. As Sider (2001, pp. 63-68) points out, it is not an easy task to provide

a clear statement of what is meant by 'an object's being wholly present at any time at which it exists.' Still, I think, a plausible common denominator of endurantists is that they reject appeal to the notion of temporal parts as a way of explaining what it is for objects to persist and change over time. Persistence on the endurantist view is thus, I suppose, pretty much what people tend to think it is. 'The man in the street' will presumably conceive of himself as an individual 'moving through time' a bit like a car moving along a road, and not think of himself and his existence over time along the lines of a spatiotemporally extended 'worm' with different parts at different times. It is not a part of Leo that is hunting in the morning and another part that is taking a nap in the afternoon. It's *Leo* who is hunting and napping, not different parts of him. Or, to put it in the snapshot terminology, taking a snapshot of Leo napping at one time and one of him hunting at another time, I take pictures of the very same thing, the lion.

Endurantists also have to find a way to deal with the problem of change over time. The problem is that if Smokey is bent, then everything identical with Smokey is also bent. But Smokey wasn't bent yesterday, so how can the putter we claim to be Smokey be the same putter as the straight putter we looked at yesterday? How can the napping lion be the same lion as the hunting lion, when it is not hunting and identity requires indiscernibility? It seems as if the endurantist who can't appeal to different temporal parts is in trouble.

To this objection the endurantist can reasonably point out that we use *tensed* language when ascribing properties to things. The lion *was* hunting and *is* napping. Smokey *was* cold and now it *is* hot. On this basis endurantists can maintain that it is one and the same wholly present object which is picked out at different times as *the napping lion* and *the hunting lion*.

Our tensed language is apt to ascribe temporary properties, but as formulated the principle of the indiscernibility of identicals doesn't seem to take this into account. Thus it may be expedient for the endurantist to ascribe properties atemporally to avoid misunderstanding (this is not an argument against using tensed verbs).

Take a situation which we would describe thus: The person teeing off in the morning is relaxed. We might interpret this as ascribing a temporary property to the person teeing off, Tiger Woods say, at a certain time, viz. the property of being relaxed. But we can also ascribe to him the permanent property *is relaxed-at-t*. Ascribing the property in this way doesn't conflict with saying that he is not *relaxed-at-t**. Thus relativizing properties to times is a way of avoiding an apparent violation of indiscernibility.

It has been a big issue in contemporary discussions of endurantism and perdurantism whether the endurantist way of accounting for how one and the same thing can have different properties at different times, is acceptable. Lewis (1986, pp. 202-204) has famously argued that temporary properties such as being bent and being straight are temporary *intrinsic* properties and not relational. Whether an object has a certain shape, Lewis insists, doesn't depend on its relation to anything, let alone to any time, and he adopts a perdurantist view. I will not discuss in detail whether purportedly intrinsic temporary properties such as being bent may instead be relations to times. The point here is to present a typical way in which endurantists deal with the problem of change. That problem is about how to account for the fact that one and the same object at different times may have different properties which seems to violate the indiscernibility of the identical, and endurantists

can either insist that they can avoid contradiction by using tensed verbs or by relativizing properties to times.¹⁰

To sum up: If we take persistence to be the existence at different times of the wholly present very same thing and think that identity entails indiscernibility, the possibility of change seems to force a dilemma upon us, in that we will accept that the very same thing both is and isn't, albeit at different times, e.g. relaxed and not relaxed.

One way to avoid this is to adopt a perdurantist view of persistence and change in general. But I have suggested that change need not be problematic for the endurantist. The apparent contradiction can be avoided by noting that we use tensed language when ascribing temporary properties or by relativizing properties to times.

I have now introduced two general views about persistence and change. Whatever we take to be the right account of how objects persist and change over time, there still remain issues to do with 'the special persistence question', which has to do with the specific identity criteria for various kinds of objects such as lions, persons and ships. It is after all by recognising that the facts constituting e.g. personal identity between the defendant and the culprit that we can conclude from the indiscernibility principle, that the defendant *was* e.g. stealing the car. I now turn to consider in more detail what a criterion of personal identity can be said to be within an endurantist framework.

1.6 PERSONAL IDENTITY AND PERSONAL HISTORIES

When we state a criterion of identity, we must be careful to do so in terms of properties and relations between the Ks in question, which can be grasped

¹⁰ There are other endurantist moves, e.g. adverbialism, but I think these two are the more attractive ones partly on the basis of discussion in Hawley 2001, pp. 21-24..

independently of awareness of the identity criteria for Ks for our effort to be non-trivial and informative.

The question about K-identity is this: What are the necessary and sufficient conditions for the truth of the claim that for Ks x and y , the x which is F at t is identical with the y which is G at t^* ? An answer will typically take this form:

For Ks x and y , the x which is F at t is identical with the y which is G at t^* if and only if the x which is F at t stands in relation R to the y which is G at t^* .

Sometimes the left-hand side of a criterion of personal identity is stated in a different way: 'Person P1 at t_1 is the same person as person P2 at t_2 if and only if ...'¹¹

If we are endurantists it is not quite clear how to think of the expressions 'person P1 at t_1 ' and 'person P2 at t_2 .' I briefly discuss some possible interpretations and thereby motivate the formulation I have used.¹²

On one interpretation, the referring expressions will be substituted with personal names such as 'Bob' and 'Bill.' We then get a statement of this form:

Bob at t is the same person as Bill at t^* if and only if ...

But given that the identity statement has proper names as referring expressions, the temporal modifiers seem out of place. First we can note that unlike e.g. definite descriptions, proper names (we assume) don't change their reference over time. So if Bob is Bill today, then this is the case at any time you ask. (Compare: 'The man in the window is the British Prime

¹¹ See e.g. Swinburne 1984.

¹² Here I am indebted to Olson 1997, pp. 37-41.

Minister.’ In this case we’d naturally ask for some temporal modifiers of the referring expressions to complete them.)

We could take the temporal modifier to be a temporal adverb qualifying the identity predicate, but this is not an easy option either, since identity (we assume) isn’t relative to times. If Bob is identical to Bill this is so whenever those names refer. If there was identity yesterday there’s identity today and *vice versa*.¹³

I agree with Olson who suggests that the better interpretation is one according to which the way to read the left hand side is to take the identity statement to be stating that two definite descriptions, such as ‘the general being awarded a medal at t^* ’ and ‘the boy stealing apples at t ’, have the same referent.¹⁴ The temporal qualifications are components of the noun-phrases and complete the definite descriptions by indicating when they are taken to pick out a person.

Now a criterion of personal identity is supposed to tell us what it takes for a person to persist and thus an account of personal identity is required not to be stated in terms the understanding of which presupposes knowledge of such conditions. The persistence conditions must be explained such that they could come to be understood independently of knowing whether the identity statement is true or not. But, furthermore, as knowledge of persistence conditions are part of grasping what a person is, these conditions must be such that someone without a full grasp of what it is to be a person can come to understand them. This brings us to how to interpret the right hand side of the biconditional stating that for Ks x and y , the x which is F at t is identical

¹³ In chapter 6 I will argue that we should adopt the temporary identity thesis when confronted with fission scenarios. On that view identity is not a relation which holds between individuals across time. However, in my presentation of the debate about personal identity I assume the necessity of identity in the discussion.

¹⁴ Olson (1997), p. 40.

with the y which is G at t^* if and only if the x which is F at t stands in relation R to the y which is G at t^* .

1.7 THE PERDURANTIST INTERPRETATION¹⁵

Here's a remark from Hirsch (1982, pp. 3-4): "When we ask with regard to physical objects what their identity through time consists in, we are asking for an account of the unity of a physical object's career. Any physical object has a career which stretches over a period of time, a career which we *can think of as* (italics added) comprised of a temporal succession of momentary stages.'

To think in terms of a succession of stages (or parts since I want to reserve 'stage' for endurantist purposes) of an object as a way of analysing its existence over time comes natural for a perdurantist.¹⁶ Call the relation which holds between person-parts which are parts of the same perduring person 'the I-relation' (this might e.g. be a relation of psychological continuity). On the perdurantist view, something is a person if and only if it is a *maximal* I-interrelated aggregate of person-parts. Thus (proper) temporal parts of temporally extended persons are not themselves persons. Nor are any two person-parts which are I-related parts of *the* same person. A single person-part can be part of distinct perduring persons, just as a stretch of road can be part of both A90 and A91, which are different roads. What is claimed is that if person-parts are I-related, then there is a perduring person of whom the one is a part who is identical to a perduring person of whom the other is a part (cf. Lewis 1976, pp. 22-23).

¹⁵ I take Lewis 1976 as my exemplary perdurantist.

¹⁶ A perdurantist doesn't *just think* in terms of person-stages, he is committed to their reality. An endurantist might be tempted to think in terms of stages, but will not accept an ontological commitment.

For a perdurantist both sides of the criterion should be interpreted in terms of K-parts. The definite descriptions on the left hand side of the biconditional should be taken to pick out person-parts at different times and state that there is a perduring person of whom they are both parts. Hence, strictly speaking, 'the Ks x and y ' are not entities which are claimed to be one and the same, but distinct entities about which it is claimed that there is a perduring K of which they are both parts. The right hand side states what relation must hold between different K-parts for this to be true.

Identity, on this view, holds between whole spatiotemporal worm(s) which are identical when they have all the same temporal parts; it is not the I-relation which is the identity relation, the I-relation holds between different parts of the same perduring object. I will assume that the perdurantist will take the left hand side really to be the claim that there is a person of which two person-parts are both parts.

I suppose one might be tempted to criticise this interpretation for rendering the original identity question highly misleading, since it is presented as a claim about identity, not as a statement claiming something about distinct parts being parts of the same thing. This complaint is, I think, similar to the complaint that perdurantism doesn't tell us how things persist and change. We want to know when individuals picked out at different times are one and the same, and instead we are told that the question is really about when different entities are parts of a single individual. However, I think the perdurantist can meet this criticism by pointing out that just as it give an account of how to understand claims about persistence and change, it gives an interpretation of how to understand claims about identity over time, and it will beg the question against it to reply that it doesn't give an account of

what we mean when we ask what it takes for an object to be identical over time.¹⁷

Now Lewis writes: 'But identity among continuant persons induces a relation among stages: the relation that holds between the several stages of a single continuant person. Call this the *I-relation*. It is the I-relation, not identity itself, we must compare with the R-relation' (1976, p. 21).

By the 'R-relation' Lewis has a particular psychological relation in mind (the relation that matters in survival), but I leave it open for now what I think is the best candidate for being the R-relation (this is the subject of chapters 2 and 3). The important thing to observe here is that for Lewis, the *right hand side* of the biconditional is supposed to tell us what relation holds between distinct person-parts of a single perduring (or continuant) person.

An identity criterion for persons on the perdurantist view may thus be stated like this:

The Perdurantist Interpretation There is a person of whom both a person-part at t and a person-part at t^* are a part if and only if the person-part at t is R-related to the person-part at t^* .

1.8 THE ENDURANTIST INTERPRETATION

Endurantists are not happy to use the notion of a temporal person-part. They want to take seriously that the expressions 'the x which is F at t ' and 'the y which is G at t^* ' refer to a person wholly present at those times and are identical across time. The left hand side is a claim about enduring persons. How may the right hand side best be interpreted?

¹⁷ I will not go into modal issues related to this claim such as the problem about entities which actually share all their temporal parts, but which possibly do not (see e.g. Gibbard 1975 for discussion of this issue).

The endurantist can't appeal to temporal parts but he can, I think, adopt a way of formulating identity criteria which, as Hirsch suggests, involves *thinking about* identity over time of an object as *the unity of stages* of the object, without committing themselves to the perdurantist ontology of temporal parts. Different stages are stages in the history or life of one and the same object if and only if they are related by what some call 'the unity relation' for objects of the kind in question. I think this is the best an endurantist can do when it comes to interpreting the right hand side. What he can do is to take a criterion of identity of enduring persons really to be a criterion for when person-events are related such as to be events in the temporally extended event we call a personal history. It tells us what events can go in our autobiography and which can't.¹⁸ For an endurantist the criteria of personal identity will best be stated in an 'eventful' way.

Taking our clue from Hirsch's suggestion and following Perry (1975, p. 8 ff.) we can start out by considering the concept of a football game and the notion of events being events in the same football game at a time and at different times. Take a situation in which you have a friend who doesn't have a clue as to what we mean by a football game. A way to introduce the concept of a football game to him is to take him to the local sports centre where plenty of football games are taking place on Sunday afternoon. Walking around with him, you make him aware of a number of events which are part of a football game. You make him able to identify that there's a corner kick, that a goal's being scored, and that a player is off side. Soon he has the ability to single out different 'football game events,' and he may even be assumed to be able to say *at a given time* whether a group of football events are events in the same football game. Thus on the playing field there might be several pitches with several games on, but we assume that our friend can tell whether a

¹⁸ Here I agree with Perry 1975, p. 10.

corner kick in one of the games and a simultaneous goal in another are events in a single game at that time.

However, you realise that though he manages to identify many individual football game events, say goals, he doesn't know whether when seeing someone score a goal it is a goal in the same football game as a corner kick he saw being taken earlier. He doesn't seem to have a grasp of how football game events relate *over time* for them to be parts of the same football game.

Now what is asked for is a relation which takes *sets of synchronically unified football events* as its values, call them *football game stages*, and relates them in such a way as to be stages in one and the same football game event. The relation is thus not supposed to be a relation which holds between football games, but between football game stages in the same football game. The question about the identity of a football game thus becomes what to take as relation R in this biconditional:

Football Game Identity Criterion The football game in which FS is a stage is the same football game as the football game in which FS* is a stage if and only if FS has relation R to FS*.

The stages FS and FS* are different stages in the same football game due to their being R-related. In this way the R-relation holds between different entities, not between an entity and itself. Hence the R-relation is not identity. It is better thought of as a *unity relation*.

We can now generalise these considerations about the identity of an event such as a football game, and we can distinguish between two subquestions about football game identity and K-identity (i.e. K-histories or K-lives) in general:

Synchronic Identity Question What relation obtains between *simultaneous* K-events that are events belonging to the same K?

Taking the notion of a K-stage to be of a set containing all and only those K-events which belong to a K at a given time, we can also ask with respect to Ks:

Diachronic Identity Question What relation obtains between K-stages when there is a K(-history) of which they are both stages?¹⁹

Consider now the question about personal identity. If we think about the question of personal identity in terms of the unity of a personal history or life, we can reformulate the question as being about what it takes for two person-history stages to be stages in the history of a person, in one and the same personal history. It seems reasonable to assume that the notion of a personal history can be appealed to as a way of explaining how to conceive of a criterion of identity for persons. If we have the notion of enduring things such as persons, we also have the notion of events, lives or histories in which enduring objects are 'caught up' or involved.

Take the case in which one has identified two person-involving events, the awarding of a medal and the stealing of apples. Instead of asking whether the general being awarded the medal today is the same person as the boy stealing the apples fifty years ago, we may ask whether the two person-involving events, 'the stealing apples fifty years ago' and 'being awarded a medal today,' are events in the same personal history. The criterion, the right hand side of the biconditional, is supposed to tell us how these two distinct events must be related for that to be the case.

¹⁹ Here I am following Perry 1975, p. 9. But I formulate the diachronic identity question differently. Perry's formulation uses the expression 'the same K.'

On the suggested interpretation the identity statement on the left hand side, 'the x which is F at t is the same person as the y which is G at t^* ', is interpreted in terms of distinct person-history stages at different times being parts of the same personal history, and the right hand side states that this is true if and only if the person-history stages are related in the way that constitutes the unity of a personal history, relation R . The endurantist can now move on to consider what that relation (relation R) might be.

Formulated in this way the endurantist adopts an interpretation which in many respects is like the one given by the perdurantist. But the endurantist doesn't thereby commit himself to the view that that *the person* doesn't wholly exist at different times. Nor does the endurantist end up saying that a person is a construction of or reducible to a series of events.

I submit that on an endurantist view, criteria of personal identity over time are best formulated in terms of person-history stages. When one knows how person-history stages must be related to be stages in the same personal history, one will be in a position to answer the question 'What are the necessary and sufficient conditions for a person picked out at one time to be the same person as a person picked out at some other time?'

1.9 BROAD AND NARROW QUESTIONS ABOUT PERSONAL IDENTITY

Eric Olson makes an interesting distinction between 'broad' and 'narrow' questions about personal identity. Olson suggests that we distinguish between two ways of asking about our persistence (Olson 1997, p. 25 ff.):

Broad Question What are the conditions under which something that is a person at one time is identical with *anything at all* that exists at another time?

Narrow Question What are the conditions under which something that is a person at one time is identical with something that is a *person* at another time?²⁰

By making this distinction Olson wants to point out that formulating the identity question in the narrow way isn't metaphysically innocent. The identity question is about numerical identity, and Olson's contention is that the narrow question assumes that for an individual picked out at time t^* to be numerically identical to something that is a person picked out at some other time t , the individual picked out at t^* must itself be a person. The expression ' x is the same person as y ' is supposed to mean that x is a person and y is a person and that they are numerically identical. But it is by no means obvious that in order to reidentify an individual person x with a future individual y that individual y must be at *person* at the future time. Maybe that which is a person now can be identical to something in the future which doesn't have the properties necessary for being a person. We can't assume that an individual of which it is now true to say that it is a person can't be identical to individuals at other times which are not persons at those times. But stating the identity question in the narrow way prejudices the issue in that it assumes that persons must persist *as persons* (1997, p. 24). The narrow question may be an appropriate statement of the problem if persons are essentially persons, but this is a controversial assumption. On some views such as animalism being a person is of no more relevance to an individual's persistence than being a student or the Prince of Denmark.

Now there is of course a common sense of 'same person' which is more like 'student' in that we say that 'he is not the same person as he used to be', where it is implicit that it is the numerically same individual which is not the same person as he used to be. And it may also be that in the same way, we

²⁰ Olson (1997), p. 25.

may sometimes say that an individual is the same person as some numerically distinct individual used to be, e.g. if someone develops the character traits and habits of, say, his father. My mother might say (perhaps in a somewhat disenchanted manner), 'you're the same person as your father,' without thereby taking us to be one and the same. But this doesn't show that we don't also have a notion of 'same person as' which expresses a numerical judgment. Olson and other animalists ultimately rejects that there is a sense in which 'personal identity' expresses a relation of numerical identity. To him nothing is no more essentially a person than it is a student or the Prince of Denmark.

Olson's concern that the narrow question is inappropriate is that it is unclear whether something that is a person at one time during its existence must be a person at any moment of its existence. The narrow question will only be appropriate if the things which are persons will only exist at other times *as persons*, which is to say, it will only be appropriate if the concept of a person 'tells us in a special sense, what the object is, and not merely what it does or where it is located or some other accidental feature of it' (see Olson 1997, p. 28). Concepts which do tell us *what an object is*, and here I take this to mean 'what an object *essentially* is' or 'what an object *persists as*', Olson calls 'substance concepts.'²¹ The appropriateness of the narrow question stands and falls with whether the concept of a person is a substance concept or not, i.e. it depends on whether Locke was successful in distinguishing an

²¹ Olson (1997, p. 28) adopts a Wigginsian theory of individuation and kinds and notes that he will rely on it while not arguing for it. To undertake an in depth investigation of Wiggins' theory of individuation is beyond the scope of this thesis. I am sympathetic to sortal essentialism in general, the view that for every object there is a sortal property, such that knowing that an object has that property is to know *what it is* and it's metaphysically impossible for the object to exist without having that property. While being sympathetic to the sortal essentialist approach I am also somewhat sceptical about whether a good account has been given of why some sortal properties are essential rather than accidental and which principles determine whether some sortal property is essential or not. (For an excellent discussion of these two questions see P. Mackie 1994.)

essential property when characterising a person as a rational, thinking, self-conscious being ('thinking intelligent being with reason and reflection').

Now Olson points out that *if* we accept that 'person' is a substance concept this entails not only that 'once a person always a person,' but also that all persons must have the same persistence conditions. I will come back to the discussion about whether persons persist *as such* and about whether personal identity is numerical identity or not, in chapter 8, where I discuss Olson's argument for rejecting that being a person is essential to something which is a person. However, I want to end this section with some critical remarks about Olson's distinction between the narrow and the broad question.

As I see it, the narrow question asks about what it takes for something x which is picked out as a person at one time to be one and the same thing as something picked out as a person y at some other time. But why take this question to *assume* that *being a person* is necessary for y to be identical with x ? In case the question is about the conditions for a student x to be identical with a student y at some other time, it doesn't seem that the question assumes that being a student is something that an individual must be at all times at which it exists. What I do assume in these cases is that there is *something* (that which it 'must persist as,' to use Olson's own expression, 1997, p. 24), which the student or the person x is and that a y at some other time must be *that kind of thing* for y to be identical with x , but I don't see why the narrow question as formulated prejudges an answer to that question about the essential nature of the individual x .

Rather it seems plausible to contend that our *answer* to the narrow question will give us an idea of what we take to be the essential nature of the individual x picked out as a person or a student. If it turns out that we find that the conditions for x 's identity with y are such that it is necessary that y be a person in order to be identical with x , it is, I submit, a good reason to

conclude that persons such as x will be persons at any time at which they exist. In case we contemplate the identity conditions for something picked out as a student at one time and find that for *it* to be identical with some future student it is necessary that there is *a student* to be found at every point of the spatiotemporal path between the two, it seems as if studenthood is a property, which we have reason to take seriously as being significant for the persistence of the objects that satisfy it. Now clearly we aren't tempted to think so about studenthood, but then think of the naturalness with which we would answer the question about the identity of a student over time with a phrase involving 'must be the same person as.' So I don't agree with Olson that the narrow question prejudices issues about whether personhood is an essential property or not and whether 'person' qualify as a substance concept.

These considerations also reveal why I am not quite happy about Olson's broad question. If we think that there is something which an individual x persists *as*, then it is odd to ask what it takes for it to be identical with *anything at all* at some other time. Rather the question should be, 'What are the conditions for person x to be identical to some y of x 's essential kind existing at some other time?' This leaves open what person x 's essential kind is, and forces us to consider what person x essentially is before trying to answer the question. In this formulation priority is given to considering *what x essentially is* before embarking on contemplating its persistence conditions.

Olson's central complaint when distinguishing the narrow and the broad question is that the narrow question takes something for granted about personhood, namely that it is an essential property. The very phrase 'personal identity' used to denote the question about what it takes for some future being to be numerically identical with you or me may be seen as a false start on behalf of those who take our identity to consist in psychological continuity of some kind, since it seems clear that if persons are essentially

persons, i.e. beings with mental properties of some complexity, then any nonpsychological account of what our persistence consists in is doomed from the word 'go' (cf. Olson 1997, p. 29-30). But we should hesitate to think that Locke singled out an essential kind when he defined 'person.' After all, there seem to be kinds of beings, e.g. human animals, some of which do and some of which don't satisfy the concept of a person. And plausibly human animals persist *as such* (i.e. as human animals) and not as persons. This is, I think, a very important point to consider for those who like myself are compelled to think that persons are essentially persons.

These points tie in with two general strands in the debate about personal identity. In general the debate is focused on the question, 'What does it take for there to be someone or something existing at some other time to be identical with you or me?' But the question may be approached from two different directions. One may, probably following a majority of personal identity theorists, focus on what it takes for us to persist. What sort of changes and circumstances can we survive, and what will result in our ceasing to exist? Another approach is to begin by asking, 'What kind of thing are we essentially?'²² For what it's worth, I find contemplating what we think it takes for us to persist is perhaps our best guide to answering the question about what we essentially are. In the next chapter I present my reasons for adopting a psychological criterion of our identity.

1.10 CONCLUSION

In this chapter I have presented questions about persistence, change and identity over time, and I have briefly characterised two general metaphysical frameworks within which these questions are typically approached. Assuming an endurantist framework which rejects appeal to temporal parts I

²² For a similar distinction between two aspects of the personal identity debate see e.g. Thomson 1997. I take Olson to think that there has been too much focus on the persistence angle and not enough reflection on how views about our persistence conditions square with plausible answers to the question about what we can plausibly be said to be essentially.

suggested that an endurantist will do best to formulate the personal identity question thus:

What are the necessary and sufficient conditions for there to be a personal history of which a person-history stage occurring at t and a person-history stage occurring at t^* are both parts?

And the answer should take the following form:

There is a personal history of which a person-history stage occurring at t and a person-history stage occurring at t^* are both parts if and only if, the person-history stage occurring at t and the person-history stage occurring at t^* are R-related.

If this is the case, then we can say that the person involved in the events belonging to the person-history stage occurring at t is the same person as the person involved in the events belonging to the person-history stage occurring at t^* . What to take as the R-relation is the subject of the chapters to come. I would like to note that what we take to be the R-relation will presumably not have to do with whether we accept perdurantism or endurantism.

2. The Psychological Criterion of Personal Identity

2.1 INTRODUCTION

In chapter 1 I gave an account of how to understand the problem of personal identity as a philosophical question. I now turn to consider what may be the correct answer to the question: What are the necessary and sufficient conditions for person P at time t to be the same person as person P* at some other time t*? (I shall use the standard terminology of 'person P at time t' in my discussion and refer to the considerations in Chapter 1 as to how this expression is best understood.)

I find the view that personal identity consists in psychological continuity and connectedness very plausible. In this chapter I present my reasons for finding a psychological criterion, as opposed to a broadly physical criterion, attractive. In my view a psychological criterion does much better accommodating some of our (or at least my) beliefs about our persistence. I conclude by stating what I take to be the correct view of what personal identity consists in. Let me note that the aim of this chapter is modest in that it is primarily meant as briefly presenting why I find the psychological criterion to be correct. I end by introducing some immediate worries about it. In later chapters I discuss what I take to be the most pressing objections to the psychological criterion.

2.2 IN FAVOUR OF THE PSYCHOLOGICAL CRITERION

For those who think that there is an informative non-circular criterion of personal identity to be given, the debate about personal identity has largely been focused on whether one should accept a physical criterion or a purely psychological criterion. I will take any criterion which holds that some kind of physical continuity is a necessary condition for sameness of person to be

physical, and use 'psychological criterion' to denote criteria which hold that only psychological relations are necessary for personal identity. Thus, someone who thinks that while psychological continuity is necessary it is also necessary that there is some sort of physical continuity for there to be personal identity will qualify as proposing a broadly physical criterion. The criterion of personal identity which I shall take as my paradigm example of a physical criterion is this:

Biological Criterion Person P1 at t1 is the same person as person P2 at t2 if and only if P1 at t1 is the same organism as P2 at t2.

Often proponents of a physical criterion has claimed that personal identity consists in sameness of body. Let me note here that there is an ongoing debate about the terms 'my body' and their use in the debate about personal identity. Van Inwagen (1980) and Olson (1997) both criticise the usefulness of the expression 'my body' which is common in the typical statement of the physical criterion (cf. Williams 1960, 1970a, and 1970b, and Thomson 1997). Olson I think presents a very good case for the claim that it simply isn't clear what sort of object is picked out by the term 'my body' (see Olson 1997, p. 142 ff.) For my purposes it will do to characterise the physical criterion in terms of the view that personal identity consists in sameness of human organism. After all, the philosophers I will focus on discuss how personal identity relates to the fact that we are human organisms and the biological approach to our persistence and, furthermore, I don't think that proponents of the physical criterion who prefer 'body talk' will for that reason have a stronger case for the physical criterion than those who take sameness of organism to be constitutive of personal identity. What is important here is that whether we accept sameness of body or sameness of organism as the criterion of personal identity, we accept a *nonpsychological* criterion, and that is the view I'm arguing against.

Another view which has been called 'the physical criterion' is one stating that we have sameness of person if we have sameness of 'enough of the brain' to continue the life of a person (Parfit 1987, p. 204, Unger 1990, p. 109.). However, it seems clear to me that this doesn't qualify as a nonpsychological criterion, since sameness of brain is not what the identity consists in on this view. The identity of a person consists in the relation which we take normally to be realised by the continued existence of enough of a functioning brain, but it is not the brain itself which makes for the identity, but the relations it realises.

One reason why a biological criterion may seem attractive is that an analysis of personal identity in terms of sameness of organism doesn't seem to face problems about circularity. We can surely understand the notion of sameness of organism without a grasp of sameness of person. Presumably we have the same organism at different times if and only if there is *spatiotemporal continuity* under the concept *organism* where this involves some sort of causal relationship between the organic states we can pick out at any time along the spatiotemporal path. But the fact that circularity is avoided if personal identity is analysed in physical terms is perhaps best thought of as a valuable bonus. What appears to be compelling for those accepting a broadly physical account is that it's a natural extension of the view that persons are enduring material beings, plausibly human organisms (or bodies).²³ They think that when we are ensnared by the psychological criterion having contemplated various exotic cases this is partly because we forget to keep in mind that we are enduring material beings, plausibly human organisms.

One thing that both adherents of the physical criterion and of the psychological criterion assume is that *if* the ontological claim that persons are

²³ Thomson writes: "The simplest view of what people are is that they are their bodies. (...) Certainly we all, at least at times, feel inclined to think that we are not merely embodied, but that we just, all simply, are our bodies." She continues: "What stands in the way of adopting this simple and attractive view?" Thomson (1997), p. 202.

identical to enduring material beings is true, then it follows that they have nonpsychological persistence conditions. What has stood in the way of adopting the view that persons are e.g. human organisms is a line of argument which supports a non-physical and purely psychological criterion of their identity. It does indeed sound plausible that if persons are material beings, then they'll have physical identity criteria, presumably spatiotemporal continuity under a concept. I will however deny this entailment in the course of this thesis.²⁴

Recently the biological criterion has been supported by animalists who argue that the fact that are identical to human animals entails that we should accept the biological criterion (prominent animalists include e.g. Olson 1997; Snowdon 1990; Carter 1990; and van Inwagen 1990, 1997).²⁵ The biological criterion was this:

Biological Criterion Person P1 at t1 is the same person as person P2 at t2 if and only if P1 at t1 is the same organism as P2 at t2.

A first criticism of this criterion may be this: Clearly it is not sufficient for a person to exist at different times that the same body exists at those times. There is an abundance of cases where persons die in their sleep leaving behind a dead organism. If I die in my sleep, I will not exist anymore. But if I die in my sleep, there will still be a spatiotemporally continuous material

²⁴ I consider two ways of doing so under the headings of 'personalism' and 'temporary identity' in chapters 7 and 9 respectively.

²⁵ I will assume for the time being that a living organism can survive as a dead organism, i.e. may survive the cessation of its biological functions. Now while there is agreement that personal identity consists in organism identity, there is disagreement amongst animalists about when an organism ceases to exist. Some think that an organism such as a tree or a cat can become a dead tree or a dead cat (e.g. Carter 1999). Others (Olson 1997, chapter 6) argue that an organism ceases to exist when its biological life is no longer continued. However, this is not of my main concern, since it doesn't directly make a difference as to whether we should accept the physical criterion. Indirectly though it might rule out certain arguments against the psychological criterion (I discuss this in chapter 8). For a recent discussion of 'the corpse problem' see e.g. Olson (2004) and Hershenov (2005).

being, a corpse, lying in my bed the next morning. Hence, the continued existence of the same organism is *not sufficient* for the continued existence of a person.²⁶

A second argument for rejecting the physical criterion is due to considering the following (at least conceptually) possible situation involving brain transplant:

Cerebrum Transplant

We may imagine that the person Brown is subjected to the following procedure: The cerebrum is removed from his head and then hooked up in the head of another organism which used to house the cerebrum of Robinson. The functioning organism, including the lower brain, is destroyed. The result of the cerebrum transplant is a person, call him Brownson, who we might say is composed of Brown's cerebrum and the Robinson's organism. Since Brown's cerebrum directly supports a continuous mental life Brownson is psychologically related to Brown. E.g. he seems to remember Brown's past life from 'the inside.' Brownson is however also continuous with Robinson in that Brownson's organism appears to be the same organism as the organism of Robinson. But, importantly, Brownson doesn't have any psychological states related to those of Robinson.

Presented with this scenario when considering our persistence conditions we have, I contend, a strong inclination to respond that:

(a) Brownson is the same person as Brown.

²⁶ Interestingly this argument sounds less plausible with respect to organisms than 'bodies', but if we assume that organisms become corpses thanks to the spatiotemporal continuity between them the argument may stand its ground.

Call this response the Transplant Intuition. Now it also seems plausible to respond that:

(b) Brownson's organism is not the same organism as Brown's organism.²⁷

But from these two claims it can be concluded that sameness of organism is *not necessary* for sameness of person. Contemplating our responses to Cerebrum Transplant we reach the conclusion that sameness of person and sameness of organism may come apart. Hence from the recognition that the organism might survive the person and from the Transplant Intuition we have two arguments jointly supporting the view that sameness of organism is neither a necessary nor a sufficient condition for the identity of a person. It is at best good, but fallible evidence.

Some may be tempted to rejoin that in Cerebrum Transplant the belief that Brownson is Brown is guided by the fact that there is some physical sameness between them. After all they have the same cerebrum.

But suppose that in order for the cerebrum to be accepted by Robinson's organism it has to be manipulated so as to realise an arbitrary psychological profile completely unrelated to the psychological profile of Brown. In that case we'd not be inclined to say that Brownson is the same person as Brown. It seems that what makes the cerebrum important in Cerebrum Transplant is that it realises a person's psychological profile.²⁸ Had it been some other part of Brown which played that role, the organ we call the upper brain or cerebrum wouldn't be of the same relevance for the decision. Rather what our response to Cerebrum Transplant seems to reveal is that we take

²⁷ Of course this doesn't go against the brain criterion, but there is another procedure which seems to do so. I describe this as a brain-state transfer below.

²⁸ By 'psychological profile' I mean the complete (or almost complete) set of a person's psychological states at a given time. I shall also call it 'the complete psychological state.'

continuity of psychological life to be sufficient for the persistence of a person, and that it isn't necessary for there to be sameness of organism.

There is another scenario which seems to reveal this belief about ourselves and which rules out the temptation to think that it is the spatiotemporal continuity of some physical entity or other which ensures our persistence. We may imagine a continuation of Brown's psychology without transplanting his cerebrum (or any other organ or part for that matter) and instead describe the scenario in a Williamsesque manner (See Williams 1970a):

Brain-State Transfer (BST)

Williams puts forward the idea of a machine which can erase and transfer the complete psychological state of a person into a 'blank' brain. The procedure we might imagine to involve Brown and Robinson. In this case there is no cerebrum transfer, but a continuous taping and erasing procedure followed by a recording procedure. (It is important that the taping and erasing takes place analogously in order to avoid complications which might arise if we have, say, Brown's complete psychological profile both realised by a cerebrum and taped at the same time.)²⁹

The outcome of this case with regard to Brown is similar to the transplant case, except that the causal mechanism underlying the relation of Brown's psychology would be abnormal. In the BST the psychological continuity traces a different path than the usual one involving the cerebrum. But there is still counterfactual dependence of Brownson's psychological profile on that of Brown. Brownson's memory of his 10th birthday originates in the experiences and memories of Brown. Some find that we should not take such abnormal causal paths to be sufficient for the psychological continuity

²⁹ I will briefly discuss the methodology of thought experiments at the end of this chapter.

constituting personal identity. I find that we should be liberal about what to take to be the appropriate causal ground for psychological continuity and thus I shall exploit the possibility of BST in subsequent discussions.³⁰

What these responses seem to point to is the claim that a psychological relation and not spatiotemporal physical continuity is necessary and sufficient for personal identity. Considering some albeit extraordinary cases has provided us with an argument against the view that personal identity consists in sameness of some physical object such as an organism and shown that we take it to be a matter of psychological relatedness. The person picked out as 'Brown' before the transplant/BST is the same person as 'Brownson' after the transplant/BST due to the psychological relatedness.

2.3 PSYCHOLOGICAL CONTINUITY

I find that the considerations presented so far (which form the standard background for accepting a psychological criterion) provide a good case for the claim that the persistence of a person necessarily involves psychological continuity.³¹ I have furthermore suggested that there is no principled reason why the cause of such continuity must be the normal cause. We should be liberal in this respect. However, when it is claimed that psychological continuity is necessary for the persistence of persons more has to be said about the content of this relation. What is psychological continuity supposed to be?

Traditionally advocates of a psychological criterion take *memory connections* to be a central element for sameness of person, but also other mental states

³⁰ E.g. Unger (1990) finds it important that there is physical continuity underlying psychological continuity. Garrett (1998) also seem to make a case for there being 'a right cause.'

³¹ I am not claiming that I have presented *the argument* for the psychological criterion. Many issues still have to be dealt with. However, I think it is fair to say that the attraction of a psychological criterion as opposed to a physical criterion is largely due to the considerations presented,

such as the holding of a belief or retaining a previously formed intention to perform an action will be elements in the psychological relation which constitutes personal identity. For simplicity I shall focus on memory connections, but presumably similar considerations will apply to other psychological relations.³²

So a typical suggestion is to say that a person at one time is the same person as a person at some other time if and only if there are memory connections between them, i.e. if the one can remember 'from the inside' experiences had by the other.

There are obvious problems about this suggestion. One is that clearly there will be plenty of events in the past which are part of my history and that I am no longer able to remember. If an old general does not remember stealing apples as a child this does not free him from being someone who did that. It is not as if he ceases to be the boy who was stealing apples just because he now doesn't remember doing it. If the psychological criterion involves commitment to such a claim this will surely count against it.

A popular way to avoid this sort of problem is to distinguish the notion of *psychological continuity* from *psychological connectedness*. Psychological continuity is a relation which is weaker than connectedness in that it may hold between persons at different times even if there are no direct

³² Unger (1990) distinguishes between 'distinctive psychology' and 'core psychology.' He argues that all that can plausibly be required for personal identity is the continuation of core psychology which comprises general and basic mental faculties, which we share with other persons. He criticises traditional psychological approaches as too demanding with respect to the distinct content of psychological continuity, while they are being too liberal about the appropriate cause of psychological continuity. I think Unger's distinction is an important one to be aware of, but I shall not enter into a detailed discussion of how the inherited distinctive psychological states must be for a person to have a psychological continuer. However I am inclined to disagree with Unger that only a very limited if any distinctive psychology is necessary. After all, the psychological profile is supposed to provide the basis for reidentifying a person as some person in the past, and this seem futile if all there is to go on is core psychology shared by every person.

psychological connections. All that is needed for psychological continuity is that there is a chain of overlapping direct connections. Thus the old general may be said to be psychologically continuous with the boy stealing apples even if he doesn't directly remember that event. In case he remembers being on boot camp as a young man, and the young man at that time remembered stealing apples as a boy, there will be psychological continuity between the general and the boy. So on the psychological criterion, when stated in terms of the concept of psychological continuity, the general will be the same person as the boy. (There may be questions about the degree to which there must be direct and indirect connections for this relation to constitute personal identity, but I shall leave such matters aside here. Parfit (1987) suggests how this might be worked out in more detail. Lewis (1976) also has some reflections on this when discussing the problem of longevity in relation to a perdurantist approach).³³

2.4 CIRCULARITY, PSYCHOLOGICAL CONTINUITY, AND QUASI-MEMORY

Another adjustment that proponents of the psychological criterion have had to make is to think of the elements of the continuity relation as e.g. *quasi-memory* in order to avoid the criterion becoming viciously circular. On our ordinary concept of memory one can only genuinely remember experiences had by oneself in the past. Our ordinary concept of memory implies that one doesn't remember what someone else did or experienced in the past, where remembering is 'remembering from the inside.' If the psychological criterion analyses personal identity in terms of a relation the holding of which can only be acknowledged to obtain if one already possesses a criterion of identity for persons, then the analysis will be unsatisfactory as an explanation of what personal identity consists in. It won't work to analyse personal identity in terms of a relation which we can only distinguish from 'seeming to remember' if we possess a criterion of personal identity.

³³ For more discussion of Reid's 'Brave Officer Paradox' see, Grice (1941), Quinton (1962) and Perry (1975).

Proponents of the psychological criterion typically appeal to the notion of *quasi-memory* to avoid this objection (cf. e.g. Shoemaker 1970, 1984; Parfit 1987). I devote a chapter to discussion of the coherence of this notion. Presently I will merely provide a brief characterization of the notion of quasi-memory.³⁴

A quasi-memory is a mental state which is phenomenologically just like an ordinary memory presenting a past event from the inside. However, unlike ordinary memories, quasi-memories don't require identity between rememberer and original subject. Nevertheless a quasi-memory must not only *seem to be* or *present itself as* an ordinary memory, but also be *caused* by *someone's* experience in the past of the event which seems to be remembered. Thus there are two features of quasi-memories which they have in common with ordinary memories: They share their phenomenology, and they must occur as the result of a causal path from a past experience.

What, on this account, distinguishes quasi-memory from ordinary memory is that mere quasi-memories are not subject to what Shoemaker calls the 'strong previous awareness requirement' which is that the subject of the quasi-memory is identical to the subject of the experience remembered. Characterised this way, ordinary memories become a subset of quasi-memories satisfying a further constraint, namely that there be identity between quasi-rememberer and original experience subject.

Assuming the coherence of the notion of quasi-memory the psychological continuity relation arguably constituting personal identity can be recast in terms of a relation of psychological *quasi-continuity*.³⁵ We are now in a

³⁴ The notion of quasi-memory was introduced in Shoemaker (1970).

³⁵ 'Quasi-continuity' is not to be understood as a less genuine relation of psychological continuity. The 'quasi'-prefix is merely intended to indicate that the continuity relation is to be understood in terms of quasi-memory, quasi-intention, etc.

position to state a preliminary version of the psychological criterion of personal identity:

Psychological Criterion Person P1 at t_1 is identical to person P2 at t_2 if and only if (1) P1 at t_1 is psychologically quasi-continuous with P2 at t_2 ; and (2) the psychological quasi-continuity is appropriately caused.

2.5 THE FISSION PROBLEM

I have suggested that we should be liberal with respect to the process underlying psychological continuity. Not only will the procedure of a brain-transplant be sufficient, but also the more exotic possibility of a brain-state transfer will do. Consider the BST. If we allow for this procedure to be sufficient for personal identity we face the following objection: In case Brown's brain-state is read off and our fantastic machine reads his psychological profile into the brains of two distinct bodies, we end up with two persons, call them Lefty and Righty, both of whom are related to Brown in the way which according to the psychological criterion guarantees identity with Brown. But now we seem to have a situation in which what appears to be two distinct persons are identical to Brown.

But as was made clear in chapter 1 identity is a relation which respects transitivity. What the proponent of the psychological criterion seems committed to is the problematic claim that Lefty is identical to Righty. This is a consequence of the truth of the claims that Lefty is identical to Brown and that Righty is identical to Brown, together with the transitivity and symmetry of identity. But this contradicts the fact that Lefty is not identical to Righty.³⁶

³⁶ A similar situation can be described in terms of half-brain transplants. Assuming that Brown may survive having merely the left (right) half of his brain (the left/right hemisphere) transplanted into Robinson's body and the other destroyed we can imagine that both of them are transplanted into distinct brainless bodies. The names 'Lefty' and 'Righty' I use due to the fact that this is a common way of referring to these two resulting persons.

The fact that the psychological criterion analyses personal identity in terms of a relation, which may in this way hold in a one-many form, has been taken to be a strong objection to accepting it. How can it be accepted that a relation such as identity, which is logically a one-one relation respecting transitivity and symmetry, is analysed in terms of a relation which does not do so? It seems incumbent upon the proponent of the psychological criterion to come up with a description of the result of fission which avoids the threat of its being committed to the set of inconsistent sentences. The more popular way to avoid this objection is to add a 'non-branching' clause to the psychological criterion. I discuss this and other moves which one might make in chapter 5.

2.6 THE NATURE AND PERSISTENCE OF PERSONS

Traditionally the question about personal identity has been focused on what changes persons can survive or undergo as opposed to those which result in them ceasing to exist. It is the fact that many find that persons have psychological persistence conditions which, presumably, has counted against identifying persons with human organisms. Persistence conditions tell us something about the nature of the kind of thing in question, and having certain persistence conditions might be taken to rule out being of a certain metaphysical nature. Recently animalists have pointed out that the psychological criterion has had too easy a time in that it hasn't been considered in appropriate detail what kind of thing could possibly be said to have the persistence conditions of the psychological criterion if we accept materialism and endurantism as our metaphysical framework³⁷.

I will come back to whether the psychological criterion can plausibly be maintained together with the claim that persons are enduring material beings in chapter 8. What I want to note here is just the general claim that the

³⁷ See e.g. van Inwagen (1997), Thomson (1997), and Olson (1997).

metaphysical nature of a thing, its belonging essentially to some kind, is taken to determine what it takes for it to persist.

Now I don't want to reject that it is plausible to claim that it is in virtue of being of a certain kind that a thing has the persistence conditions it has. However, I do think that one of our best ways to determine what kind something belongs to is by considering what we think it can survive. This means that if it turns out that we take ourselves to have certain persistence conditions, perhaps conditions which don't involve biological properties as essential properties, then we should take this seriously and accept that despite the fact that we are enduring material beings which have biological properties, we belong to a kind whose nature isn't essentially biological. That is to say, singling out persistence conditions is one of our best ways of coming to know what kind of object something is. The question about personal identity is about what changes beings like us can survive. If we come to have good reason to accept a certain criterion of e.g. our own identity, then this provides us with very good reason to acknowledge that we belong to a kind whose members have the persistence conditions singled out. Thus the 'metaphysical priority' of kind-membership to identity criteria doesn't rule out that 'the methodological priority' of identity criteria should make us recognise a substantial kind to go with it.³⁸ We can conclude something about our nature from considerations about our persistence conditions.

2.7 IMAGINARY CASES

In the argument for the psychological criterion I have appealed to imaginary cases such as Williams's brain-state transfer procedure. The debate about personal identity is notorious for its use of imaginary cases such as the brain-state transfer procedure just described. When considering the nature and

³⁸ I take the distinction between being metaphysically and methodologically prior from Garrett 1998, p. 12.

persistence of some kind of object it comes quite naturally to ask, what would happen to it in extraordinary situations which one might imagine. While it might be a reasonable supposition that whether an object can survive certain adventures depends on what kind of object it essentially is, a way to become more clear about its essential properties is to imagine that something happened to it and consider whether the imagined event is a possible future for the object in question.

The case of personal identity invites us to consider imagined cases since under ordinary circumstances the satisfaction of psychological and physical criteria go together. I am both physically and psychologically continuous with the person who moved to St. Andrews some years ago, but this identity judgment will come out true on both psychological and physical criteria. If we want to find out which criterion is the better account of what our identity consists in it is helpful to consider cases where they come apart.

Now it might turn out that we find it hard to come up with a definite answer to what happens to a person given some procedure. Some might find that what seems to be the same case, when described in different ways, provokes different replies and take this to be a symptom that intuitions generated by contemplating imagined cases are unreliable as a way of forming philosophical views e.g. about the nature and persistence of persons.³⁹

I agree that we must be careful when using imagined cases as 'intuition pumps' to be aware of the danger of letting prejudice shape our description of them. However, I don't think that the fact that a method must be applied with a certain vigilance counts against it as such. We can imagine plenty of scenarios, which we simply wouldn't know how to describe given the meaning invested in our concepts, but this doesn't mean that there aren't

³⁹ Johnston 1987, p. 67 criticises 'the method of cases' for not generating reliable intuitions. See also Williams 1970 for a similar worry.

imagined cases the consideration of which may bring about certain insights into what we take personal persistence to involve.

The question about the use of imaginary cases does I think deserve more attention than it is traditionally given in texts on personal identity and more than I can give it here. However, my main aim in this chapter is to present what I take to be the basis for accepting a psychological criterion, and while some may argue against the psychological criterion on the basis of a criticism of the use of imaginary cases in its support, this is not the line of criticism I am going to focus on.

Now imagined cases such as BST are presumably physically and conceptually possible. However there are important questions about how imaginability or conceivability and metaphysical possibility relate. Can we maintain that conceivability entails metaphysical possibility? Now it may not be clear that imagining a possibility guarantees that it is a metaphysical possibility. However, I will assume that the cases I have presented are metaphysically possible.

Now it might turn out that we come to know *a posteriori* that the psychological continuity relations, which I suggest constitute personal identity, necessarily involve some physical continuity, e.g. the continuity of the brain. Perhaps the relevant relation of psychological continuity and connectedness can only be realised by physical relations very similar to those which actually do it. This will of course put some pressure on the view that what constitutes personal identity is purely psychological. However, even if metaphysical necessities may be justified *a posteriori* and it turns out that there could only be a transfer of psychology if certain physical conditions were satisfied, I think it can still be maintained that it is not the satisfaction of the physical conditions which are constitutive of personal identity. The importance of the physical circumstances, insofar as the psychological

relations are concerned, doesn't undermine the conceptual conclusion that personal identity is an entirely psychological matter. The importance of the physical conditions being satisfied is entirely due to their role in maintaining psychological relations. Thus I can maintain the conclusions about personal identity I draw from the BST procedure even if that procedure doesn't turn out to be a genuine metaphysical possibility.

2.8 CONCLUSION

In this chapter I have lined out the line of reasoning which grounds my support for the psychological criterion. I have also presented some of the main objections and challenges facing proponents of the psychological criterion, challenges which I will take up in chapters to come.

In the next chapter I consider an objection which has been brought forward by Olson 1997. Olson argues that we need not take the Transplant Intuition to provide a justification for accepting that psychological continuity is sufficient for personal identity. Our reasons for thinking that Brownson is Brown in Brain-Transplant are based on broadly practical concerns which do not support a claim about the numerical identity of a person over time.

3. Personal and Practical Identity

3.1 INTRODUCTION

My main reason for finding the psychological criterion attractive is what I call 'the Transplant Intuition': In cases such as BST and Transplant I find that Brownson is the same person as Brown and not the same person as Robinson. I am, as Olson says, 'in the grip' of the Transplant Intuition. To me the fact that a future person is psychologically continuous with me is both necessary and sufficient for me to exist in the future.⁴⁰ In the chapters to come I consider more specific problems for this view such as the problems of fission and circularity. However, in this chapter I will discuss Olson's argument that the Transplant Intuition is not a reliable guide to our beliefs about our persistence.

Olson argues that our inclination to identify Brownson with Brown doesn't lead 'by simple and elegant arguments to an acceptable general principle about persistence' (Olson 1997, pp. 51-52). But, as Olson is aware, if we are 'in the grip' of the Transplant Intuition we may be reluctant to recognise that this is so. Even if there appear to be problems facing the psychological criterion, e.g. how to respond to the possibility of fission, these problems are unlikely to make someone give up on that criterion if in the grip of the Transplant Intuition.

Olson's contention is that we find the Transplant Intuition compelling because it explains certain responses about the practical concerns that someone like Brown would prudentially have for someone like Brownson. What our responses suggest is a view about whether Brownson falls within the scope of prudential concern from the point of view of Brown. However, if it turns out that one may be concerned for someone else in the future in just

⁴⁰ Note that I don't qualify psychological continuity with a non-branching requirement here. I will argue that it is not satisfactory to add a non-branching clause in order to handle the possibility of fission. I come back to this in chapters 5 and 6.

the way one is, under normal circumstances, concerned for oneself (what might be called 'quasi-prudential' concern analogously to 'quasi-memory'), then such practical concern doesn't justify claims about numerical identity, which is what the psychological criterion is purportedly all about. If the relation that grounds prudential concern is not *identity*, but some other relation, which need not coincide with identity, then the fact that we take Brownson to fall within the scope of Brown's prudential concern may not guide us when it comes to the question about personal identity.

What Olson argues is that the Transplant Intuition is based on who we'd take it to be correct for Brown to be prudentially concerned about when contemplating an upcoming BST or Transplant. Thus it is better thought of as an intuition about 'practical identity,' i.e. about who falls within the scope of prudential concern, not numerical identity.

Olson suggests that the reasoning lending support to the psychological approach on the basis of the Transplant Intuition is this (cf. Olson 1997, p. 69-70):

- (1) In Transplant, Brownson is practically the same person as Brown, and
- (2) x is practically the same person as y if and only if x is numerically identical to y .
- (3) Hence, Brownson is identical to Brown.

Olson's 'bold conjecture' is that when we take Brownson to be Brown the 'is' is an 'is of practical identity.' But if that is so, then the Transplant Intuition only supports a criterion of personal identity of metaphysical concern if premise (2) is true. However, Parfit has notoriously argued that personal identity is only practically important in a derivative sense. It is not necessary for someone to fall within the scope of my prudential concern that that person is strictly speaking me. The relation which grounds prudential

concern is not personal identity, and our ordinary notion of prudential concern as being about *oneself* is really just a case of 'quasi-prudential' concern, namely what we have when we are quasi-prudentially concerned about someone who also happens to be oneself.⁴¹ What is of nonderivative practical significance is whether he is psychologically continuous with me in the right way. If Parfit is right, Olson thinks he can reject premise (2), and if Olson can deny premise (2), then the conclusion expressed in the Transplant Intuition doesn't follow.

I think it is fair to say that many are attracted to the psychological criterion when contemplating cases such as Transplant. Olson's criticism is that they draw the wrong conclusion when they think their responses reveal something about what it takes for them to persist. What their responses do tell them about is their patterns of concern from a prudential point of view. However, as Olson himself is aware, we might not accept the Transplant Intuition on the basis of such reasoning. As an autobiographical fact I find that intuition to be compelling independently of reasoning involving reference to rational practical concerns. However, arguing from what we take to be a case of genuine prudential concern to conclusions about personal identity is an influential line of argument, so it is worth considering whether Olson has a case for getting us out of 'the grip.'

3.2 PERSISTENCE AND PRACTICAL CONCERN

It is commonplace in the debate about personal identity to invoke considerations about 'strictly prudential concern' when discussing whether some future person will be identical to some present person. And there seems to be a strong basis for favouring psychological criteria when bringing patterns of practical concern on to the scene. Unger explicitly claims that an inquiry into our persistence conditions must 'appreciate what is involved in

⁴¹ I should note that Parfit, to my knowledge, doesn't himself use the expression 'quasi-prudential' concern.

a philosophically adequate conception of ourselves. (...) such an adequate concept must be well suited for engagement with our central prudential thoughts and concerns' (Unger 2000, p. 326). Agreeing with Unger many advocates of a psychological criterion have appealed to what he calls 'Avoidance of Future Great Pain Tests' (or imagined cases of a similar nature).

Avoidance of Future Great Pain Test (AFGPT)⁴²

Before undergoing some procedure you are asked to consider whether from a strictly prudential concern you will choose to have yourself suffer considerable pain before the procedure if your not suffering this will result in the infliction of far greater pain on the being resulting from the procedure. On the basis of our answer to this kind of scenario it will presumably be possible to reveal some of our deepest beliefs about our persistence conditions.

In case the procedure is eating tomato soup or having a heart transplant I take it that everyone from a strictly prudential concern would take the considerable pain before the procedure so as to prevent the person digesting the tomato soup or receiving a new heart from suffering far greater pain. Similarly, if you were Brown and you were told about the Transplant procedure or BST, the claim is that given our ordinary practices it will be required of you to be concerned about Brownson from a strictly prudential point of view, whereas the being left behind when Brown's brain is removed doesn't seem to fall within the scope of such concern.

If it is objected that prudential concern is by definition rational concern for *oneself*, then the argument is better thought of as stating that Brown would have a deep felt rational concern for the welfare of Brownson *just like* the

⁴² See Unger (1990) for an elaborate development of this 'test'.

concern he has for himself. The claim is that the reasons Brown has for being specially concerned about himself (assuming there are such reasons), he also has for being concerned about Brownson. But if it is rational for Brown to be concerned about Brownson the way he is usually concerned about himself when the processes are more down to earth such as eating tomato soup or (less ordinary I suppose) having a heart transplant, then we have justification for thinking that Brownson is Brown. When some future person is related to a present person in such a way that the future person falls within the scope of prudential concern, we may say that they are *practically identical*. In Transplant there is general agreement that Brownson is practically identical to Brown, but should we also take this to entail that they are numerically identical? Is practical identity a sufficient condition for numerical identity? Parfit has notoriously argued that personal identity 'isn't what matters.' In the next section I present Parfit's argument.

3.3 PARFIT'S ARGUMENT FROM FISSION

Olson bases his rejection of premise (2) on Parfit's claim that personal identity isn't what practically matters. It is thus important to get clear on exactly how Parfit supports his conclusion that identity isn't what matters.

Parfit, after having suggested that the right criterion of personal identity is the psychological criterion with a non-branching condition, appeals to the possibility of fission in his argument. When contemplating my future fission two general questions seem pertinent to ask: (i) How should we describe the result of the fission with respect to personal identity? This is a semantic-cum-metaphysical question about the truth conditions of identity statements about persons; (ii) What attitude would it be rational for me to have when facing such a procedure, e.g. brain division?

Parfit writes: 'Some people would regard division as being as bad, or nearly as bad, as ordinary death. This reaction is irrational. We ought to regard

division as being about as good as ordinary survival. As I have argued, the two 'products' of this operation would be two different people. Consider my relation to each of these people. Does this relation fail to contain some vital element that is contained in ordinary survival? It seems clear that it does not. I would survive if I stood in this very same relation to only one of the resulting people. It is a fact that someone can survive even if half his brain is destroyed. And on reflection it was clear that I would survive if my whole brain was transplanted into my brother's body. In the case that we are now considering, my relation to each of the resulting people thus contains everything that would be needed for me to survive as that person. It cannot be the *nature* of my relation to each of the resulting people that, in this case, causes it to fail to be survival. Nothing is *missing*. What is wrong can only be the duplication' (Parfit 1987, p. 260).

Parfit's claim is that if I accept all this and still take my impending fission to be as bad as ordinary death, then I am irrational: 'Double survival is not the same as ordinary survival. But this does not make it death. It is even less like death' (Parfit 1987, p. 261). In fission the relation of psychological continuity turns out not to be coextensive with personal identity. But the relation, which makes my survival something that matters to me, must still hold, since there is no intrinsic change in my relation to either of the post fission persons. When I find that it is important that I *survive* this is really a *derivative* concern. What *underderivatively* matters is the relation which doesn't change intrinsically in the multiplication situation. Hence, what is underderivatively significant is that there is *someone* in the future who is psychologically continuous with me. This is Parfit's argument from fission.

An important assumption of Parfit's is that the relation that matters cannot cease to hold due to merely extrinsic facts. Interestingly this makes it different from my identity with some future person which, on Parfit's view, may depend on whether someone else exists. For an interesting discussion of

this assumption see e.g. Johnston (1997). My aim here is not to assess whether we should accept Parfit's argument, but to consider whether, if it holds, it can do the work Olson needs it to do.⁴³

3.4 THE PRACTICAL IDENTITY OF PERSONS

Olson writes: 'We believe that one survives if and only if one's mental features are preserved because that is what it takes for someone to be worthy of our prudential concern, and because it is natural to suppose that prudential concern always coincides with identity. But if we were wrong to think that rational prudential concern always implies identity, it ought not to surprise us much if [Brown] were not [Brownson], but Brainless [i.e. the functioning organism left behind when Brown's cerebrum is removed]' (Olson 1997, p. 56).

If Parfit's thesis is accepted it may be that Transplant is a case in which we come to think of Brownson as identical to Brown even though all that can be assumed is that Brownson is someone whom we'd take to fall within the scope of Brown's quasi-prudential concern, and not within Robinson's. But then it is not clear that the Transplant Intuition supports any particular view about personal identity.

One of the attractions of the debate about personal identity is that it appears to be concerned with the practically quite significant question: who will be *me* in the future? It seems very hard to accept that there can be someone existing tomorrow, who is literally me and for whom I have no reason to be practically concerned. Even in case there might be other relations, which we also take to be relevant for who we have reason to be e.g. prudentially or quasi-prudentially concerned about, it will be very hard to accept a view

⁴³ One might take Parfit to have shown that if the psychological criterion leads to the conclusion that identity isn't practically significant, then so much worse for that criterion. This is not a conclusion Parfit would find of much concern. What we should be discussing is 'what matters' in practical concerns normally taken to be based on personal identity.

according to which there may be someone in the future who is identical to me, and for whom I can now do something to avoid him undergoing severe torture, yet I have no reason to be concerned for this person from a strictly prudential point of view. What this suggests is that it will be very hard to accept a view which doesn't count personal identity as sufficient for having reason to be practically concerned about some future person. Even if practical identity isn't a sufficient condition for numerical identity, it may seem quite plausible to think that numerical identity is sufficient for practical identity. I shall come back to this point in the next section. But first it must be made clear what Olson means when he claims that 'sameness of person' doesn't pick out a relation of numerical identity, but rather a practical identity.

Here's what Olson seems to understand by 'being the same person as' and what I shall refer to as 'practical identity':

x is at time t practically the same person as y at a later time t^* if and only if x ought to be prudentially concerned, at t , for y 's well-being at t^* .⁴⁴

Here's Olson: 'In unusual cases, a single human being such as you or I might be one particular person at one time, in this [practical] sense of 'person',⁴⁵ and another particular person later on. (...) And two numerically different human beings might be the same person, though perhaps not at once. That might be the case if you were destroyed and replaced by a perfect duplicate: perhaps the duplicate would be the same person as you without being you. This is meant to be analogous to the sense in which Clinton is now the same

⁴⁴ Olson (1997), p. 66. Olson also discusses other practical relations than prudential concern such as responsibility and 'the treatment argument.' I have just focused on the prudential case. Notice that practical identity as opposed to numerical identity is time relative. I discuss the temporal identity thesis about numerical identity in chapters 6 and 9.

⁴⁵ I'm not sure Olson is allowed here to speak of human beings as being identical to 'you and I.' Whether 'you and I' are human beings in the sense of being identical is one of the main questions to be discussed.

elected official as Reagan was ten years ago, even though Clinton is not Reagan' (Olson 1997, p. 66).

We certainly have a use of 'the same person' which doesn't entail anything about numerical identity. If someone undergoes some traumatic experience we might say that he is not the same person he used to be. It is quite a common phenomenon that people change in 'personal' respects. Diseases such as Alzheimer's Disease may be characterised as 'personality changing.' The loss of memory and change in personality traits such as interests, values, beliefs, habits and ceasing to recognise or associate with family and friends are changes which we'd often describe as someone's becoming a different person. But in all these cases we seem to think that there is a single individual, who undergoes all these changes. It is not as if the traumatized patient has just come into existence, or that we cease to exist and are replaced by a distinct person when suffering from Alzheimer's Disease.

Even if we do have a use of 'the same person as' which doesn't entail anything about numerical identity, I am not convinced that this is all there is to it. I think that Unger is correct when he claims that an inquiry into our persistence conditions must 'appreciate what is involved in a philosophically adequate conception of ourselves. (...) such an adequate concept must be well suited for engagement with our central prudential thoughts and concerns' (Unger 2000, p. 326). I think identity does have something to do with our practical concerns, and isn't just of derivative importance. But at this point my aim is not to show that Parfit is mistaken in claiming that identity doesn't matter.⁴⁶ I am going to do is to undertake a more modest line, which is to show that if we accept a nonpsychological criterion of personal identity, we must accept a stronger thesis than the one advocated by Parfit, namely the thesis that identity isn't even sufficient for what matters

⁴⁶ I will come back to Parfit's claim in chapter 10.

in practical concern. Someone who is literally identical to *me* may be related to me in such a way that he is of absolutely no interest to me in any special way. Practically speaking 'he could be anybody.'

Let us first explore what Olson's distinction between practical and numerical identity entails with respect to cases such as BST? If Brown and Robinson undergo a BST they stay put. Two organisms change their psychological profile, but not in an arbitrary way. There is a *causal* background for the psychological profile that each organism has after the BST. It is not as if the life they wake up and remember 'as their own' hasn't been led. It is just that it hasn't been led by themselves. There are some interesting consequences of maintaining, as Olson does, that Brownson is numerically identical to Robinson and practically identical to Brown.

Let's say that Brown and Robinson (on Olson's view persons which are essentially human organisms) are told that there will be a BST between them tomorrow. Olson admits that the relation which in ordinary circumstances grounds Brown's prudential concern in that case holds between Brown and Brownson, but since Brownson is Robinson, Brown will be quasi-prudentially concerned about what's going to happen to Robinson after the BST. On the other hand, Olson's view entails that since the relation that matters for quasi-prudential concern doesn't hold between e.g. Robinson and Brownson, Robinson won't have grounds for being prudentially concerned about Brownson despite the fact the Brownson is Robinson himself. Practically speaking, Brownson is another person than Robinson.

When Unger writes that an adequate concept of personal identity must be well suited for engagement with our central prudential thoughts and concerns, he underlines that a view of personal identity, which radically dissociates facts about who will be *me* in the future from who falls within the scope of my prudential concern, must be a view which doesn't get it right.

And it does indeed appear as if there is an urgency to our concern about ourselves, which must be taken into account when discussing personal identity. There is after all an enormous difference in one's response to being told that someone will be hit by a truck this afternoon, and to be informed that that someone is oneself. Indeed a great part of our practical life is governed by, what appear to be, identity-based patterns of concern.⁴⁷ A theory which doesn't accommodate this may I think reasonably be criticised for being inadequate to answer the philosophical question about our persistence.

If we describe the BST on the basis of the physical criterion, Brownson would be Robinson with a new psychology inherited from Brown. However, it seems correct to assume that Robinson has little reason to be concerned about Brownson from a prudential point of view. This response to the BST scenario reflects an additional feature of it. Not only does the BST scenario support the belief that personal identity has something to do with psychological continuity, it also strongly suggests that *mere* physical continuity is not relevant to our patterns of practical concern and not for our identity either. Thus even if it turns out that the BST doesn't provide a master argument for the psychological criterion, it should be clear that this fact doesn't as such support the physical criteria either. It may still seem far-fetched to claim that Brownson is Robinson.

I now turn to consider whether Olson's use of Parfit doesn't extract more from Parfit's argument than Parfit put into it. As it turns out, Parfit's argument doesn't allow for the possibility of the situation I just sketched.

⁴⁷ Perry has interesting discussion of this in Perry (1976).

3.5 A RESPONSE: IDENTITY IS A SUFFICIENT CONDITION FOR WHAT MATTERS EVEN IF NOT A NECESSARY CONDITION

I find that Parfit's claim that identity isn't what matters can be given at least three interpretations: a radical, a moderate, and a minimal interpretation. Olson must give it a radical interpretation for his argument to go through, and I think that the claim is better given a moderate interpretation.

When Parfit claims to have shown that we do not ever care about identity underderivatively this is perhaps to conclude more than he has shown. The argument he presents us with shows that in fission situations, we would find it rational to extend the concern, we in ordinary circumstances have for our own future selves, to future persons who are not strictly identical to us. He then concludes that it is *never* identity that matters underderivatively. Yet it certainly seems to most of us that it is the fact that some future person is oneself which rationalises prudential concern for that person.⁴⁸

A *minimal interpretation* of Parfit's argument would be to say that he has pointed out that there is a relation, which I may stand in to future persons who aren't me, along which it may be rational to *extend* the concern I ordinarily only have for myself. But this is not to say that identity doesn't *normally* have underivative importance. It does matter, but in special circumstances, like the ones considered in fission cases, other relations might also matter in the absence of identity. This is a minimal interpretation. (It doesn't seem as if we appreciate our future selves (or other persons for that

⁴⁸ It is an interesting question whether it is ever justified to treat someone differently because they are identical to a particular person. I do not consider this problem here. However, in his *The Possibility of Altruism* Nagel (1970) interestingly argues that one would not have an integrated conception of oneself, if one didn't have prudential concern. He then utilises this to build up an argument for altruism. Noticing that one should be equally concerned for one's future selves independently of how far into the future they are, I should recognise that I should be equally concerned for present selves who are not identical to me. (This is of course a very simplified description of Nagel's book.)

matter) merely for realizing a personal profile or type.⁴⁹) On this interpretation there is a relevant difference between quasi-prudential and prudential concern in that the latter is essentially different since *identity* has practical significance. The minimal interpretation doesn't give Olson what he needs.

Now to the moderate interpretation. Parfit takes his argument to show that the following conditional doesn't hold:

(2a) if x has prudential concern for some future y (where this is not due to factual error about the identity of y), then y is identical to x .

Now this 'moderate' interpretation only provides Olson with a rejection of half of the biconditional he identifies as premise (2) in the reasoning purportedly governing the Transplant Intuition. Olson must also reject the following conditional:

(2b) if x is identical to some later y , then x has reason to have prudential concern for y .

That is to say, Olson must reject that it is a sufficient condition for someone to fall within the scope of one's prudential concern that it is oneself. This is, I suspect, a more radical claim than Parfit's argument by itself supports. By itself Parfit's thesis only allows that a future person needn't be me in order for him to fall within the scope of my prudential concern, but Olson assumes that it allows that some future person, who is me, need not fall within my prudential concern. But Parfit's thesis doesn't seem to allow the conclusion that Olson wants to draw, namely that a criterion of personal identity may be

⁴⁹ I take Johnston (1997) to be make this line of response to Parfit.

such that a future person, who, according to that criterion, is *me*, isn't related to me in such a way as to fall within the scope of my prudential concern.

3.6 CONCLUSION

My discussion in this chapter has not shown that Olson is wrong claiming that it may be misleading to take the Transplant Intuition to support the psychological criterion. I have merely shown, I think, that his reliance on Parfit's thesis doesn't give him a satisfactory basis for rejecting it. Olson and those sympathetic to a physical criterion may simply just deny the Transplant Intuition), or just deny that identity is a sufficient condition for what matters in prudential concern. What I think can be concluded is that until further argument is presented, the Transplant Intuition still stands as a strong motivation for accepting a psychological criterion. But the psychological criterion faces theoretical objections which are not easy to answer. In the next two chapters I consider two of the main objections to the psychological criterion. I first consider whether the criterion is circular, then I discuss the problem of fission.

4. The Circularity Objection

4.1 INTRODUCTION

I have suggested that personal identity consists in psychological continuity. A central element in the analysis of personal identity in terms of psychological relations is memory. In this chapter I discuss the claim that a criterion which partly analyses personal identity in terms of memory is viciously circular.⁵⁰ The alleged problem is that even if memory is a sufficient condition for personal identity, it can't be part of an informative criterion or reductive explanation, which tells us what personal identity consists in, for the simple reason that in order to assess whether I genuinely remember some past event 'from the inside', it must first be established whether I am the person who had the experience of the event I claim to remember.

4.2 INTRODUCING THE CIRCULARITY OBJECTION

Suppose someone asks me to explain to him what it takes for him to be identical to some past person, and I respond by telling him that 'a person x today is the same person as a person y yesterday if x today remembers what y experienced yesterday.' In that case I assume that person x is familiar with how to establish whether he remembers the experiences of a person at some past time. But in order for him to know, whether he truly remembers what someone experienced in the past, he must be in a position to determine, whether he is the same person as a person, who had the experience he apparently remembers. Hence my criterion in terms of memory is no help as an informative and non-trivial answer to his question. It is really no different from telling him that he is the same person as some past person if he is identical to him.

⁵⁰ Historically this is an objection that goes back to Butler and Reid's criticism of Locke's memory criterion and it has been of great concern to followers of Locke. See Perry (1975, part II) for the historical discussion.

A reason why it seems plausible to think that memory presupposes personal identity is that the statement, 'I remember experiencing E or doing A' is really a claim to *know* that I, myself, experienced E or did A. Our concept of memory is such that one only remembers what one experienced or did oneself in the ordinary sense of 'remember.' Thus the locution, 'I remember what you did last summer', will normally be understood as stating that I remember seeing you do such and such from 'the outside', not that I remember doing what you did.⁵¹ The way I remember what you did last summer contrasts with the way I remember what *I* did last summer. In the latter case the mode of representation of the remembered event is from a first-person perspective or 'from the inside.'

Thus the notion of memory with which the debate on personal identity is concerned is what might be called 'experiential memory' or 'memory from the inside.' What these locutions are supposed to pick out is the way in which we remember doing or experiencing something in the past. This is a different way of remembering than the one I express by saying, 'I remember that the battle of Hastings was in 1066', which may be called 'factual memory' of a past event, though it is no more a way of remembering the past than it is to remember that the atomic number of gold is 79. It is also different from 'practical memory,' which is what I have when I remember how to swim or ride a bike.

As a matter of fact we only remember our *own* past from the inside, and given this fact it is not clear whether our everyday memory statements involve any reflective view on whether personal identity is a condition for veridical experiential memory or not. However, for personal identity theorists it doesn't really matter much whether ordinary speakers take

⁵¹ Of course there is a sense in which I can remember what you did. E.g. if you went water-skiing I might remember doing what you did in case I also went water-skiing, but this is not the sense of remembering from the inside in question.

memory to be conditioned by personal identity. With respect to the philosophical enterprise, what we must consider is whether and how knowing about the past in the way we do, when we have memory knowledge, is related to personal identity. Is it a condition that there is personal identity between rememberer and the subject of the experience or the agent of the action remembered for an apparent memory to be genuine?

It will be helpful to have a case to focus our attention on: Let's say that Bob went to a ball last night and the next day he comes to his friend Bill and happily reports, 'I remember dancing with Suzy last night.' The question is whether it is correct that in order for Bob to have a genuine memory and not just an apparent memory which is illusory, he must be the very same person as the person whose experience of dancing with Suzy he claims to remember. Does Bob claim to remember something that never happened, if he himself didn't dance with Suzy last night?

Following Shoemaker (1970, 1984) and Parfit (1987) I want to suggest that we can maintain that there is a sense in which Bob remembers dancing with Suzy last night which comes out true, even if it is not the case that Bob himself danced with Suzy last night. Reserving 'remember' and 'memory' to denote cognitive mental states, whose veridicality is conditioned by personal identity, I will follow Shoemaker and call this other sense of remembering 'quasi-remembering.' Truly quasi-remembering dancing with Suzy last night doesn't entail that the quasi-rememberer himself had the experience remembered, but merely that *someone* had the experience claimed to be quasi-remembered.

If we can make intelligible the notion of quasi-memory as a way of knowing about the past from the inside, we may be in a position to appeal to that notion, when we explain what personal identity consists, and thereby avoid the charge of circularity. We can say that if *x* today quasi-remembers an

experience had by y yesterday and certain other conditions are satisfied, then x today is the same person as y yesterday. What this means is that Bob will have a genuine memory of dancing with Suzy last night if he quasi-remembers doing so and certain other conditions which are satisfied, and, importantly, we can establish whether the conditions of quasi-memory, together with the other conditions, are fulfilled without presupposing any awareness about personal identity. Or so the suggestion goes.

Let me sum up the problem: There is a difference between apparently remembering dancing with Suzy and truly or genuinely remembering doing so. I might end up with an apparent memory of dancing with lovely Suzy due to Bob's detailed account of his experiences at the ball last night. But I will say something false, if I state that I remember dancing with Suzy. A natural way to distinguish between an apparent memory and a genuine one is to appeal to personal identity. It is because I am not identical to a person who danced with Suzy last night that my memory claim is false. Thus in order to draw the distinction between apparent and genuine memories, one must already be able to establish whether a personal identity claim is true or not. But then it is no good to appeal to the concept of memory in an account of personal identity, which is not merely supposed to state sufficient conditions, but is also intended to have explanatory power. The suggestion is that we can appeal to a notion of quasi-memory which is phenomenally just like memory except that there is no requirement of personal identity for quasi-memories to be genuine.

4.3 MEMORY

In order to assess whether we can appeal to a notion of quasi-memory to avoid circularity it will be helpful to consider in more detail what it is to remember having an experience or doing an action. What's the nature of the epistemic faculty we refer to as memory?

A first requirement for me to remember an experience of an event E can plausibly be said to be this:

Representation I represent or seem to remember experiencing event E.

When considering what are necessary conditions for me to be in a state correctly described as a memory state, it seems clear that I must presently perform some activity, and to begin with it seems plausible to think that I represent a past event to myself. I shall leave aside more specific discussion of what it is to represent a past event and think of it in terms of 'mental imagery,' but representation might not involve imagery at all. However, one way to think about my representation of an experience of a past event is to think of it as my having a mental image or entertaining a sequence of such images. A momentary event we might think of as a 'mental snapshot,' and a temporally extended event we might think of as a 'mental motion picture.'⁵²

A second condition for my remembering a past event appears to be:

Previous Awareness Someone was aware of event E at the time of its occurrence.

If I remember experiencing an event, the remembered experience must have been had by someone. Someone must so to speak have taken the snapshot at the time of the occurrence of the event remembered. If not, then my apparent memory is of course an illusion as it would be in case I seem to remember having lunch with Napoleon on the Moon. Surely I might conjure up a memory impression of such an event, but since there is no one who ever had

⁵² For discussion of the notion of representation see Martin and Deutscher (1966) which in general is a very important article on the subject and contains further discussion of representation p. 166 ff. and p. 172 where there is recognition of the lack of more specific explanation of what 'represents' is supposed to involve. My discussion of quasi-memory owes a lot to Shoemaker (1959, 1970, 1984).

lunch with Napoleon on the moon, since there is no such experience for my memory image to represent, there can be no memory of such an event. There simply isn't any experience corresponding to the content of my apparent memory. So the apparent memory is an illusion.

But is it not also part of our concept of memory that it can't be just any old person's past experience I remember? If I seem to remember having lunch with Churchill in Downing Street and my representation happens to correspond to an experience that someone once had of lunching with Churchill, then on the conditions stated so far, I would remember that experience. But surely this is not what we'd normally understand to be a memory of lunching with Churchill. E.g. I might come to be in a mental state indistinguishable from that which some person who actually did lunch with Churchill is in, when remembering that event due to reading that person's very detailed memoirs, but that doesn't provide me with a genuine memory of the event. The fact that I entertain a memory image of lunching with Churchill, which is an accurate representation of an event that took place, doesn't make me remember that event. The fact that someone has given me a mental snapshot of the event, which I can entertain and in virtue of which I can be said to have quite a bit of knowledge about what happened at the lunch, doesn't make it true that the snapshot is of an experience that I remember. I don't remember lunching with Churchill despite my accurate knowledge of the event.

What makes us wary of thinking that my apparent memory state is genuine is presumably that the following condition isn't satisfied:

Strong Previous Awareness I was aware of E at the time of its occurrence

However Representation and Strong Previous Awareness may not be thought to be sufficient for memory. I might have a representation of an

experience of event E, which corresponds to a past experiential state of mine, but for this to be a case of genuine memory a further requirement must be satisfied. This can be brought out by considering the following situation:

Suppose my uncle is a hypnotist who once recorded my detailed description of being at my first jazz concert. Due to an accident I have completely lost my memories of this experience where this is to say that they are irrevocably gone, not just 'dormant.' Now my hypnotist uncle takes me to his study and bringing me into deep hypnosis he 'installs' (as opposed to recovers) vivid apparent memory states in me, which are indistinguishable from those I had before the accident (I am of course not aware that they are completely similar).

In such a situation there seems to be a rather extraordinary *causal* link between my original experience and my apparent memory of it. Some might find that such extraordinary causal connections will undercut my apparent memory states as genuine memories. I must confess that I don't agree. It would be a rather unusual case, but I don't think the case is much different from a situation in which memories are 'dormant' in the sense that they can't be recalled at will, but nevertheless prompting circumstances may bring them back. I suppose this is often the case when people suffer from traumas and the like, but I will not go into a detailed discussion of this issue.

I think, however, that the case for adding a separate causal requirement on memory can be made stronger if we imagine that my hypnotist uncle after my accident, without any knowledge of my past experiences, just happens to install vivid memories of being at a jazz concert, which brings me to entertain mental states of apparent memories completely similar to the one I was in before the accident. In this situation my apparent memories of being to the jazz concert would be phenomenally like the ones I used to have, but given the complete lack of causal connectedness between the apparent

memory states and the experiences allegedly remembered, I think few of us would accept this as a case of genuinely remembering being at the concert (even if we were willing to accept the former case as one of remembering).

Whatever detailed account we want to give of the kind of causal relation that is relevant to memory, it seems correct to conclude that there's some causal requirement on apparent memory states for them to be genuine. Here's a list of four plausible conditions for my remembering experiencing E:

Representation I represent or seem to remember experiencing a past event E.

Previous Awareness Someone was aware of E at the time of its occurrence.

Strong Previous Awareness I was aware of E at the time of its occurrence.

Causal Requirement My apparent memory is appropriately causally linked to an experience of E at the time of its occurrence.

If by 'remembering' we mean a faculty of knowing about the past which involves the strong previous awareness condition as necessary for the veridicality of seeming memory states, there will be a problem about circularity facing a criterion of personal identity appealing to memory. Let's assume that 'memory' as ordinarily used conceptually involves the satisfaction of the strong previous awareness condition. That is simply how we use the word 'remembering.' In that case we may attempt to make sense of an independently intelligible faculty of knowing about the past which is just like 'memory,' except that it doesn't involve the strong previous awareness condition, and then go on to formulate our criterion of personal identity on the basis of this weaker notion, which Shoemaker calls 'quasi-memory.' A quasi-memory is a state which is phenomenally on par with an ordinary memory. For me to quasi-remember I must seem to remember event E, and there must be someone who had the experience of E which I seem to remember, and my apparent memory must be appropriately causally related to someone's witnessing of E.

The sort of mental states we denote with our ordinary concept of memory will qualify as quasi-memories, but they will be a subclass of the quasi-memories, namely those quasi-memories which also satisfy the strong previous awareness condition. But importantly I could determine whether I genuinely quasi-remember some past event without having to settle any questions about personal identity, and thus a criterion involving only reference to that notion would avoid the vicious circularity objection. One can come to understand what it is to quasi-remember, without having any knowledge of personal identity.⁵³

But it now becomes important to consider whether we can say that the mental states we ordinarily take to be memories can justifiably be said to be quasi-memories that just happen to be states of the same subject as the subject which had the original experience quasi-remembered? And it becomes a pressing issue whether the strong previous awareness condition involved in the concept of memory is really redundant when it comes to characterising what it is to remember, i.e. when it comes to the metaphysical nature of memory as an epistemic faculty for knowing about the past. If 'the strongness' of the strong previous awareness requirement is irrelevant to the epistemic status of ordinary memories, then quasi-memories, whose causal origin might not entail its satisfaction, can be established as a faculty of knowing about the past on par with ordinary memory.

Memory is a faculty of knowing about the past. The suggestion is that the cognitive status of memories is grounded in the satisfaction of the conditions for quasi-memory, and that the fact that the cognitive states we call

⁵³ It must of course be possible to establish that the apparent quasi-memory state is appropriately causally connected to some past experience of E for it to qualify as quasi-memory knowledge of E, without having to obtain knowledge of facts about personal identity.

'memories' must also satisfy the strong previous awareness requirement is not relevant to their epistemic value.

One of the reasons why it is tempting to think that memory in some way provides us with knowledge of our own identity is that it appears to be a way of knowing about our own past in a direct way unmediated by testimony and inferences from present data. When asked by his sceptical friend Bill, how he knows that he danced with Suzy last night, it is quite natural for Bob to say, 'Well, I just remember doing so.' And, the tempting conclusion is, that if Bob's apparent memory is veridical, then we can conclude that he is identical with the person whose experience he remembers. Thus it is tempting to say that personal identity consists partly in memory relations.

The suggestion is that if quasi-memory is an intelligible and possible faculty of knowing, we will in quasi-memory have a way of knowing of 'past events which we did not ourselves witness, of experiences we did not ourselves have, and of actions we did not ourselves perform, that is in all important respects like the knowledge we have of past events, experiences, and actions in remembering them' (Shoemaker 1970, p. 271).

Thus the task becomes this: 'we need to consider (...) whether there could be a kind of knowledge of past events such that someone's having this sort of knowledge of an event does involve there being a correspondence between his present cognitive state and a past cognitive and sensory state that was of the event, but such that this correspondence, although otherwise just like that which exists in memory, does not necessarily involve that past state's having been a state of the very same person who subsequently has the knowledge. Let us speak of such, supposing for the moment that it is possible, as "quasi-memory knowledge," and let us say that a person who has this sort of

knowledge of a past event "quasi-remembers" that past event' (Shoemaker 1970, p. 23-24).

4.4 MEMORY AND IMMUNITY TO ERROR THROUGH MISIDENTIFICATION

Shoemaker picks out two principles which point out the conceptual interdependence between personal identity and memory. The first principle is the condition which I have already presented as part of the ordinary concept of memory, viz. the strong previous awareness condition. The second principle has to do with the fact that memory is immune to error through misidentification with respect to the reference of the first-person pronoun. Quasi-memory has a weaker previous awareness condition than our concept of memory. As it turns out dropping the strong previous awareness conditions entails that quasi-memory will differ from ordinary memory in not being immune to error with respect to the first-person pronoun.

In 'Personal Identity and Memory' (Shoemaker 1959) Shoemaker argues that first-person memory claims to knowledge such as 'I broke the window yesterday' are noncriterial. When remembering that I broke the window yesterday, I do not use a criterion of identity to ascertain that the person whom I remember breaking the window is identical to me.

In the article Shoemaker argues that 'while it is true that one doesn't use bodily identity as a criterion of personal identity when one says on the basis of memory that one did something in the past, this is not because one uses something else as a criterion, but is rather because one uses no criterion at all' (p. 873). How may one come to think that claims to remember experiences of past events don't involve applying a criterion of identity? After all, how can I take myself to remember some past experience without making some judgment about my identity with its original subject? If I think that I

genuinely remember some event E, doesn't that involve judging that I am the person whose experience I remember?

Shoemaker considers the statement, 'I broke the front window yesterday', where this is asserted only on the basis of memory. Two requirements have to be satisfied for it to be a statement based on an identity criterion: (1) I must know by remembering that someone broke the window, and (2) I must have used the criterion to determine that that person was myself and this must be partly on the basis of what I know by memory about the person at that past time.

But from this it follows that the statement must be a conclusion from what is remembered about the person who broke the window and perhaps other facts as well. And the conclusion that I broke the window yesterday, when drawn on the basis of the facts I remember about the window-breaker, and perhaps also facts about myself at present, must be justified by my criterion of identity. Thus it must be presumed that if I had remembered different facts about the person who broke the window, this would have led me to a different conclusion about whether I am him or not.

Shoemaker now notes that if this is correct, then it is not true, strictly speaking, to say, 'I remember that I broke the window.' If 'I remember I broke the window' is a *conclusion from* facts I remember, then it is not itself a memory statement, which is to say, it is not a way of expressing what I remember. As I understand Shoemaker this means that if the statement is a conclusion, then in 'I remember that I broke the window' the idea that the second 'I' refers to the rememberer is really a conclusion from what is remembered, and not a deliverance by the content of the memory. It is not the memory content which reveals an identity.

It is important to distinguish statements which are based on memory from statements which are simply memory reports. Memory reports or statements, as opposed to conclusions based on them, can't be false unless I misremember. On the other hand, I might conclude something which is false on the basis of an accurate memory report. I might be in a memory state which correctly represents the past as it was experienced at the time and draw the conclusion that I saw Suzy at the party, e.g. due to features of the person I remember seeing. But this conclusion might be mistaken if the person I saw really was Jane, Suzy's identical twin. It is one thing to be in a state which is characteristically remembering some past experience of an event, and another thing to be in the state of claiming, on the basis of memory, that such and such facts obtained at that time.

The claim Shoemaker is considering here is that first person past tense statements such as, 'I broke the window,' are memory statements which necessarily involve *application* of a criterion of identity. But if I state by application of a criterion of identity that I broke the window yesterday, then the memory I have of this can't be that *I* broke the window. Rather that *I* broke the window is what I conclude after having consulted my memory knowledge of what happened yesterday, viz. that someone with such and such features broke the window. And from this I conclude on the basis of my criterion of identity that I was that person and hence I broke the window yesterday.

The concept of memory as ordinarily conceived implies that one only remembers some past event if one witnessed the event. So if I truly remember someone breaking the window yesterday, then I was there to witness it. Is my witnessing of the remembered event an element in what is remembered? Is my witnessing of the remembered event a conclusion from what is remembered? It seems to be neither. Rather it is a conclusion from the fact that I genuinely remember.

If I am justified in stating that 'I remember someone breaking the window yesterday,' it then follows that I am justified in stating that 'I was there when the window was broken yesterday.' Now the latter statement is of course a first person past tense statement, and Shoemaker asks us to consider whether it could be grounded on any criterion of personal identity and quickly asserts that it can't be, since 'it is not, (...), a conclusion from what I know about someone who existed in the past. What I know about the past, in the case we are considering, is what I remember, but this statement ['I was there when the window was broken yesterday'] is not a conclusion from what I remember at all; it is a conclusion from that fact that I remember something, not from any of the facts I remember' (Shoemaker 1959, p. 125).

Generally it seems reasonable to claim that my memory statements will only be false insofar as I misremember. However, conclusions based on memory may be false even when I don't misremember anything. Suppose my memory representation of what I saw last night is accurate, and on that basis I conclude that I saw Suzy. In this case I don't misremember, it is not my memory which is false, it is the conclusion I draw on the basis of the information delivered to me in memory. This brings us to the second principle besides the strong previous awareness condition characterizing ordinary memory, namely that it is immune to error through misidentification with respect to the first-person pronoun.

We can distinguish between first-person memory claims such as 'I danced with Suzy last night' and third-person ones such as 'Suzy danced on the table last night.' The latter memory statement might be false due to a misidentification of the person dancing on the table last night, either at the time of the event or later when accurately recalling remembered features of the scene. Let's say that I at the time of the table dancing took myself to see

Suzy dancing on the table, though it really was her twin sister Jane who did so. Third-person memory reports are thus subject to misidentification.

On the other hand, my memory that I danced with Suzy last night is immune to misidentification with respect to the reference or identity of 'I', though not with respect to other aspects of the memory. A memory of dancing with Suzy, while it might be false with respect to the identity of my dancing partner, it may not be false with respect to who I remember dancing with Suzy. It is not as if I remember dancing with Suzy the only mistake being that it wasn't me dancing with her, but someone else. If a memory statement is false, it is not due to the fact that the first-person pronoun doesn't refer to the rememberer, but because some third-personal feature of the memory is mistaken.

Shoemaker writes: 'In general, if at some past time I could have known of someone that he was F, and could at the same time have been mistaken in taking that person to be myself, then the subsequent memory claims I make about the past occasion will be subject to error through misidentification with respect to the first-person pronouns. But if, as is frequently the case, I could not have been mistaken in this way in the past asserting what I then knew by saying "I am F," then my subsequent memory claim "I was F" will be immune to error through misidentification relative to 'I'; that is, it is impossible in such cases that I should accurately remember someone being F but mistakenly take that person to be myself. We might express this by saying that where the present-tense version of a judgment is immune to error through misidentification relative to the first-person pronouns contained in it, this immunity is *preserved* in memory' (Shoemaker 1970, p. 21).

What seems to make for the immunity to error of the statement 'I was F', when made at some time t_2 , is not only that it was immune to error when uttered at t_1 in the present tense, but *additionally* that it was immune to error

at t1 when said *by the same person* who says 'I was F' at t2. Thus, as Shoemaker himself acknowledges, the immunity to error has satisfaction of the strong previous awareness condition as a necessary condition.

However, assuming a faculty of quasi-memory, a way of knowing about the past 'just like' memory not subject to the strong previous awareness condition, it is a genuine question whether the quasi-rememberer is identical to the subject of the experience quasi-remembered, and in that case application of a criterion of identity would be appropriate to determine whether the experience quasi-remembered was had by the quasi-rememberer himself. A criterion would be used to determine whether a quasi-memory is also a memory. A quasi-memory claim thus doesn't seem to be immune to error with respect to the first person pronoun. The fact that I genuinely quasi-remember a past experience doesn't entail that I was the subject of the experience quasi-remembered.

4.5 QUASI-MEMORY AND THE PREVIOUS AWARENESS CONDITION

We have seen that the notion of quasi-memory doesn't respect immunity to error through misidentification with respect to the first-person pronoun and that this is a consequence of giving up on the strong previous awareness condition. The strong previous awareness condition is replaced with a weaker one according to which quasi-memory requires that:

Previous Awareness Requirement on Quasi-Memory My apparent quasi-memory was caused in an appropriate way by someone's experiencing E at the time of its occurrence.

Replacing the strong previous awareness requirement with this weaker one may help us avoid the objectionable circularity, but only if it is possible to spell out what 'an appropriate causal connection' is, without appealing to personal identity considerations.

Shoemaker points out, that in cases such as Transplant, we seem to have good reason to think, that when Brownson wakes up after the transplant procedure and claims to remember the past life of Brown from the inside, the fact that Brownson's cerebrum is physically continuous with the cerebrum underpinning Brown's mental life will provide justification for the claim that Brownson's apparent memories are caused in an appropriate way for them to be genuine. In this case it is tempting and justified to say that Brownson remembers doing and experiencing the things Brown did and experienced in the past. And, importantly, we seem to be able to establish that the appropriate causal relation holds independently of knowing anything about Brownson's identity with some past individual. Our justification for thinking that an appropriate causal relation holds between the apparent memories of Brownson and the past mental states of Brown for it to be a case of memory is not based on any view about their identity.

Consider now the case in which Brown's cerebrum is divided and each hemisphere is transplanted into two empty heads. Presumably we end up with two persons, 'Brownson 1' and 'Brownson 2', both of whom claim to remember Brown's past life, and the total mental state of both seem to stand in an appropriate causal relationship to the total mental state of Brown before the cerebrum division. Now assuming that Brownson 1 and Brownson 2 are distinct persons, it can't be that they are both identical to Brown, and since it would seem arbitrary to think that the one but not the other is identical to Brown, Shoemaker suggests that the best description of the case is that Brown is *replaced* by two persons, who both apparently remember his past life, but neither of whom is identical to him (see Shoemaker 1984, p. 85). Shoemaker takes this to suggest that personal identity is constituted by *non-branching* psychological continuity, a suggestion which I will discuss in the next chapter.

What seems clear is that reflecting on cases of cerebrum transplant and cerebrum division we realise that 'the appropriate causal relation' referred to in the previous awareness condition stated for quasi-memory, can be recognised to obtain without presupposing knowledge of facts about personal identity. Contemplating cases of cerebrum transplant and cerebrum division we find that someone has knowledge from the inside of the past without settling any questions about personal identity first. Thus the notion of quasi-memory may play the role in our criterion of personal identity for which the ordinary concept of memory was unsuited.

4.6 THE CONTENT AND INDEPENDENT INTELLIGIBILITY OF QUASI-MEMORY

Quasi-memory is a faculty of knowing which is *just like* memory in the way it *represents* the past. And while it is not immune to error through misidentification in the way ordinary memory is, it does give *mere* quasi-rememberers, if ever such there be, knowledge of the past from the inside. This means that the first-personal aspect of the way memory delivers the past to the rememberer is not tied up with there being identity of rememberer and past experiential subject. The intelligibility of the concept of quasi-memory thus entails that the first person mode in which we ordinarily represent the past in memory doesn't reflect the identity of rememberer and past subject even in the ordinary memory cases. It has been thought to be controversial to claim that the intelligibility of quasi-memory demonstrates that it is possible to know about the past in the direct and noninferential way we do in ordinary memory in a way that doesn't represent it as knowledge of one's own past. I turn to this criticism in the remaining sections of this chapter.

The problem about circularity which is supposed to be avoided by appealing to quasi-memory is that when we correctly can be said to remember something, it involves establishing (or assuming established) that there is personal identity between rememberer and original subject of the

remembered experience. In case we can't determine whether I myself did or experienced what I appear to remember experiencing or doing, we can't ascertain whether my apparent memory is genuine or not. But to ascertain that I have a veridical quasi-memory of doing or experiencing something in the past there is no such identity requirement.

But it is also part of the notion of quasi-memory that it represents the past in just the way ordinary memory represents the past. The mode of representation is first-personal or 'from the inside.' Evans has criticised the notion of quasi-memory thus:

(...) it is somehow supposed that the intelligibility of the notion of *q*-memory, (...), demonstrates the possibility of a faculty which is both like our memory in giving subjects knowledge of the past, and unlike it in that the content of memory states in no way encroaches upon the question of *whose* past is concerned. The informational states of a *q*-memory faculty announce themselves, so to speak, as *merely q*-memories, so that it seems to the subject that someone or other *F*-ed without its in any way seeming to him that *he* *F*-ed. Obviously this is a fallacy (Evans 1982, p. 248).

McDowell criticises the notion quasi-memory in a similar way when he describes the notion of quasi-memory as being of 'a capacity whose exercises intelligibly constitute retention of knowledge of past states and occurrences 'from within', but in such a way that the identity of their subject – in particular his being one and the same as the quasi-rememberer – is not represented in the content of the retained knowledge' (McDowell 1997, p. 239). McDowell doubts this. He continues: 'We have not been equipped to make sense of mere quasi-memories other than by supposing that they would present themselves as memories – that is, that they would embody an impression, which must be illusory, that the subject of the recalled state or occurrence was oneself' (McDowell 1997, p. 240).

When I entertain a quasi-memory the content of the quasi-memory is not identity-neutral. The way quasi-memories represent the past is not really in accordance with the previous awareness condition merely that *someone* or other did or experienced such and such. Even if it is veridical just as long as *someone* did what I quasi-remember doing, the content of the quasi-memory represents the past action as something *I* did. The way I'd express a quasi-memory will be in the first person: 'I quasi-remember that I danced with Suzy last night.' Pointing out this feature of the content of quasi-memory it might be objected that quasi-memory is not intelligible independently of personal identity, since I simply can't get the content right, unless I have a belief that I am the same as the person whose experiences and actions I apparently quasi-remember.

Thus objection is that the very way in which quasi-memories are entertained and expressed is in the first person mode and that mode is tied up with a belief about the identity of the quasi-rememberer with the original subject. The content of a quasi-memory is represented in a way which we'd naturally express in terms of a first-person statement, e.g. 'I quasi-remember that I danced with Suzy last night.' And that sort of claim is typically understood only on the basis of a prior understanding of claims such as 'I remember dancing with such and such', or 'I danced with such and such', which are true only if I am identical with the person who danced with such and such. But then some grasp of the notion of personal identity is presupposed even when the notion of quasi-memory is concerned.

The way in which I, following Shoemaker, have made a case for the possibility of quasi-memory as a way of knowing about the past on par with ordinary memory is by contemplating various extraordinary ways in which we'd take someone to be in a state just like the one we are ordinarily in when we apparently remember some past event from the inside. One such extraordinary way of setting up a situation in which we'd say that one

person has an apparent memory of the past experience of another person is fission (by cerebrum division).

Evans and McDowell agrees that this will indeed make intelligible the notion of mere quasi-memory, but only as an apparent ordinary memory which is illusory with respect to its first-personal content. It is not as a way of knowing about the past from the inside, which doesn't give the (false) impression of personal identity. We simply don't get a notion of quasi-memory which is an 'autonomously intelligible faculty of knowing the past from a participant's perspective but without commitment to the participant's having been oneself' (McDowell 1997, p. 240). Again the criticism is that in order for me to grasp the notion of a quasi-memory, I must still be able to single out an aspect of my apparent quasi-memory as illusory, and I must be in a position to be aware of it as a false appearance of *personal identity*.

I think that what these criticisms point to is that in the actual world, it is very hard to make sense of quasi-memories as representing the past in an identity neutral way, i.e. in a way that doesn't present the quasi-rememberer as quasi-remembering his own past. If we want to consider a world in which this is not so, we seem to have to move to a world, where mere quasi-memories are commonplace, i.e. were e.g. children inherit quasi-memories of their parents' past lives and during their upbringing are taught only to report quasi-memories in the first-person form, when they are sure that the remembered experience belongs to their own past (cf. Noonan 2003, p. 156). But I think that in such a world, quasi-rememberers would not automatically think of their representation in first-personal terms, i.e. as representations of themselves having done such and such, even though the representation of the past would be 'from the inside.' They would take it to be something to be established independently of the deliverance of their apparent quasi-memory.

Also, in a world where fissions and fusions were commonplace, the first-person mode would not be the standard way of thinking and reporting one's quasi-memory states. It would be a way of presenting the content, which had to await evidence that no fusion or fission had taken place between the occurrence of the quasi-remembered experience and the quasi-memory. In both these possible worlds, the content of quasi-memories would be intrinsically just like the content of quasi-memories in our world. Still, someone could be in a memory state intrinsically like one in our world and not automatically believe himself to be the subject, who had the experience quasi-remembered. Just as we automatically take a quasi-memory to be a memory until evidence to the contrary is presented, so in this other world, one would take one's quasi-memory states to be mere quasi-memories until further evidence came about.

Considering these two possible scenarios I think justifies the suggestion that the *content* of quasi-memories in these two possible worlds is the same due to the sameness in intrinsic features of the quasi-memory states. However, Evans and McDowell would seem to have to claim that the content of a quasi-memory is not determined only by its intrinsic features, but also involves what beliefs are generated on the basis of them, and hence that the *content* of the quasi-memory states in these possible worlds differ from the similar content in our world, despite their similarity in intrinsic features. (Or alternatively they would have to claim that there isn't really a sense in which intrinsic features of a quasi-memory can be held constant across these possible scenarios). So it does seem as if we have a case for claiming that it isn't the content of the quasi-memory, which 'encroaches upon' the identity of the person whose past is quasi-remembered. Rather the belief that one is identical to the original subject of a quasi-memory is separable from the content, and has to do with the fact that in the actual world, we automatically take the content of our quasi-memories to represent an identity due to the

fact that whenever they are veridical they also qualify as memories due to the satisfaction of the strong previous awareness condition.⁵⁴

4.7 CONCLUSION

In this chapter I have considered a reply to the circularity objection against the psychological criterion which partly analyses personal identity in terms of memory. Memory, the objection goes, presupposes personal identity, and hence someone without an understanding of what makes for personal identity can't know whether an apparent memory is genuine or not. In reply to this, Shoemaker has famously introduced the notion of quasi-memory, which involves a relaxed previous awareness condition compared to memory. Instead of the requirement that to be veridical, a memory must be of the past of the rememberer himself, quasi-memory merely requires that the apparent quasi-memory is appropriately causally related to someone's past experience of it. What this means is that when I quasi-remember, I might be mistaken about who had the experience remembered, even though I have an accurate representation of what was experienced. Thus quasi-memory is not immune to error with respect to the first-person pronoun.⁵⁵ In order for quasi-memory to have any claim to be a way of knowing, there must be satisfaction of some causal requirement. A veridical quasi-memory must be causally linked to the past mental state which its content represent. Reflecting on cases such as cerebrum transplant and cerebrum division (or fission) I justified the claim that we can establish that the appropriate causal relation holds without presupposing anything about personal identity, and that one can be in a cognitive state *just like* memory in which one has

⁵⁴ I have taken inspiration from Noonan's discussion of quasi-memory when writing this section (Noonan 2003, p. 153 ff.).

⁵⁵ Shoemaker does think there is a sense in which he can maintain that we can reasonably assume that we were the subject of the experiences we quasi-remember, and that there is a sense in which immunity to error is maintained but I leave those issues aside here (cf. Shoemaker 1970).

unmediated knowledge of the past from the inside without it being one's own past.

Now the fact that quasi-memories present themselves just like ordinary memories, i.e. in the first-person mode, has been claimed to be open to the objection that due to their content quasi-memories aren't intelligible as such independently of beliefs about personal identity. In the last section of this chapter I have shown that it is a mistake to think that the content of quasi-memories is intrinsically identity-biased, as opposed to identity-neutral. The fact that we tend to find it hard not to express the content of quasi-memories without use of the first person is plausibly due to the fact that we automatically come to think of them as representing an identity, since they always do when veridical in our world. But considering possible worlds in which e.g. fissions often occur suggests that we shouldn't take this to be a sign that the characteristic way in which the content of quasi-memories represents the past from the inside can only be understood as representing an identity.

5. Fission, Closest Continuers, and Multiple Occupancy

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;⁵⁶

5.1 INTRODUCTION

I have suggested that we should accept the psychological criterion which states that a person *P* at time *t* is identical to a person *P** at time *t** if and only if *P** at *t** is psychologically continuous with *P* at *t* and the psychological continuity is appropriately caused. A prominent argument against this version of the psychological criterion is that psychological continuity may hold in a one-many form. It is recognised by both opponents and adherents of the psychological criterion that it seems possible that there be more than one psychological continuer of a single person and thus, the argument goes, we should reject the psychological continuity view.

The objection to the psychological criterion which I will discuss in this chapter was first presented by Bernard Williams and has been at the centre of the contemporary debate of personal identity. Williams writes with respect to identity with some item *A*: 'the reduplication problem arises if a supposed criterion of identity allows there to be two distinct items, *B* and *C*, each of which satisfies the criterion in just the way that it would if the other did not exist' (Williams 1970b, p. 77-78).

⁵⁶ Frost, 'The Road not Taken.'

I begin by fleshing out what Williams seems to be getting at when raising the reduplication problem by describing a situation in which, pre-theoretically, it is tempting to say that a single person becomes or divides into two distinct persons. I then move on to discuss responses which have been suggested on behalf of the psychological criterion. I should note that Williams actually took the reduplication problem to be a problem which is especially pertinent to the psychological criterion. According to Williams, his preferred bodily criterion is not vulnerable to the possibility of duplication in the way that the psychological criterion is, but it is doubtful whether Williams is right in thinking that the reduplication problem doesn't put pressure on every continuity account of identity (though, for the record, I think it does). However, in this chapter my aim is merely to assess whether an advocate of the psychological criterion can respond coherently to the challenge posed by the possibility of reduplication.

Is reduplication conceptually impossible? I see no reason to think so. Swinburne has claimed that fission and fusion are impossible scenarios (1984, pp. 18-21 and p. 45), but I find his argument confused. The possibility envisaged in the possibility of reduplication is that the process of psychological continuity, which some claim is not just evidence for, but constitutive of, personal identity, may hold between two distinct persons at one time and a single person at some other time. Swinburne correctly points out that there is a problem about how to deal with this possibility for someone who takes the relation in question to be constitutive of personal identity. He argues that the possibility shows that the continuity relations are at best fallible evidence, but that personal identity must really consist in sameness of indivisible substance. But then he strangely takes this to show that fission is impossible (1984, p. 20). Now of course fission of indivisible persons is conceptually impossible, and fission is perhaps in a sense conceptually impossible if the possibility is taken to be the possibility of 'two equalling one.' But that is exactly what the possibility of fission is supposed

to *show* that some views are committed to, and thus they should be rejected, and as a matter of fact Swinburne appears to employ that line of reasoning himself. It is not fission which is conceptually impossible, surely it is conceptually possible as Swinburne himself describes it in a case of brain division (1984, p. 18). What is impossible is that a process that allows for fission is constitutive of numerical identity, since it would entail that 'two equals one.' But fission as such is merely the situation in which we have two distinct entities related to a single past entity in the same way.

5.2 A CASE OF FISSION

The literature on personal identity contains numerous discussions of fission cases in which two persons seem both to have an equally good claim to being identical to some past person. Sometimes this is described as a case in which a person 'amoeba-like splits into two persons.'⁵⁷ In this chapter I will engage in this debate. The upshot of my discussion is to point out a view which I find plausible. It will be my contention that adopting a temporary identity thesis is the most satisfactory reply for the psychological approach. But first it will be helpful to have a story of fission on which to focus the discussion.

It seems plausible to claim that if a person has a hemispherectomy, that is, if a person has one hemisphere removed from his head, this procedure need not result in the death of the person. The person may continue to exist with just one functioning hemisphere, even if this means that his mental faculties will be somewhat reduced. Take for instance a case in which Brown suffers from a malignant brain tumour and that in order for the tumour not to spread to the rest of his brain it is necessary to remove the right hemisphere. In this case the resulting person will be someone with Brown's left hemisphere, who is in many important ways psychologically continuous

⁵⁷ Wiggins 1967, p. 50 is presumably a *locus classicus* for a case of fission. Contemplating this possibility is allegedly a major reason for some to embark on studying philosophy (cf. Parfit 1987, note 40 Part III).

with Brown and able to underpin a psychological life. In this case it is natural to think that Brown survives with just one hemisphere. And an analogous case can of course be made with respect to the other hemisphere. So here's a claim about the survival of Brown which seems to be quite plausible: Brown might survive without his left hemisphere due to the continued functioning of the right hemisphere and vice versa.

In chapter 2 I described a situation in which Brown's brain is removed and transplanted into Robinson's brainless body.⁵⁸ The result is a person, 'Brownson', with Brown's psychological profile, who thinks he is Brown and so on. I suggested that contemplating this scenario, we should think that Brownson is the same person as Brown. If we find that brain-transplantation is a situation in which Brownson is identical to Brown, who goes with his brain and psychology, we may now imagine that Brownson undergoes hemispherectomy. The result of the removal of Brownson's left hemisphere may again be thought not to upset the survival of Brownson. But this means that Brown may survive having his whole brain transplanted, i.e. as Brownson, and then he may survive a subsequent hemispherectomy.

What this suggests is that Brown could survive having one hemisphere transplanted into Robinson's body. If we are willing to accept that Brownson survives with one hemisphere, why not think that Brown can survive even the transplant of merely one of his hemisphere's into Robinson's body? We get the result that Brown's continued existence is secured if enough of his brain continues to realise the appropriate degree of mental continuity. I shall refer to this as 'the process securing Brown's survival.' (If we are very liberal about the appropriate cause of psychological continuity, we may even think that a way for Brown to survive is to have his psychology transferred

⁵⁸ In this chapter I follow the standard jargon and use the expressions 'brain' and 'body' instead of 'cerebrum' and 'organism.' It is convenient since those are the expressions generally used in the debate.

through a brain-state transfer device outside his brain and body, but I'll leave that possibility aside for now.)

But now we have allowed for the possibility of fission: There seems no reason to rule out that each of Brown's hemispheres is transplanted into distinct recipient bodies. In that case we may call the recipient of Brown's left hemisphere 'Lefty' and the recipient of Brown's right hemisphere 'Righty.' That is to say, there seems to be a possibility of multiplying the process securing Brown's survival. In the case of fission the process securing Brown's survival co-occurs with another instance of the process securing Brown's survival. There are two distinct processes, the transplant of the left hemisphere and the transplant of the right hemisphere, and each of them, it is agreed, secures Brown's survival when occurring without the other occurring. In the singular cases we'd say that Brown survives as or is identical to Lefty and Righty respectively. Fission seems to be a situation in which Brown survives twice over. He survives as Lefty and he survives as Righty, since according to our criterion of personal identity, both Lefty and Righty are psychological continuers of Brown, who satisfy the conditions for being identical to Brown.

One general assumption which must be in place for the possibility of fission to get off the ground is that our use of terms such as 'person' and expressions such as 'same person as', 'brain' and 'hemisphere' doesn't rule out the conceptual possibility of fission. It must be conceptually possible that the process which in one scenario grounds the survival of a person can occur in the manner presented without change in its intrinsic character when co-occurring with a similar process. That is to say, there must be conceptual room for genuine co-occurrence. Again, I shall not doubt this.

If we think there is conceptual room for genuine co-occurrence or multiplication of the process securing Brown's survival, we face the

possibility that it is true in the post-fission scenario to say of Lefty that he is Brown, and also true to say of Righty that he is Brown. But it also seems very plausible to claim that Lefty is not Righty. After all, they display no unity of consciousness and they are constituted by distinct organisms etc. But then we seem to have a case in which, as Parfit puts it, 'double success is a failure' (1987, p. 256). Brown can survive the occurrence of 'the process securing his survival', but not the co-occurrence or multiplication of that process.

5.3 THE REDUPLICATION PROBLEM

It is important to get clear about what is supposed to be problematic about all this. What is it about the possibility of reduplication which puts the psychological criterion in a tight spot?

One immediate problem about the possibility of reduplication is that it shows that we have analysed identity in terms of a relation which may not respect transitivity. The relation of psychological continuity can hold between A and B and A and C, but from this it doesn't follow that it holds between B and C. On the other hand, identity is formally an equivalence relation. Williams claims that 'no principle can be a criterion of identity for things of type T if it relies on what is logically a one-many or many-many relation between things of type T' (Williams 1960, p. 21). On the assumption that Lefty and Righty are distinct persons, we apparently get a conflict with the transitivity of identity from analysing personal identity as psychological continuity. Presumably the adherent of the standard psychological criterion is committed to the following set of statements about the situation after fission has taken place:

- Lefty is not identical to Righty.
- Lefty is identical to Brown.
- Righty is identical to Brown.
- Lefty is identical to Righty.

I shall discuss three endurantist ways in which one might avoid commitment to the truth of these four statements. I shall end up suggesting that specifying the times of the identities and non-identities stated brings out that the four statements, or better, the statements with additional temporal qualification, are not inconsistent after all (the statements are of course inconsistent as stated above).

A central assumption behind the reduplication objection is the following principle:

The only x and y principle If x at time t1 is the same individual as y at a later time t2, that can depend only on facts about x, y and the relationships between them. No fact about any other existing thing is relevant to (deciding) whether x at t1 is (part of the same continuing individual) as y at t2 (cf. Nozick 1981, p, 31-32).

The principle rules out the possibility that there is a relation R constituting the identity over time of e.g. a person at one time and a person at another time not just in worlds where there is branching of that relation, but even in the actual world in which it only holds in a one-one form. It is thus not enough for a relation R to constitute the identity of individuals at different times that there has actually not been any branching.

One might respond to the reduplication objection by rejecting the only x and y principle. This is the option which has appealed to some of the leading personal identity theorists defending a psychological criterion of personal identity, prominently Shoemaker (1984), Garrett (1998) and Parfit (1987), though Parfit merely thinks this is 'the better description' of the fission case. Nozick (1981) of course also defends this approach and I shall come back to it in the next section. On what has been called 'the closest continuer view,' it

should simply be denied that psychological continuity constitutes personal identity when it takes a branching form, but otherwise it does.

An immediate reason why one may feel the attraction of the only x and y principle is, I think, that it manages to rule out situations in which we can correctly say that an object x is identical to an object y partly due to the absence of some third object z. It does seem odd to suggest that I might be related to some past person in just the way I am now related, but not be identical to that person, where this is so, not because there is something missing in my relation to him, but because there is another individual who is related to him in a similar way. I will defend the only x and y principle in more detail later in this chapter.

In what follows I will start out by discussing the so-called 'closest continuer view' developed by Nozick and appealed to by Shoemaker and others. I consider whether its rejection of the only x and y principle is damaging to it. I agree that it is problematic to reject this very plausible principle. I then move on to consider a view suggested by Noonan (2003, p. 139 ff., see also Robinson 1985), 'the multiple occupancy view', which provides a description of the fission situation which is in accordance with the only x and y principle and the psychological criterion. On this view there were 'two persons all along.' I also find this view to be unsatisfactory. My main worry is that it denies what seems absolutely obvious, at least as obvious as the only x and y principle which it is developed to accommodate, namely that there is a single person before fission. Doing cost-benefit metaphysics I take the rejection of this assumption to be too high a price to pay for keeping the only x and y principle intact.

One view which I will not take up is that Lefty and Righty together compose Brown, and may be said to be 'jointly identical' to Brown. The reason why I don't take this view to be an option for the adherent of the psychological

criterion is that it will undermine a major reason for adopting that criterion in the first place, namely that where we have a unified consciousness we have a single person. I will later discuss the view that in fission 'there are two persons all along,' and criticise it for not accepting that the unified consciousness we have before fission is a mark that there is just one person. The view that in fission 'there is one person all along,' I think can be criticised on a similar basis. While it recognises that there is one person before fission, it also suggests that there is a single person made up of Lefty and Righty, despite the fact that they don't combine so as to make up a unified conscious subject, and despite the fact that we'd typically count them as two. Thus I take it that this approach will be subject to arguments parallel to those I will raise against the so-called multiple occupancy thesis.

5.4 FIVE WAYS TO RESPOND TO FISSION

There seem to be five general ways to respond to the result of fission:

- (A) Neither Lefty nor Righty is Brown.
- (B) Lefty is Brown and Righty is Brown.
- (C) Either Lefty is Brown or Righty is Brown, but not both.
- (D) It is indeterminate what the result of fission is.⁵⁹
- (E) Lefty and Righty taken together is Brown.

I will be concerned with views which are versions of either (A) or (B), but I'll briefly present my main reasons for not being drawn towards (C) and (D). As I have just stated, I think (E) can be criticised in a way parallel to the one I raise against the multiple occupancy thesis, which I return to later in the chapter.

⁵⁹ I don't intend this to exclude (C). I suppose one might agree that either Lefty or Righty is Brown, but suggest that even so it is indeterminate which one is Brown.

It seems to me that someone who accepts the view that personal identity consists in relations of psychological continuity and admits that such a relation may hold symmetrically simply can't accept (C). If it is agreed that there is a process which secures Brown's survival and which involves only facts about psychological continuity and its underlying causes, and it is possible that this process may occur simultaneously with another process securing Brown's survival, and that in the co-occurrence situation neither of them is intrinsically different, then accepting (C) one must deny that such processes are sufficient for the survival of a person whether or not they co-occur.

Typically those who accept (C) will claim that the persistence of a person doesn't consist in psychological continuity, but in sameness of 'soul' or the like. Psychological continuity is only to be thought of as good evidence for personal identity, not as a sufficient condition for it. The indivisible soul will turn out to be associated with either Lefty or Righty and thus there will be a relevant difference between the two, even if they stand in a similar continuity relation to Brown. It is not the continuity relation which secures Brown's survival, it is fallible evidence of it. The possibility of fission is taken to show this.

My reason for finding this view unattractive is not that it might not be true that our identity over time has to do with sameness of soul. Rather its unattractiveness stems from the fact that it seems to make our ordinary ways of identifying and reidentifying persons prey to a radical scepticism. There seems to be no reason why we should not be radically deceived about the correlation between psychological (and bodily) continuity and personal identity which underlies their importance in our practice of reidentification.

If one takes the line suggested in (D), the reason might be that it seems we could say different things about what is the result of fission, and there seems

to be no response that doesn't somehow conflict with principles governing our everyday talk about persons and their identity over time such as the only x and y principle, or that two distinct persons can't be in exactly the same place at the same time. We might then conclude that it is indeterminate whether Brown is either or neither or both. In fission personal identity is an indeterminate matter given that there's no unproblematic extension of the expression 'is the same person as.' Johnston seems to be of this opinion when he writes:

'... we should regard the fission case as a case of indeterminacy, a case in which there is no fact of the matter about personal identity. Each of the restricted and consistent extensions of our practice to the fission case are associated with concepts of personal identity which are good competitors to be associated with our use of the expression "is the same person as". Since the question of survival in the fission case is answered differently by each of the competitors there is no fact of the matter about how or whether one survives fission. Personal identity is here an indeterminate matter' (Johnston 1989, p. 393).

I think that there is something to be said for Johnston's view of the matter. In the ordinary way of the world persons don't fission, and in the ordinary way of the world all of the different ways of dealing with fission seem to come out with the same verdict about personal identity. It is not until we take the less parochial situation and consider how to describe it that various commonly accepted principles about personal identity seem to come under pressure, e.g. the only x and y principle. Johnston thus thinks that our hardship in describing the outcome of fission, while respecting e.g. that principle, is to be taken as a sign that there is no fact of the matter as to whether Brown survives or not. While there are various things we could say about it consistent with our ordinary ways of reidentifying persons, there is no determinate thing that we, given our ordinary practice, would say, and no

clear way of determining whether some principle or other which coheres with ordinary cases should be rejected as false (cf. Johnston 1989, p. 393).

While I agree that there are different things we could say, and it might be unclear what we would say, I think that someone who seriously proposes a particular relation as what constitutes identity should say that since both fission offshoots stand in that relation to Brown, it is correct to say on that account that they are the same person as Brown. Someone who, like me, finds that psychological continuity is not only necessary, but also sufficient for personal identity may suggest that we give up on one or another of the principles we take to be inherent in our ordinary practice, but I want to be a little ambitious and consider what we should say.⁶⁰ Of the various things we could say, I want to suggest that the thing we should say is something which takes seriously the view that personal identity consists in psychological continuity with the right cause. I think we must acknowledge that the very same relations which hold in the ordinary case between Brown and some future person who is identical to him also hold between him and his fission products. And I think that this should make us consider in some detail the option of saying that both Lefty and Righty are identical to him and weigh it against the suggestion that Brown is neither of the offshoots.⁶¹ That is to say, I want to discuss whether the typical reply that Brown is replaced by two distinct persons is really the better one for adherents of the psychological criterion.

⁶⁰ Note again that I work within an endurantist framework and thus won't take up e.g. Lewis 1976 and his suggestion that fission is a case of temporal part sharing.

⁶¹ As it turns out I think the option of temporal identity is plausible. Lefty *was* Brown and Righty *was* Brown, but neither of them *is* Brown after the fission.

I now turn to the claim that Brown doesn't survive fission at all, and that Lefty and Righty come into existence only when fission occurs. This is option (A).⁶²

5.5 NOZICK'S CLOSEST CONTINUER SCHEMA: 'TO BE SOMETHING LATER, IS TO BE ITS CLOSEST CONTINUER'⁶³

Responding to fission quite a few psychological identity theorists have appealed to what Nozick calls the closest continuer schema of identity.⁶⁴ In this section I discuss the merits of this way of dealing with the possibility of fission.

According to the closest continuer schema we assess who is the closest continuer of some individual person along the dimensions relevant for the continued existence of persons. The psychological criterion takes the relevant dimension to be psychological continuity, where this involves the requirement that the relevant succeeding states of a continuer 'grow out of' or are causally related to those of its predecessor. On the closest continuer view, the survival of a person from one time to another has as a necessary condition that there is an appropriately related or close enough continuer. But furthermore there must be a single close enough continuer. If there is no such closest close enough continuer the person will cease to exist.⁶⁵

In my discussion of responses to the fission situation I shall assume the 'local' version of the closest continuer schema which has it that the person exists from moment to moment in accordance with this schema. On an

⁶² I will also sometimes refer to this description as 'the replacement thesis', since it describes fission as a case in which Brown is replaced by two new individuals.

⁶³ Similarly, to be something earlier is to be its closest predecessor.

⁶⁴ Notably Unger (1990), Parfit (1987), Shoemaker (1984), and Garrett (1998).

⁶⁵ Thus it seems that for any relation R which is close enough, there might be a relation R* which is closer, and hence B's having R to A is never sufficient for B to be identical with A, what must also be the case is that there is no other entity C which is either R or R* related to A. Again this just strikes me as a very odd view about whether A and B are identical or not.

alternative 'global' version, a person may not have as his continuer the immediately closest continuer. E.g. in a case where the closest continuer is extremely short-lived and there is a close enough continuer, who has continuers further into the future, this may be a relevant consideration. These are issues of detail which one must work out if one accepts the closest continuer schema. Presently I am concerned with what to say about a situation in which there is a tie for the title of being Brown's closest continuer, and I assume that the longevity of Lefty and Righty is not material to our response.⁶⁶ Lefty and Righty are equally close (and close enough) continuers of Brown along the dimension deemed relevant for personal identity. According to the closest continuer schema, we should thus respond to this case as one in which Brown doesn't survive. It is this response I will focus on.

When suggesting the closest continuer schema Nozick is quick to point out that he doesn't accept the only x and y principle on which Williams bases his reduplication objection. Here it is again:

The only x and y principle If x at time t1 is the same individual as y at a later time t2, that can depend only on facts about x, y and the relationships between them. No fact about any other existing thing is relevant to (deciding) whether x at t1 is (part of the same continuing individual) y at t2 (Nozick 1981, p, 31-32).⁶⁷

⁶⁶ I mention the local/global distinction because it is natural to ask, when presented with a brain division case in which the two hemispheres are both around for a while before one of them is destroyed, whether it then isn't a genuine case of fission, even though only one of them survives for more than a short while.

⁶⁷ Nozick accepts a temporal parts ontology. I have suggested that the endurantist formulate the question about personal identity as being about what makes events at different times parts of the history of a single person. The principle then will say something like this: If two events are parts of the history of a single entity of a certain kind in one situation, then they must also be parts of the history of a single entity of the kind in any second situation in which both they, and all the events which are parts of the history of the entity in the first situation, remain present. Noonan suggests that we accept a 'Cambridge criterion' for transworld identity of events according to which 'it is a sufficient condition of the events

Nozick motivates his closest continuer schema by pointing out the way it allows us to deal with puzzle cases such as the Ship of Theseus case:

The Ship of Theseus⁶⁸

Theseus leaves port in his ship, call it 'Original,' and sails around the Mediterranean Sea continuously changing Original's parts to repair it. A clever collector manages to collect each part that Theseus throws away and replaces when repairing his ship. After some years, Theseus sails around on a ship, call it 'Repair', which doesn't have a single part in common with the ship he started out with, while our clever collector has managed to reassemble all the original parts continuously thrown away by Theseus to a ship, call it 'Reassembled.'

Now the question is what happened to the Ship of Theseus? Does it still exist? And if so, which of the competing candidates for being the original ship has the better claim? Repair is spatiotemporally continuous with the original ship, but has no parts in common with the original ship. Reassembled on the other hand can claim to have the very same parts as the original ship, but there is no spatiotemporal continuity under the concept 'ship' between the original ship and Reassembled. According to Nozick,

occurring at a certain location in two situations being identical, that with respect to what happens at that location, there is a mere Cambridge difference between the two situations. The claim is that event identity can be decided independently of object identity (Noonan 2003). A 'Cambridge change' is a change we might say an object undergoes merely due to the change in the truth-value of a statement such as 'the Prime Minister is Tony Blair.' I change from having the property of living in a world where Tony Blair is the British Prime Minister when he one day leaves office, but this is not a change in 'what happens at my location.' Now presumably the idea that we can have sameness of events determined independently of what object's history they are part of entails that the fact that the object whose history they are part of is *this object*, i.e. has the property of being *this particular thing*, its *thisness*, is a 'Cambridge property.'

⁶⁸ This is a version I make up. The original discussion of Theseus ship can be traced back to antiquity, but I suppose Hobbes presents the problem in the form it is typically discussed today (cf. Lowe 2002, p. 25).

adopting the closest continuer schema allows us to 'sort out and structure the issues' which this case presents us with.

The more specific question about which of the two ships we should take to be the Ship of Theseus will depend on what weighted sum of dimensions we take to be relevant for closeness of ships. Let's say that we find that the two candidates for being the Ship of Theseus are equally close continuers on the relevant dimensions; Repair has a strong claim due to spatiotemporal continuity and continuity of parts, whereas Reassembly has a strong claim due to identity of parts. On the closest continuer schema this means that neither of them is the Ship of Theseus. But, Nozick submits, 'even when the two properties receive equal weight, if there actually had been one ship existing without the other, then it, as the closest continuer, would be the original ship' (Nozick 1981, p. 34).

The closest continuer schema presents a necessary condition for the identity of y at t_2 with x at t_1 , namely that it is x 's closest continuer where this involves being closer than all other individuals; if there is some other individual z at t_2 which is equally close, then neither y nor z is identical to x . Closeness is measured along the dimensions relevant for the kind of individual in question. Furthermore, the notion of something's being a closest continuer involves a causal requirement such that its properties, at least those relevant for identity, must be causally dependent on those of the predecessor.⁶⁹

What makes the closest continuer schema attractive with respect to the Ship of Theseus situation is that we want to maintain that even if Repair is the

⁶⁹ It is worth noting that Nozick allows for temporal gaps in the continuity relation if that is the best sort of continuity we have. As a matter of fact, while it is plausible to think that there must be causal relations involved in the closest continuer schema, one might, I suppose, adopt a closest continuer schema which doesn't even involve causal relations.

original ship in the case described, Reassembly surely would be the original ship in case there had not been continuous repair. One and the same ship can be dismantled and reassembled and that is what would be the case in a situation which didn't involve continuous repair of the Ship of Theseus. We have two processes securing the survival of Theseus's ship, which, when they co-occur, seem to rule out the survival of that very ship. So it is natural to suggest that had it not been for the existence of Repair, Reassembled would be the Ship of Theseus and vice versa. But this contradicts the only x and y principle.

When it comes to the case of the fission of a person, we find ourselves confronted with a situation in which there are two equally close, and close enough, continuers, Lefty and Righty, and thus on the closest continuer schema, we should conclude that neither of them is Brown; Brown doesn't survive fission despite there being psychological continuers aplenty. The central challenge to the closest continuer theorist is to make good his claim that had it not been for the existence of Lefty, Righty would have been Brown and *vice versa*.

The reduplication objection against the psychological criterion involves the acceptance of the only x and y principle. Nozick's claim is that it is false and hence no criticism against the psychological criterion if it doesn't respect it. If identity conforms to the closest continuer schema we can just add a non-branching clause to the psychological criterion as a response to the problem of fission. In order to reject this way of dealing with the reduplication objection it must be shown that we'd do better to accept the only x and y principle than to reject it. We are considering three possible situations:

Mono-1: The left hemisphere is transplanted, the right destroyed⁷⁰

Mono-2: The right hemisphere is transplanted, the left destroyed

Fission: Both the right hemisphere and the left hemisphere are transplanted

One criticism of the closest continuer view is that its rejection of the only x and y principle leads to a conflict with the necessity of identity.⁷¹ If Lefty is Brown in Mono-1 and Righty is Brown in Mono-2, and names are rigid designators, then Fission will be a possible situation in which, say Lefty, is not identical to the person he is identical to in another possible situation.

To this objection the closest continuer theorist may reply that the correct description of the modal features of the case does not entail that he is committed to denying the necessity of identity. If we call the person with the left hemisphere in Mono-1 'Lefty*' and the person with the right hemisphere in Mono-2 'Righty*', and reserve the names 'Lefty' and 'Righty' for the persons resulting from fission, then treating these names as rigid designators we get that the referent of 'Lefty*' is not the referent of 'Lefty.' In the situations depicted in e.g. Mono-1 and Fission, the person picked out as 'the person with Brown's left hemisphere' does not pick out the same person in Mono-1 as it does in Fission.

Assuming the necessity of identity, we get the result that Fission is not a case in which we have Lefty* and Righty* co-existing without either of them being Brown. It can be maintained that Lefty isn't Lefty* and Righty isn't Righty*. So the closest continuer theorist is not committed to denying the

⁷⁰ Doesn't the hemisphere which is not transplanted realise a continuer for at least a short while just after the brain-division, and if so how can we have a genuine case in which there is just one continuer? I think we can get around this problem by rephrasing the fission situation in terms of a brain-state transfer device and take the mono-cases to be those in which only one brain is 'rewired' with Brown's psychology. The focus should be on the fact that there is a relation of psychological continuity which may multiply which is taken to provide an analysis of personal identity. Another way of dealing with this is to invoke a notion of closest continuer involving longevity as a parameter.

⁷¹ See Nozick 1981, chapter 1, note 9.

necessity of identity. It can say that the relata in Fission are different from those in the mono-cases.⁷²

But this is not a satisfactory reply. On the sketched account, it seems as if Fission is a case of which we could say that Lefty and Righty should count themselves lucky that the other one exists, since neither of them would exist were it not for the existence of the other. In case Righty didn't exist, we would have the situation in which Lefty* exists, and that is a situation which doesn't involve Lefty in it.

This is a rather unhappy result that seems to lead to a pretty implausible claim: The process which brings about Lefty* in Mono-1 does not lead to the existence of Lefty* in Fission, but to the existence of a different person, Lefty, which is to say that the processes securing Brown's survival in Mono-1 and Mono-2 don't secure it in Fission, even though they both occur in Fission. The identity of the processes seem to be constant in the three sketched situations even if the identity of the persons involved is different as suggested by the closest continuer view.

The problem is this: In Mono-1 and Fission we have the same surgeons removing the same hemisphere of presumably the same person and transplanting it into the same brainless body etc. The series of events in the two situations appear to be the same, but the contention is that even so they may be events in the life of one person in the one situation and in the life of another person in another situation, and this is because there is something going on in a different room with Brown's right hemisphere.

⁷² Another option might be to bite the bullet and claim that the same persons are picked out in all three cases, but that there is only identity in the mono-cases and thereby maintain that in fission the *existence* of the one doesn't depend on the existence of the other, but that the answer to the identity question does.

The advocate of the closest continuer view may reply that we can't individuate events without individuating the object in whose history they occur. But this simply seems a far fetched reply. As Johnston writes:

'Intuitive conceptions [of the processes securing a person's survival] will also allow that such processes, made up as they might be of many bodily and mental sub processes, can be individuated without first settling the question of whether they have secured some person's survival. Given these intuitive conditions, we may suppose that the two diverging processes in the fission case are just the processes which in each of the other two worlds secured the existence of Brown' (Johnston 1989, p. 381).

I conclude that all this should make us hesitate to adopt the closest continuer view. In the next section I consider an account of Fission which purportedly accepts both the only x and y principle and the psychological criterion.

5.6 THE MULTIPLE OCCUPANCY THESIS

I have concluded that if we accept the psychological criterion, it is difficult to maintain that Brown doesn't survive fission merely due to the fact that there is a tie for being his closest continuer.

If we reject the closest continuer theory and accept the only x and y principle, we seem to advocate a view according to which Brown survives fission as both of the fission products. I.e. we must accept a view which makes sense of the claim that Lefty is the same person as Brown and that Righty is the same person as Brown. What we seem to have to say is that both Lefty and Righty existed all along, i.e. before the fission. I think that's what we should say if we accept the psychological criterion, and I don't think that it should be thought of as very odd or counterintuitive. It is a clear consequence of taking the psychological criterion seriously that Lefty and Righty must have been around before the fission, and the challenge is to provide some explanation

of how to make sense of this claim in a consistent way, i.e. in a way which respects identity as an equivalence relation.

I now turn to consider a way to respond to fission which falls under (B). Noonan (2003, p. 139 ff., and also Robinson 1985) suggests that we should accept a 'multiple occupancy thesis' according to which it is a mistake to think that 'Brown' denotes the same person in both Mono-1 and Mono-2. In Mono-1 it denotes Lefty. In Mono-2 it denotes Righty. This follows from recognizing that Lefty and Righty both occur after fission and are distinct. But then the question becomes who 'Brown' denotes in Fission? According to Noonan it must be acknowledged that in cases of fission, there are as many pre-fission persons as post-fission persons, i.e. in the case at hand there are really two persons all along, Lefty and Righty. So the task for the adherent of the multiple occupancy thesis is to make sense of the claim that Lefty and Righty exist as distinct persons before the fission, and some explanation has to be given of the fact that we seem to take 'Brown' to denote a single person before the fission.

One immediate problem about the multiple occupancy thesis is that it simply seems very implausible to think that there is more than one person before fission. Surely the way fission is described is as a case in which one thing of some kind becomes two things of that kind. An amoeba splits into two amoebas, a road diverge into two roads, and a person called Brown fissions into two persons, Lefty and Righty. We don't want to say that there are two persons before fission.

Noonan points out two responses to this criticism. First he tries to deflate the counter intuitiveness of the claim that there is multiple occupancy in the fission case. He agrees that we normally describe it as a case in which one person becomes two persons, 'but', he writes, 'is it not also in some degree tempting to describe them as cases in which two people were present all

along, though this did not become obvious until after the [double] brain transplant? I submit that it is' (Noonan 2003, p. 141). I must confess that I find there to be reasons to resist this temptation if one has it at all.

The second line of response to the claim that we think there is one person before fission involves revising our concept of counting persons at a time. When at a given time I count the persons in my office, I usually do so on the assumption that if the persons in my office are not identical, then they should not be counted as one person. For Bob and Bill to be counted as one person, the person referred to by those names must be one and the same.

Yet, Noonan (and also Lewis 1976) points out, that we can count by other relations than identity. We can count the persons in my office by the relation 'has the same height as.' I.e. we can count by equivalence classes of persons which groups a plurality of persons into a single class due to their standing in some equivalence relation. In general we can make sense of counting a plurality of Fs as one, if for the Fs x and y , xRy . This is basically to say that we can count by a relation R , where R is not identity. The claim Noonan makes is that the common-sense man, when counting persons, does so in terms of an equivalence relation such as spatiotemporal coincidence at all times before the fission and not in terms of numerical identity.

Adopting the multiple occupancy thesis and the revised notion of what we count when counting persons at a time we can avoid the inconsistent set of sentences. Fission should be seen as a case which shows that we do not really count by identity. Lefty and Righty were around before fission. However, there was not a single person picked out as 'Brown' before the fission. I have worries about the multiple occupancy thesis.

One problem about the alternative view of counting persons suggested is that it is not merely the suggestion that there is a fact of the matter as to how

many persons are sitting in my chair at the moment, numerically speaking, but that we don't count by numerical identity. It is not merely the claim that there might be more than one person sitting where I am sitting, say two, but that there is no fact of the matter, determined by the present state of the world, which determines how many persons there are at present. When we ask how many persons are sitting in my chair, the answer implied by multiple occupancy thesis is not just that it depends on what happens in the future, but that there isn't a determinate answer at all. I find this to be too contentious to accept.

I assume here a view about the future according to which the future does not exist. If future facts determine what facts obtain now e.g. how many persons are present in my room, and it is open what happens in the future, then presently there are no facts about how many persons are in my room. Maybe tomorrow it turns out that there were two, but then maybe 10 years from now it turns out that there were ten. This worry is not epistemological but metaphysical. That we don't know how many persons there are now due to our lack of knowledge of the future is not my main worry. The main worry is, so to speak, that there is no fact of which we are ignorant, which determines whether there is one or more persons sitting in my chair now.⁷³

A second problem I think is this⁷⁴: If Lefty and Righty exist before fission, then there must be something which distinguishes them from one another to make them two. But what could that be? Presumably they are composed of

⁷³ There is an analogous question with respect to fusion and the existence of the past. How many persons are sharing my consciousness and body now depends on whether there has been any fusions in the past. If there are not facts about the past, then there seems to be an analogous complaint against the multiple occupancy and the way it would have to deal with fusions. There are intricate metaphysical issues in the philosophy of time which are relevant to this discussion. But my aim here is merely to present a consideration which I think goes against the multiple occupancy view. While I do find it hard to deny that the future is open, I am not confident that the same is the case with respect to the past. Hence I am not committing myself to presentism here.

⁷⁴ Here I am inspired by Burke (1994) and his argument against material coincidence of numerically distinct entities due to a difference in their persistence conditions.

exactly the same parts now, and tracing the one back in time will not bring us to a time and place which isn't simultaneously occupied by the other. However, it seems as if there are future properties which distinguish them. Lefty, we may assume, wakes up in room 102 and Righty in 104, hence they have different future properties. But I think this suggestion fails. Let's say that due to being persons they both have the same persistence conditions. Now it is not clear that we are tracing the one and not the other into the future when, starting out before fission, we single out a branch of psychological continuity. So why think there is someone before fission, Righty, who doesn't have the future property of being in e.g. 102, if we ascribe that property to Righty applying the very same persistence conditions to him as we'd use to trace Lefty into the future? There seems to be nothing which explains their alleged distinctness before fission.

I find there is a third problem about the multiple occupancy thesis.⁷⁵ It seems as if there is a possible world in which Brown dies the day before fission. In that case it would seem that even the multiple occupancy thesis will have to say that 'the life of Brown' denoted a single life, but on the multiple occupancy view the very same events might have constituted the life of two persons in case Brown had undergone fission. That is to say, it is natural to assume that the facts in the fission world and the nonfission world up to the time of fission or the death of Brown are the same. But it seems as if the very same facts in one possible world constitute the existence of two persons and in another possible world constitutes the existence of a single person. This is just very hard to believe.

Finally I also find that the claim that there are two persons all along makes it unclear why we should think that there is not one person all along, i.e. that a single person exists both before and after fission. It seems arbitrary to think

⁷⁵ See e.g. Noonan 1985, p. 198-199. Again I assume non-realism about the future.

that unity of consciousness is necessary but not sufficient for there being a single person *at a time*. One of the main reasons to accept that fission is really a case in which there are two distinct persons in the post-fission situation is that there is no unity of consciousness with respect to Lefty and Righty's mental life.⁷⁶ At least this is something that an adherent of the psychological criterion will naturally be inclined to think. But if unity of consciousness is not the mark of a single person, something it is not on the multiple occupancy thesis, it becomes less obvious why we should not think that Lefty and Righty are a single person. Similarly, fusion would not be a case of two persons becoming one, but one in which two persons somehow shared one consciousness while being distinct conscious beings.

5.7 CONCLUSION

I think there are too many problems facing the multiple occupancy view for it to be a viable response. I am sympathetic to its acceptance of the only *x* and *y* principle, but my main objection is that it involves that there is more than one person thinking and doing the things Brown does before fission.⁷⁷ However, I do think it worthwhile to maintain that we can say of Lefty and Righty that the events before fission are events in their past life. So in some sense to be developed they were around before fission. I now move on to discuss a view which seems to accommodate both the claim that there is a single person picked out as Brown before fission and two distinct persons picked out as Lefty and Righty after fission which are not replacements of Brown, but can be said to have been identical to him.

⁷⁶ Of course we might also cite the fact that we have two independently functioning human organisms etc. But we are considering ways for a proponent of the psychological criterion to respond.

⁷⁷ If one adopts a temporal parts metaphysics and take 'Brown' to refer to a shared person-stage, which is a part of two distinct temporally extended persons it may be thought that there is then just a single entity picked out before fission, the person-stage named 'Brown.' But then one e.g. faces a question about the status of the person-stage. Is it itself a person?

6. Fission and the Temporary Identity Thesis

6.1 TEMPORARY IDENTITY AND LIFETIMES

In this chapter I will argue that in order to give a satisfactory account of the possibility of fission, we'll do well to adopt the thesis that identity is temporary as opposed to eternal, which in turn entails that identity is contingent as opposed to necessary. Now this is not something many will want to accept. As it happens quite a few philosophers might even think that the necessity of identity is 'part of the concept' of identity. However, I will end up suggesting that for someone who, like me, wants to defend the claim that both Lefty and Righty in some sense have a history which stretches back to before fission, and who is also keen to maintain that Brown is not going to cease to exist because there's a plurality of satisfactory candidates for being him, will do well to subscribe to the view that identity is temporary.

I don't think the problems pertaining to the alleged possibility of person fission that I'm concerned with alone justify adopting the temporary identity thesis. But I think there are reasons to think that the thesis that identity is temporary may in general provide us with an elegant way of dealing with a number of the standard puzzle cases at the centre of much modern metaphysics. The project of providing a general defence of the temporary identity thesis is given its most elaborate defence in Gallois (1998). When I began to be attracted to the idea that temporary identity is an appropriate response to the problem of fission, Gallois's book boosted my confidence in the feasibility of the temporary identity thesis. I'm afraid that a complete discussion of all the pros and cons of the temporary identity view is not possible within the space of this thesis. However, I will address what some

might think of as short-comings of temporary identity at the end of this chapter.⁷⁸

In the rest of this chapter I consider the suggestion that Lefty and Righty are identical before the fission, but not after the fission.⁷⁹ The general thesis which this claim reflects I will call the temporary identity thesis, or 'the thesis' for short:

Temporary Identity Thesis $(\exists x)(\exists y)(\exists t)(\exists t^*)[\text{at } t: x = y \ \& \ \text{at } t^*: \text{not } (x = y)]$ ⁸⁰

The temporary identity thesis takes fission to be a case in which we should say that persons who are distinct at one time are identical at another.⁸¹ As such it entails that we can have contingent identity, but I will not presently discuss contingent identity.

I think it is important to point out that when claiming that persons and other kinds of entities may be identical at one time and not at another, this need not be thought of as a rejection of the concept of identity as it is normally conceived: An equivalence relation respecting Leibniz's Law. It is not claimed that a thing can be identical to anything but itself at any time or at any possible world. Thus I take the views to be considered here as respecting the notion of identity which is typically assumed by those working on personal identity. The thesis is not revisionist about the concept of identity.

⁷⁸ I come back to the temporary identity thesis in chapter 10, where I also outline some of the interesting consequences of adopting that thesis with respect to questions about quasi-memory and what matters in survival.

⁷⁹ Perry presents his view in 'Can the Self Divide?' (Perry 1972). Gallois defends his occasional identity thesis in his book *Occasions of Identity* (Gallois 1998).

⁸⁰ I take this formulation from Gallois 1998, p. 89. In my discussion of temporary identity I owe much to Part II of Gallois' book which helped me a great deal in structuring the issues.

⁸¹ As it turns out this is not quite what Perry ends up saying.

Identity is what we take it to be, but, the claim is, it is a relation about which we can reasonably ask *when* it holds between objects.⁸²

The complaint under consideration is that if we accept the psychological criterion of personal identity, it seems we must say that after the fission both Lefty and Righty stand in the relation to Brown which constitutes personal identity. On the face of it we thus seem to have to accept the following statements:

- Lefty is identical to Brown.
- Righty is identical to Brown.

But we also assume that it is correct to say that

- Righty is not identical to Lefty.

But these three statements taken together violate the principle of the transitivity of identity⁸³:

Transitivity $(x)(y)(z)[(x = y \ \& \ y = z) \rightarrow x = z]$

If Lefty is identical to Brown and Righty is identical to Brown, Transitivity entails that:

- Lefty is identical to Righty.

⁸² It might be thought that it sounds strange to say that '2 + 5 = 7' is true today, as if it might not be true tomorrow. However the strangeness of this claim I think has more to do with the fact that it is a mathematical truth, than with the fact that it's an identity claim. The possibility of fission and other puzzling cases I think justifies taking the temporary identity thesis seriously.

⁸³ I will use the locution 'violate the transitivity of identity,' as a way of saying that it violates that identity is an equivalence relation and thus also involves symmetry and reflexivity.

But then we seem both to accept and deny that Lefty is identical to Righty. Some think that we then have to give up the psychological criterion. It can't be a correct analysis of personal identity, since it is committed to the truth of all four statements. I have previously discussed and rejected two ways of dealing with this objection. On the closest continuer view, Lefty and Righty replace the person Brown; on the multiple occupancy view, there is not a single person before fission; both Lefty and Righty exist before fission, but they are not identical.

On the present suggestion there is a single person before fission and two distinct persons after fission. In this there is agreement with the closest continuer view. However, the temporary identity thesis disagrees with the closest continuer view in its claim that the two persons after fission were not around before fission. The temporary identity thesis agrees with the multiple occupancy thesis that 'Lefty and Righty were there all along', but it disagrees with the claim that they weren't identical before fission. As such it combines what I take to be the plausible aspects of the two views rejected while avoiding the unattractive aspects of replacement and multiple occupation.

The question is of course whether it is a view which faces problems of its own. The aim of the coming sections is to assess whether the temporary identity view provides a satisfactory way of dealing with the transitivity objection raised against the psychological approach. The suggestion is that if we take identity to be temporary in the case of fission, we are not committed to the inconsistent set of identity statements. On the temporary identity account of fission, the statements are ambiguous as to when they are supposed to be true. When the time of the identities is made clear they turn out not to be inconsistent at all.

So the reply I suggest in response to the transitivity complaint is to point out that before the fission, there was one person and Lefty and Righty were

identical with that person before the fission. After the fission, Lefty and Righty are distinct, which is to say that there is not a single person with whom they are both identical after fission. The transitivity complaint overlooks that the alleged contradiction stems from assuming that it doesn't matter at what time the allegedly contradictory statements are held to be true. The statements, one might point out, are more perspicuously stated thus:

- Before fission: Righty is identical with Brown.
- Before fission: Lefty is identical with Brown.
- Before fission: Righty is identical to Lefty.
- After fission: Righty is not identical to Lefty.

If one restricts transitivity of identity to identity holding at the same time, there doesn't seem to be a problem maintaining these four statements. The adherent of the temporary identity thesis will only subscribe to a temporally restricted principle of the transitivity of identity:

T-Transitivity (x)(y)(z)(t) [(t: x = y & t: y = z) → t: x = z]

This temporally indexed version of the transitivity principle entails that if Lefty is identical to Brown at t and Brown is identical to Righty at t, then Lefty is identical to Righty at t. But it doesn't entail that if Lefty and Righty are identical before the fission, then they are identical after fission.⁸⁴

It might be objected that even this description is untenable and that Lefty will still be said to be identical to Righty after fission, which we agree is unacceptable. Here's an argument for that conclusion⁸⁵:

⁸⁴ I will later consider how to deal with the derivation of 'x will F' from 'x=y and y will F.'

⁸⁵ Cf. van Inwagen 1981 and Gallois 1998, p. 75 ff. This way of stating it originates from van Inwagen's paper. My presentation of this argument is based on Gallois, pp. 75-76.

- (a) The person who is Lefty after fission is identical to the person who is Lefty before fission.
- (b) The person who is Righty before fission is identical to the person who is Righty after fission.
- (c) The person who is identical to Lefty before fission is identical to the person who is Righty before fission.
- (d) The person who is Lefty after fission is identical to the person who is Righty before fission (from a, c and transitivity).
- (e) The person who is Lefty after fission is identical to the person who is Righty after fission (from b, d, and transitivity).
- (f) Lefty is identical to Righty after fission.

However if this argument is to show that the view under consideration leads to contradiction, then it must be made clear when the identity statements are supposed to be true. The first four premises must be formulated in a way that respects that one only subscribes to T-Transitivity, so (a)-(e) should be prefixed with 'Before fission ...' But in that case it is not clear that the conclusion in (f) must be accepted.

- (A) Before fission: the person who is Lefty after fission is identical to the person who is Lefty before fission.
- (B) Before fission: the person who is Righty before fission is identical to the person who is Righty after fission.
- (C) Before fission: the person who is identical to Lefty before fission is identical to the person who is Righty before fission.
- (D) Before fission: the person who is Lefty after fission is identical to the person who is Righty before fission.

But from this we get,

(E) Before fission: the person who is Lefty after fission is identical to the person who is Righty after fission (T-Transitivity).

And if we adopt temporary identity is not clear that (E) is equivalent to (f) which says that Lefty is identical to Righty after fission. The adherent of the psychological approach adopting a notion of temporary identity will of course say that (E) is equivalent to:

(F) Before fission: Lefty is identical to Righty after fission.

But (F) doesn't contradict the view that it is not the case that after fission Lefty is identical with Righty.

Now the argument (a) to (f) was supposed to show that if we accept that Lefty and Righty are identical before fission, then even in case only T-transitivity is accepted, the absurd conclusion that after fission Lefty is identical with Righty follows. But it has been demonstrated that in that case, the premises and conclusions must also be indexed with the time at which the transitivity is supposed to hold, i.e. 'before fission.'

Now one might wonder whether the argument goes through if instead of 'before fission' we prefix 'after fission.' However, the argument is based on granting the temporary identity theorist his restricted T-transitivity, and since 'after fission' picks out a time, where Lefty and Righty are not claimed to be identical, and thus can't be assumed to have all their properties in common at that time, the argument won't go through. The 'reverse argument' one might have in mind if suggesting the 'after fission' prefix, is presumably one in terms of 'after fusion', since after fusion there would be T-transitivity of Lefty and Righty. Thus I suppose a similar argument could be suggested to show that if we accept that Lefty is identical to Righty after

fusion, we are committed to thinking they are identical before fusion. An 'after fusion' argument would not only be one in which one changed the prefix, but also the occurrences of 'after fission' to 'before fusion' and 'before fission' to 'after fusion' in the content of the premises.

To sum up: What the transitivity complaint states is that the person who is Lefty after fission is identical to the person who is Lefty before fission, and the person who is Lefty before fission is identical to the person who is Righty after fission. But if that conjunction is true, then the person who is Righty after fission is the person who is Lefty after fission. However, on the temporary identity thesis, the antecedent conjunction is only true before fission, and thus that thesis needn't accept that the consequent is true at some other time than before fission.

6.2 RESTRICTING TRANSITIVITY AND LEIBNIZ'S LAW

One might be sceptical about whether we can just restrict transitivity to times the way the temporary identity thesis does. Doesn't a notion of identity according to which transitivity is time-indexed violate the transitivity of identity? That is to ask, can we maintain that temporary identity is really a transitive relation given this restriction?

I find that it is unfair to criticise the temporary identity thesis for not respecting the transitivity of identity when restricting transitivity of identity to times.⁸⁶ There are relations which are obviously transitive but only so when transitivity is time-indexed. Take the case in which Bob and Bill are the same height in 2000. In 2005 Bill is the same height as Pete. Does it follow that Bob is the same height as Pete in 2005? Surely Bill might have grown taller and become the same height as Pete, while in 2005 Bob has the height he had in 2000. But this should not lead us to think that being the same

⁸⁶ See also Gallois 1998, pp. 75-79.

height is not a transitive relation. It is just that it is only transitive at a time. Similarly it can't be correct to claim that indexing transitivity of identity to times is a way not to respect transitivity. It might lead to other problems (which I shall discuss below), but the restriction of the transitivity of identity to times doesn't violate the transitivity of identity as such.⁸⁷

If we only want to accept T-Transitivity we should also modify the standard formulation of Leibniz's Law:

Leibniz's Law $(x)(y)(F)[x = y \rightarrow (Fx \rightarrow Fy)]$.

Leibniz's Law entails T-Transitivity, but T-Transitivity is also entailed by the version of Leibniz's Law which allows for change in properties (including the property of being identical to x):

T-Leibniz's Law $(x)(y)(F)(t)[t: x = y \rightarrow (t: Fx \rightarrow t: Fy)]$.⁸⁸

But, as Perry (1972) points out, it is not obvious that time-indexing avoids the objection that the adherent of the psychological criteria is committed to accept inconsistent statements. We might assume that Lefty wakes up in room 102 after the fission and that Righty wakes up in room 104. Given T-Leibniz's Law, we seem to face the following reductio of the thesis that before fission Lefty is Righty and that this is not the case after fission⁸⁹:

⁸⁷ A relation R is said to be *temporally invariant* 'if it is not possible for x to stand in R to y at some time t unless x stands in R to y at any time at which x and y exist,' Gallois 1998, p. 79. If we take R to be 'the same height as' it clearly isn't temporally invariant. And similarly, psychological unity is not temporally invariant either. Persons A and B may be psychologically unified (i.e. one person) at one time, but not at another time, according to the temporary identity thesis combined with the psychological criterion.

⁸⁸ We can see that transitivity can be derived from Leibniz's Law by considering the following: Let F be the property '= z.' From $x = y$, and x has the property '= z', it follows that y has the property '= z.'

⁸⁹ Cf. Gallois's discussion of 'Slide' and 'Pond', Gallois 1998, p. 81 ff.

- (1) Before fission: Lefty is identical to Righty.
- (2) Not (After fission: Lefty is identical to Righty).

But we also have to accept that:

- (3) After fission: Lefty is in room 102.
- (4) Not (After fission: Righty is in room 102).

But it seems we then face a reductio in that the truth of (3) and (4) are inconsistent with (1), the claim that before fission Lefty is identical with Righty, assuming T-Leibniz's Law. The idea is that (1) can't be true given that Lefty and Righty don't share all their properties in common after fission, and that they don't is entailed by (2), since being identical with Lefty is a property only one of them has. Hence (1) and (2) can't be maintained together, which is what is being claimed by the view presently under consideration. Given (3) it must be correct to state that:

- (5) Before fission: Lefty is in room 102 after fission.

If Lefty is around before fission and will be in room 102 after fission, then it must be correct to ascribe the property of being in room 102 after fission to Lefty before fission. Similarly, if I have the property of being writing a philosophy thesis in 2005, it must also be correct to state that in 2000 I had the property of being writing a philosophy thesis in 2005.⁹⁰

Now from T-Leibniz's Law, (1) and (5) we then get:

⁹⁰ The properties in question are time-indexed properties of the form 'is F at t.' The underlying assumption seems to be that what makes it true for me to have at some time t the property of being F-at-t* is that I am F at t*. And furthermore, as we shall see, that I can't change my time-indexed properties over time.

(6) Before fission: Righty is in room 102 after fission.⁹¹

If Lefty and Righty are identical before fission, they share all their properties at that time. If Lefty before fission has the property of being in room 102 after fission, then before fission Righty also has the property of being in room 102 after fission. But then it seems we must also accept:

(7) After fission: Righty is in room 102.⁹²

Just as I must have the property in 2000 of writing a philosophy thesis in 2005 if I do write a philosophy thesis in 2005 (and exist in 2000), we should accept that if Righty has the property of being in room 102 after fission, then after fission Righty is in room 102. But (7) contradicts (4) which states that it is not the case that after fission Righty is in room 102. On the basis of T-Leibniz's Law we seem to have generated a reductio of the consistency of: (1) before fission Lefty is Righty, (3) after fission Lefty is in 102, and (4) not, after fission Righty is in 102.⁹³ Call (1)-(7) the 'T-Leibniz's Law Argument.' How might the adherent of temporary identity reply to this argument?

6.3 RESPONDING TO THE T-LEIBNIZ'S LAW ARGUMENT (TLL-ARGUMENT)

I think it can be shown that the T-Leibniz's Law argument can be blocked in two places. I will suggest that given a very plausible principle, (7) doesn't follow from (6). It seems plausible to think that we get (5) from (3). Here's another case: If I in 2003 have the property of being in New York, it seems

⁹¹ We can substitute 'Lefty' with 'Righty' in (5) and get (6), since they fall under the same temporal index 'before fission.' As will become clear, substitutivity of identicals is also relative to times, even when the referring expressions are proper names (Gallois 1998, p. 70 ff. thinks of names as *quasi-rigid designators*, Perry (1972) finds them to be temporally non-rigid).

⁹² I see no elegant way to avoid referring back to (1)-(7) in the discussion here. I take some comfort in the fact that Perry, when discussing issues related to fission in 1972 lists no less than 22 claims.

⁹³ Presenting this argument I am indebted to Gallois 1998, pp. 81-82. The problem apparently goes back to Prior 1957.

correct to claim that I have the temporally indexed property of 'being in New York in 2003', at other times than 2003. In 2005 I have the property of being in New York in 2003. It also seems correct to assume that if in 2003 I have the property of being in New York, then everything identical with me at that time, i.e. in 2003, has the property of being in New York in 2003. So assuming that I (SH) am in New York in 2003, the following implication seems true:

(x) (2003: x = SH \rightarrow 2003: x is in New York)

It also seems plausible to think that if I have the property of being in New York in 2003, then at any time at which I exist, say in 2005, I will have that temporally indexed property, and it also seems to be sufficient for my having the property of being in New York 2003 in 2005, that in 2005 there is something which I am identical to and which has the property of being in New York in 2003. As a matter of fact, it seems impossible that it is not so (but I leave out the modal operator in what follows). Hence we get:

(2005: SH is in New York in 2003) \leftrightarrow $(\exists y)$ (2005: SH=y & y is in New York in 2003).

More generally we get the following principle:

E-Principle $(x)(t)(t^*)(F)[\text{at } t: Fxt^* \leftrightarrow (\exists y)(\text{at } t: x=y \ \& \ Fyt^*)]$ ⁹⁴

(5) says that before fission Lefty is in 102 after fission. Now we get (6) from (5) by application of T-Leibniz's Law. Everything identical with Lefty before fission will before fission have the property Lefty has after fission, i.e. the property of being in 102 after fission. But we also get (6) from (5) via the E-Principle:

⁹⁴ Gallois 1998, p. 84. I have omitted the wide scope necessity operator.

- (Before fission: Righty = Lefty & Lefty is in 102 after fission) → Before fission: Righty is in room 102 after fission.

So both Leibniz's Law and the E-Principle will give us (6) from (5). But does (6) give us (7) as argued by the TLL-Argument? I think we get (7) from (6) if we think that in order for me to have in 2005 the property of being in New York in 2003 everything identical with me in 2003 must in 2005 have the property of being in New York in 2003. Now, living in a non-fission world, it is quite plausible to think that at any given time (e.g. 2005) at which I have the property of being in New York in 2003, everything identical with me at that time will have the property of being in New York in 2003. But when contemplating fission cases, it will beg the question against the temporary identity theorist to assume that for me to have the property of being in New York in 2003 at some time t , all the entities with which I am identical at any given time must have the property of being in New York in 2003. I might be identical with Eddy in 2005 due to a fusion which took place after my visit to New York, and Eddy might never have been to New York at all. Nevertheless, the temporary identity theorist is allowed to claim that I may be identical to him in 2005, without him having been in New York in 2003 for the purposes of the *reductio*. The principle which I think we should thus not accept is this:

E-Principle* $(x)(t)(t^*)[(\exists y)(\text{at } t: x=y \ \& \ \text{at } t^*: Fy)] \rightarrow \text{at } t^*: Fx$.⁹⁵

If we accepted that principle, we couldn't even get the TLL-Argument off the ground, because it would be ruled out that we could assume both (1), (3) and (4) to begin with. If we think the E*-Principle is correct, we rule out the temporary identity theorist's description of the fission case as such, but the

⁹⁵ Gallois 1998, p. 87.

reductio argument at hand is supposed to show that even assuming the four claims in the description, we end up with a conclusion, viz. (7), which contradicts (4).

Now the reasoning with which we begin is the observation that everything identical to Lefty after fission is in room 102 (after fission). What we started out with was of this form:

$(x)(\text{after fission: } x=\text{Lefty} \rightarrow \text{after fission: } x \text{ in } 102).$

Considering (6) what we should say given the temporal identity description of fission is that it is sufficient for Righty to be in 102 after fission that everything identical with Righty before fission is in room 102 after fission:

$(x)(\text{before fission: } x=\text{Righty} \rightarrow x \text{ is in } 102 \text{ after fission}) \rightarrow \text{Righty is in } 102 \text{ after fission.}$

But this conditional comes out false on the assumptions of the temporal identity theorist. Not everything identical to Righty before fission is in room 102 after fission. Righty is identical with Righty before fission and Righty is not in room 102 after fission. Rather the principle we should accept is this:

A-Principle* $(x)(t)(t^*)(F)[(y)(\text{at } t: x=y \rightarrow \text{at } t^*: Fy) \rightarrow \text{at } t^*: Fx]$ ⁹⁶

What has been suggested is that it is not sufficient that something identical with Righty before fission is in room 102 after fission for us to get (7). In order for that to be so, everything identical with Righty before fission will have to be in 102 after fission, and this is not so.

⁹⁶ Gallois 1998, p. 87.

But now one might ask whether a similar complaint can't be made with respect to the claim that before fission Lefty has the property of being in 102 after fission? After fission Lefty is identical to something with the property of being in room 102 after fission. Isn't it mistaken to think that this fact isn't sufficient for Lefty having that property at some other time?

It might be suggested that everything identical with Lefty before fission must be in room 102 after fission for it to be the case that before fission Lefty is in room 102 after fission. So perhaps the E-Principle should be substituted with:

A-Principle $(x)(t)(t^*)(F)[\text{at } t: t^*:Fx \leftrightarrow (y)(t: x=y \rightarrow t^*: Fy)]$

However, while it might be tempting to adopt the A-Principle as opposed to the E-Principle, this will not help the proponent of the TLL-Argument much. The A-Principle entails that given the assumptions (1)-(4) of that argument, we don't get (5) from (3). While (3) comes out true on the A-Principle:

- after fission: after fission: Lefty is in room 102 \leftrightarrow (x)(after fission $x=Lefty \rightarrow$ after fission: x is in 102),

the claim in (5) will be false since:

- before fission: after fission: Lefty is in room 102 \leftrightarrow (x)(before fission: $x=Lefty \rightarrow$ after fission x is in room 102).

But this is false since there is something, Righty, which is identical with Lefty before fission and not in room 102 after fission. As a matter of fact, the temporary identity theorist gives a description of the fission case which allows him to explain how, given the A-Principle, (5) can be false and (3) true. But then while one might disagree as to whether (A) or (E) is the better principle for stating the necessary and sufficient conditions for the truth of

claims of the form 'at t: x is F at t*', it will be possible for a temporary identity theorist to block the T-Leibniz's Law Argument by denying either that (3) implies (5) or that (6) implies (7).⁹⁷

It is important to appreciate that we are trying to describe a situation in which there is fission, and the temporary identity theorist is suggesting what to say about it. It is, I think, quite clear that allowing for fission scenarios to be described in terms of temporary identity, we can't assume that it is sufficient for me to be in New York tomorrow, that today I am identical to some x and x is in New York tomorrow. The temporary identity analysis is exactly the view that this is not sufficient if fission may occur between today and tomorrow. On the other hand, I think it would be rather odd given the temporary identity description to claim that in case I was not in New York tomorrow, someone who is today identical with me, and who is in New York tomorrow, is not correctly said today to have the property of being in New York tomorrow. Again this seems to me to go against the whole spirit of the temporary identity thesis.

6.4 CAN THINGS CHANGE THEIR TIME-INDEXED PROPERTIES?

If we deny the move from (6) to (7), which is the option of Gallois and Perry, we must also reject that a thing can't change its time-indexed properties over time. Righty is said to have the time-indexed property of being in room 102 after fission, before fission, but not after fission. This goes against the view that if things have time-indexed properties, and the temporary identity theorist seems to have to allow for this, they can't change those properties. This means that the temporary identity theorist will have to reject the following principle assumed by the T-Leibniz's Law objection:

Transmission (x)(t)(t*) {at t: Ex and at t* Ex → [at t*:Fxt* ↔ (at t: Fxt*)]}⁹⁸

⁹⁷ This is pointed out by Gallois 1998, p. 90. Gallois acknowledges Graham Priest for this observation.

This principle states that if some thing exists at different times t and t^* , then it has a property F -at- t at t if and only if it has the property F -at- t at the other time t^* . Transmission expresses the general metaphysical thesis that things don't change their time-indexed properties. If I have the property of being writing a philosophy thesis in 2005, there doesn't seem to be any time at which I exist, where I don't have the property of being writing a philosophy thesis in 2005. However, a counterexample to Transmission is the truth of

(6) Before fission: Righty is in room 102 after fission

and the falsity of:

(7) After fission: Righty is in 102 after fission.

(7) says that after fission Righty has the time-indexed property of being in 102 after fission. So if (7) is false and (6) is true, then Righty has changed a time-indexed property. Before, but not after, fission, he had the time-indexed property of being in room 102 after fission, before fission, but not after fission.

Now the E-Principle says that:

E-Principle $(x)(t)(t^*)(F)[\text{at } t: Fxt^* \leftrightarrow (\exists y)(\text{at } t: x=y \ \& \ \text{at } t^*: Fy)]$.

On the temporary identity thesis and the E-Principle, it is enough for me in 2000 to have the property of writing a philosophy thesis in 2005 that someone writing a philosophy thesis in 2005 is identical with me in 2000. Of course in the case of fission, we might suggest that I am Righty and that in

⁹⁸ Cf. Gallois 1998, p. 96. 'Ex' should be taken to mean 'x exists.'

2005 Lefty is the person writing a philosophy thesis. In that case I would not be writing a philosophy thesis in 2005 even though this was true of me in 2000. Why? Well, due to the fact that in 2000 I was also identical to someone who is not writing a philosophy thesis in 2005. Not everything with which I was identical in 2000 (in case I am the product of fission) has the property of writing a philosophy thesis in 2005. I was identical to myself at that time, and *ex hypothesis* I am not writing a philosophy thesis in 2005.

This might lead us to strengthen the conditions for me to have a temporal property F-at-t, so that everything identical to me at different times (including t) must be F-at-t. This is what is expressed in:

A-Principle $(x)(t)(t^*)(F)[\text{at } t: Fxt^* \leftrightarrow (y)(\text{at } t: y=x \rightarrow \text{at } t^*: Fy)]$.⁹⁹

In the case were we assume that I am the product of someone's fission between 2000 and 2005, and I am not writing a philosophy thesis in 2005, but the other fission product is, the A-Principle entails that it would be false to ascribe the time-indexed property to the prefission person with whom the other fission product writing a philosophy thesis in 2005 was identical. But this again entails that something has changed its time-indexed properties. In case we accept the A-Principle, it turns out that the other fission product with the philosophical proclivities in 2005 has a time-indexed property he didn't have in 2000, namely the property of writing a philosophy thesis in 2005.

If we accept either the E-Principle or the A-Principle as an account of the truth-conditions for statements ascribing time-indexed properties, and we also adhere to the temporary identity thesis, we must reject that objects can't change time-indexed properties. However, why this is so on the temporary

⁹⁹ Gallois 1998, p. 89.

identity thesis is explained on the basis of a plausible principle about what it takes for something to have a time-indexed property at different times.

6.5 REPLACEMENT AND THE ONLY X AND Y PRINCIPLE¹⁰⁰

I have already presented arguments for why I don't think it is attractive to describe fission as a case in which two new distinct persons replace the original person. One thing I find unsatisfactory about this proposal is that the relation, which normally secures personal identity according to the psychological criterion, holds between Brown and each of the fission products. If we accept that this relation does secure personal identity in non-branching cases, we must explain what it is about the situation after fission which implies that neither Lefty nor Righty is identical with Brown. Why is there nobody after fission of whom we can truly say that he is Brown, when there is some such person before fission and the relation securing Brown's identity over time holds between him and at least one future individual?

On the closest continuer view in one of its shapes, the one I have been discussing (chapter 5), Righty and Lefty are never identical to Brown, but the existence of each depends on that of the other, since they never exist in a world where there isn't fission. On another version, Righty and Lefty are identical to Brown in some possible worlds, and it is the existence of the other in fission worlds, which makes it the case that Brown isn't identical to either of them.

The temporary identity thesis I think provides an explanation of why Brown can't correctly be said to be identical to either of the fission products after fission, i.e. of why there is a sense in which the relation that secures Brown's survival doesn't do so in fission scenarios. So how does the temporary identity thesis explain that Brown is Righty at one time and not at some other

¹⁰⁰ This section draws on Gallois's discussion in his 1998, pp. 111-113.

time at which Righty exists? Some explanation must be provided for this contingency claim about the identity of Righty and Brown.

The reason why *after* fission Righty can correctly be said to be identical with Brown *before* fission is that not just something but everything that is identical with Righty after fission is identical with Brown before fission. However, not everything identical with Brown before fission is identical with Righty after fission. E.g. Lefty is identical with Brown before fission, but not with Righty after fission. This is of course analogous to the way temporary identity theorists handle other property ascriptions. For it to be true before fission that Brown is in room 102 after fission, something identical with Brown before fission must be in room 102 after fission. On the other hand, for Brown to be at a given time *t* in, say, room 100, everything identical with Brown at *t*, must be in room 100.

I think it is worth pointing out that given the temporary identity thesis, there is a sense in which we can maintain that Righty and Lefty have a past before fission. In the case of Righty, we can say that since there is something with which Righty is identical before fission and everything identical with Righty after fission is identical with that something (Brown) at that past time, Righty was around before fission (analogously for Lefty). However, there is a sense in which we will also say that Brown doesn't have a future, since there is no person after fission with whom every person now identical to Brown is identical after fission. When we are tempted to say, 'had it not been for Lefty, Righty would have been Brown', we mean to say that in that case everything identical with Brown before fission would have been identical with Righty after fission, and thus Righty would have been Brown after fission.

How does all this square with the only *x* and *y* principle? The only *x* and *y* principle is roughly that the identity of *x* and *y* should not depend on facts about other entities than *x* and *y* and their interrelations. So the complaint is

that when we say that Lefty isn't identical to Brown after fission this is because Righty exists after fission.

As I see it the temporary identity thesis renders the objection that it doesn't respect the only x and y principle innocuous, since accepting temporary identity involves a certain account of what it takes for the fact that Brown is identical with Lefty at one time to entail that he is identical with Lefty at some other time. For Brown's identity with Lefty before fission to entail that Brown is also identical with Lefty after fission, it is not enough that there is something which is identical to both Brown before fission and also identical to Lefty after fission.

Rather, on the temporary identity thesis, we will claim that Brown's identity with Lefty before fission entails that Brown is identical with Lefty after fission only if *everything* identical with Brown *before* fission is identical with Lefty after fission.

Transmission of Identity (x)(y)(t)(t*)[(z)(at t: z=x → at t*: z=y) → (at t*: x=y)]¹⁰¹

The explanation of why Brown isn't identical with Lefty after fission is that in the fission world, there exists a person (Righty) after fission, who is identical to Lefty at some other time, but not identical to Lefty and Brown after fission. Now the fact that the temporary identity thesis can explain why Brown isn't Lefty after fission on the basis of this principle doesn't as such avoid violating the only x and y principle, since it may still be said that it is the existence of Righty after fission, which rules out Brown's identity with

¹⁰¹ Gallois, p. 112. This principle is of course just one instance of the transmission of time-indexed properties in general. Thus for Brown to have the property of being in room 102 after fission, everything identical to him before fission must be in room 102 after fission. Again, this is not so, since Righty is identical to Brown before fission and not in room 102 after fission.

Lefty after fission. However, I think the complaint has much less force against the temporary identity view.

First the explanation of why Brown isn't identical to Lefty after fission is based on the general and plausible account of what it takes for there to be transmission of a time-indexed property like identity. Secondly, if we keep in mind that the principle entails that Lefty *was* identical with Brown before fission, and thus that we aren't forced to conclude from the fact that there is fission that Righty's existence after fission entails that Lefty, so to speak, loses his past, I think the violation of the only *x* and *y* principle becomes innocuous. It is one thing to say that after fission, Lefty isn't identical to Brown, because after fission there exists a person, who isn't identical to Lefty, and who was identical to both Lefty and Righty. It's another thing to say that due to the existence after fission of Lefty and Righty respectively, neither of the two is now and was ever identical to Brown. We can maintain that *every* person that was identical to Brown exist after fission, but after fission, neither of those persons are identical to Brown. But if we can maintain that every person that was identical with Brown still exists, then within the temporary identity framework it seems immaterial that neither of them has the property of being identical to Brown.

6.6 THE REOCCURRENCE OF THE TRANSITIVITY ARGUMENT AND TRANS-TEMPORAL IDENTITY¹⁰²

Some relations hold trans-temporally. E.g. the relation 'is the same height as' seems to do so. Suppose I am the same height now as Socrates was when he was 30. E.g. I am now 180 centimetres tall and so we might say was Socrates when he was 30. Given that I am 180 centimetres now and Socrates was 180 centimetres then, I can truly be said to be the same height now as Socrates was then, and this is true despite the fact that at no time *t* is it the case that

¹⁰² Writing this section I am indebted to Gallois 1998, p. 113 ff.

both I and Socrates are 180 centimetres tall. Thus it is characteristic of a relation which may hold transtemporally, such as 'is the same height as', that individuals x and y can be the same height and thus stand in the 'same height' relation, without them ever being the same height simultaneously.

We might now think that the transitivity argument can be revived in this way:

- The thing which is Lefty after fission = the thing which is Righty before fission.
- The thing which is Righty before fission = the thing which is Righty after fission.

If we interpret these two statements as involving trans-temporal identity, i.e. if we take the identity relation to hold trans-temporally, we may think we get the conclusion that :

- After fission Lefty = Righty.

The argument would go through if we substituted '=' with 'is the same height as.' The difference between an identity statement and a sameness of height statement is not that it makes sense to ask when the statement is supposed to be true with respect to the one but not the other. Taking Socrates' height to be fixed, we might ask '*when* was SH the same height as Socrates?' This makes perfect sense, just as it makes perfect sense on the temporary identity view to ask '*when* was Lefty identical to Righty?' The difference is that when we indicate the time at which SH was the same height as Socrates, that need not be a time at which there is something which is Socrates and 180 centimetres tall and something that is SH and is 180 centimetres tall. Nothing is Socrates at the time in question and Socrates doesn't have any height at all at the time at which SH is the same height as

Socrates. The relation holds *trans*-temporally between SH and Socrates at a time.

On the other hand, when the relation is *identity*, it is clear that not only can we ask *when* the identity relation holds between Lefty and Righty, but when we are told at what time it holds, then *at that time* there must be something which is identical to Righty and identical to Lefty. This makes it clear that the relation of identity doesn't hold *trans*-temporally on the temporary identity thesis. This is an important implication of accepting the temporary identity thesis.

Now it is worth pointing out that there is a relation which does hold *trans*-temporally and which it seems is very much like identity. It is not identity however, since identity is identity at a time on the view under consideration, but it is what might be called 'trans-temporal identity.' The relation of trans-temporal identity can be expressed in the following principle:

Trans-Temporal Identity $(x)(y)[x T= y \leftrightarrow (\exists t)(\exists t^*)(z)(\text{at } t: z = x \rightarrow \text{at } t^*: z = y)]$.¹⁰³

The T=-relation is not an equivalence relation. It is not symmetrical. Take the case of fission:

There is clearly a time at which for all entities that are identical with Lefty it is the case that they are identical with Brown at some other time. This is the situation after fission. So Lefty after fission is *trans*-temporally identical with Brown. However, the claim that everything identical with Brown before fission is identical with Lefty after fission is not true both before and after fission. Thus Brown isn't *trans*-temporally identical with Lefty.

¹⁰³ Gallois 1998, p. 116.

This brings out the claim that we can say that Lefty has a past because he is trans-temporally identical with Brown, while Brown doesn't have a future due to the fact that Brown isn't trans-temporally identical to Lefty. I think the machinery provided by the temporary identity thesis for making this claim makes it quite attractive for someone who thinks that personal identity can be analysed in terms of a relation which may fission.

6.7 WHAT HAPPENS TO BROWN?

A question which is worth asking at this point is, 'What happens to Brown?' After all, the one and only person we pick out before fission, we call 'Brown,' and the two persons we pick out after fission, we call 'Lefty' and 'Righty'; but what then happens to Brown?

On what I have called the replacement thesis, Brown ceases to exist and this entails that Lefty and Righty should be taken to be new persons who start to exist instead of Brown.

Another view is one which subscribes to the following set of claims:

- Before fission: Lefty = Righty = Brown = the person in room 100.
- After fission: Lefty = the person in room 102.
- After fission: Righty = the person in room 104.
- After fission: it is not the case that there is an x & $x = \text{Brown}$.

This 'third man' view agrees with the replacement thesis that Brown ceases to exist, but doesn't agree that it entails that Lefty and Righty weren't existing before fission. Before fission we say that Brown is identical to Lefty and Righty and that Lefty and Righty are identical. After fission Brown is not identical to anyone. But if Brown isn't identical to anyone after fission that may seem to indicate that he has ceased to exist and that before fission, there

was 'a third man', Brown, who is nowhere to be found after fission. But this 'third man' view isn't much more appealing than the replacement thesis. It commits the temporary identity thesis to the claim that there are three person before the fission, one of whom doesn't survive, but a reason to adopt the temporary identity thesis was to avoid having to say that fission is a case in which a person ceases to exist.

Here I think it is important to maintain that there is a third alternative according to which no person ceases to exist in fission, and according to which fission involves exactly two persons, Lefty and Righty, who are identical at one time and distinct at another.

If we think the case is one which involves exactly two persons, Lefty and Righty, we'll say that 'Lefty' names the person in room 100 before fission and in room 102 after fission, because that is what we stipulate, and similarly for 'Righty.' But why is it then not open to say or stipulate that by 'Brown' we refer to the person before fission and leave it open whether Brown is identical with Lefty or Righty after fission?

Here I think it can be pointed out that it doesn't give us a complete description of the use of the name 'Brown' in the fission case to leave it open, which of the persons it denotes. Does it denote the one who is in 102 after fission or the one who is in 104 after fission, when applied before fission to the person in 100? In case we rule out that there is a third person referred to by 'Brown' and maintain that there are exactly two persons involved in the case, we must make clear, whether we use 'Brown' to refer to the one or the other, when introducing the name to denote the person in 100 before fission. I.e. it seems that we must, when introducing a name for the prefission person, make clear whether it is intended to pick out one or the other fission product. We can't name the person prefission without stipulating which of the two survivors it picks out on this view. Let me present a difficulty with

this view and then move on to consider an alternative to both the view that 'Brown' picks out a third man before fission and the view that it must be stipulated to pick out either one or the other fission offshoot.

There are difficulties with the account which stipulates 'Brown' to refer to one or the other fission offshoot, even if it's more attractive than the one which involves three persons. Presumably we might name a person without knowing that fission is going to take place. But this seems to be ruled out by the suggestion at hand. It is of course open to claim that it isn't always possible to name an individual with which one is confronted without knowing about its future. Still that surely seems to be a rather unpalatable suggestion. Even if we are ready to accept that there are cases in which we are not able to name something not knowing about the individual's future, there is also another problem with the view that there are exactly two persons in the fission scenario which are identical before and nonidentical after the fission.¹⁰⁴

If we want to maintain that fission is a case involving exactly two persons, we face the following challenge. When fixing the reference of 'Brown' before fission, we use the expression 'the person in room 100 before fission'¹⁰⁵ to pick someone out as the referent of 'Brown.' The problem which Gallois points out is this: Which of the two following statements is true if any of them is?

(D1) After fission: the person in 100 before fission = Lefty

(D2) After fission: the person in 100 before fission = Righty

¹⁰⁴ Here I follow Gallois 1998, p. 105.

¹⁰⁵ I use the expressions 'before fission' and 'after fission' as temporal indexes analogous to 't1' and 't2.' They should not be taken to signify that one knows before fission that fission will take place. I.e. Brown could think that he is in room 100 'before fission' without being aware that fission will take place. He'd think something like 'I'm in room 100 now' or 'at t1.'

After fission there doesn't seem to be any reason for taking one but not the other of the two persons Lefty and Righty to satisfy the description. However, if we accept that both descriptions are true, we end up saying that after fission Lefty is identical with Righty. If it is true of Lefty after fission that he satisfies the description and true of Righty that he satisfies the description at the same time, then given that the description only picks out a single person, Lefty and Righty must be identical after fission. So it seems it must be claimed that both are false. However, if both (D1) and (D2) are false, then it does seem as if there is a third person before fission who doesn't exist after fission. We were supposed to avoid this.

However, Gallois points out that the adherent of the temporary identity thesis needn't give up on the claim that there are exactly two persons involved in the fission case, even if he rejects the truth of the two statements (D1) and (D2).

Here's a third statement we can compare with the two above:

(D3) Before fission: the person in room 100 before fission = Lefty¹⁰⁶

What sort of truth conditions should we assign to statements of this form? It seems fair to suggest the following analysis of the truth conditions of (D3):

(D4) $\exists x[x \text{ is a person in room 100 before fission} \ \& \ (y)(y \text{ is a person in room 100 before fission} \rightarrow x = y) \ \& \ \text{before fission: } x = \text{Lefty}]$

Of course on the temporary identity thesis we must time index all identities, so (D4) is better rendered as:

¹⁰⁶ I assume that 'in room 100' denotes a place occupied by exactly one person, just as we might say 'the person at place p1.'

(D5) $\exists x[x \text{ is a person in room 100 before fission} \ \& \ (y)(y \text{ is a person in room 100 before fission} \rightarrow \text{before fission: } x = y) \ \& \ \text{before fission: } x = \text{Lefty}]$

With (D5) in hand we can avoid having to give up on there being exactly two persons in fission while rejecting (D1) and (D2). (D1) is of the same form as (D3) but has different temporal indexes. Thus (D1) becomes this:

(D6) $\exists x[x \text{ is a person in room 100 before fission} \ \& \ (y)(y \text{ is a person in room 100 before fission} \rightarrow \text{after fission: } x = y) \ \& \ \text{after fission: } x = \text{Lefty}]$

But (D6) is false. It says that there is something which is a person in room 100 before fission and that everything which is a person in room 100 before fission is identical with Lefty after fission. But that is not true. Righty is a person in room 100 before fission, but is not identical with Lefty after fission. The same can be said about (D2). What has been pointed out is that for (D1) and (D2) to be true, it must be the case that after fission there is only one person who satisfies the description before fission. But after fission there are two such persons according to the temporary identity thesis. Hence that thesis need not admit three persons if it denies the truth of (D1) and (D2).

We still need to give an account of the use of a name like 'Brown' to name a person with whom we are confronted, when we don't know he will fission in the future. The principle used to reject (D1) and (D2) is this:

(D) (t)[At t: the Fx is G if and only if something is F, everything which is F is identical with x at t, and x is G at t].¹⁰⁷

When the fission has occurred I might ask myself, 'Is Lefty the person I named 'Brown' before the fission?' Thus I ask for the truth conditions of:

¹⁰⁷ This is my rendering of Gallois's (D') 1998, p. 106. My discussion follows his on p. 105 ff.

(D7) After fission: the person I named 'Brown' in room 100 before fission has the property of being identical to Lefty

Given (D), (D7) is true if and only if

(D8) Lefty satisfies 'is the person I named 'Brown' in room 100 before fission', and every person which does satisfy that description is identical with Lefty after fission.

But (D8) is false, due to the falsity of the second conjunct, so (D7) is false. What all this entails is that it is not correct to apply the name 'Brown' to Lefty after fission, because Lefty is not the only person after fission satisfying the description 'the person I named 'Brown' in room 100 before fission.' However, the inapplicability doesn't entail that there are not exactly two persons involved. Rather it is a case in which each of them at one time, before fission, satisfied the condition for being referred to as 'Brown', whereas neither of them do so after fission.¹⁰⁸

6.8 ON PERRY'S LIFETIME VIEW

Fission may be depicted as a Y-shaped figure in which the lower line is denoted by 'Brown,' while the upper branches are denoted 'Lefty' and 'Righty' respectively, with Lefty being the left hand branch.

Perry has suggested that the move from (6) before fission: Righty is in room 102 after fission, to (7) after fission: Righty is in room 102 after fission, will not go through if we take the names to pick out lifetimes or sets of person-stages. The names 'Brown', 'Lefty' and 'Righty' are introduced as applied to a person-stage *s*. A lifetime is the set of all and only the person-stages R-

¹⁰⁸ This is parallel to the claim that whereas they each satisfy the condition for being identical with Brown at one time, neither do so at another time (after fission).

related to a person-stage s picked out at some time t . When we at a time pick out a person-stage s naming it e.g. 'Brown' the person-stage will determine a lifetime, where this means that it picks out the maximal set of person-stages R-related to s . The name we give to s has as its primary referent the lifetime determined by the stage it picks out.

So on this view fission is a situation in which there are three lifetimes which are determined, namely the lifetimes determined by the three stages referred to respectively as 'Lefty', 'Righty', and 'Brown.' The Y-shaped lifetime is determined by 'Brown' which picks out the stage before fission, and the two lifetimes, 'left-branch lifetime' and 'right-branch lifetime,' are determined by the stages to which 'Lefty' and 'Righty' are applied after fission. If we take the prefission stage picked out as 'Brown,' we can trace R-related person-stages along both branches of the Y-shape, and thus 'Brown' determines the whole Y-shaped lifetime. On the other hand, when we pick out Lefty at the top of the left branch of the Y-shape, we trace person-stages back to before fission, but not to the right branch of the Y-shape. Hence, the lifetime determined by Lefty is the line going from the top of the left branch down to the bottom of the Y-shape. Similarly for Righty.

The lifetime determined by the stage to which 'Lefty' is applied Perry calls the name 'Lefty's *primary referent*. Perry now stipulates that names have not only a primary referent, viz. the lifetime determined by the stage to which the name is given. They also have a *secondary referent at a given time t* . Take the lifetime determined by the stage we picked out as Lefty. That lifetime has a stage occurring before fission as a member. Now while that stage doesn't determine the lifetime determined by Lefty, it determines another lifetime, the Y-shaped lifetime, and this means, Perry stipulates, that the name 'Lefty' has the Y-shaped lifetime as its secondary referent before fission. The secondary referent of a name at t is the lifetime *uniquely* determined by the

stage at t , which is a member of the lifetime which serves as the primary referent of that name.

Take the name 'Lefty.' 'Lefty' picks out a person-stage after fission. That stage determines a lifetime, 'the left-branch lifetime.' Thus for the name 'Lefty' the primary referent is the lifetime determined by the stage picked out as Lefty after fission. Now that lifetime contains person-stages at other times. E.g. it contains a person-stage before fission called 'Brown,' which itself determines the Y-shaped lifetime and thus has the Y-shaped lifetime as its primary referent. Perry stipulates that this means that 'Lefty' has a 'secondary referent' before fission, namely the lifetime determined by the stage occurring before fission in the lifetime determined by Lefty after fission, so before fission Lefty's secondary referent is the Y-shaped lifetime.

In fission there is no time at which the lifetimes determined by the stages we call 'Lefty' and 'Righty' contain exactly one person-stage. Hence it is always clear what the names 'Lefty' and 'Righty' refer to at different times. Whether we pick out the lifetime determined by Lefty before or after fission, it contains a single person-stage at that time, and hence it is determinate that 'Lefty,' when uttered at that time, will unambiguously pick out a single person-stage.

On the other hand, the name 'Brown' picks out a single person-stage before fission which determines the Y-shaped lifetime. Before fission the Y-shaped lifetime is both the primary and secondary referent of 'Brown.' However, after fission there isn't a unique lifetime determined by the person-stage of Brown's lifetime occurring after fission, since there are two such person-stages, those named 'Lefty' and 'Righty.' But then, Perry concludes, the name 'Brown' is *improper* at that time. What this means is that while 'Brown' has a secondary referent before fission, it doesn't have one after fission. However,

both 'Lefty' and 'Righty' have a secondary referent before fission, namely the lifetime determined by Brown.

Now Perry stipulates the following truth conditions for property ascriptions to persons:

'at t^* : N is F-at- t ' is true if and only if the secondary referent of N at t^* has a stage s as a member which is F at t .

' t : N is F' is true if and only if the secondary referent of N at t contains a stage s which is F at t .

We can now see how Perry's account may block the move from (6) to (7). For (6) to be true it must be the case that the secondary referent of 'Righty' before fission has a stage as a member which is in room 102 after fission. And (6) is true, since the Y-shaped lifetime is the secondary referent of 'Righty' before fission and it contains a stage after fission which is in 102.

But the fact that (6) is true doesn't entail that (7) is true. (7) turns out to be false. The truth condition for 'after fission Righty is in 102' is this: the secondary referent of 'Righty' after fission must contain a stage which is in room 102 after fission. But the secondary referent of 'Righty' after fission is the right-leg-lifetime, and that lifetime doesn't contain a stage which is in room 102 after fission. Hence, (7) is false.

On Perry's account of the truth conditions of temporally indexed property ascriptions to persons, the reason why we can't infer (7) from (6) is analogous to the reason why we can't infer from:

- On Monday: the senator from California will be in San Francisco on Tuesday,

to

- On Tuesday: the senator from California is in San Francisco

The expression 'the senator from California' may be proper at one time and not at another. We can imagine that one of the two senators of California steps down on Friday and a replacement is not appointed till Tuesday. On Monday, when 'the senator of California' picks out a single person, who as a matter of fact turns up in San Francisco on Tuesday, it is true to say that 'the senator from California will be in San Francisco on Tuesday. But on Tuesday, when there are two senators, the expression 'the senator from California,' doesn't uniquely refer, and thus the statement 'the senator from California is in San Francisco' doesn't come out true on Tuesday.

In the case at hand, due to 'Righty' not being a temporally rigid designator, but functioning as a temporally nonrigid designator, 'Righty' has distinct lifetimes as its secondary referent before and after fission, so (6) may be true and (7) false.

On Perry's view, 'Brown' picks out a single person or lifetime before fission, namely the Y-shaped lifetime. The left-branch lifetime and the right-branch lifetime are not determinable before fission, since the person-stages which determine them are distinct ones picked out after fission, and so they are never determined by Brown, nor are they at any time identical with the Y-shaped lifetime or with each other. In fission, while the names we use are temporally nonrigid, there is no time at which they denote the same lifetime or person according to Perry's view. However, this seems to disqualify

Perry's view as one according to which fission is a case best described as one in which two things were at one time identical.¹⁰⁹

On Gallois's account the move from (6) to (7) fails, not because 'Righty' is temporally nonrigid and designates distinct lifetimes before and after fission, but due to the fact that the very same person, named 'Righty' both before fission and after fission, before fission satisfies a necessary and sufficient condition for being in room 102 after fission, but after fission doesn't satisfy a necessary condition for being in room 102 after fission.

6.9 SHORT REMARK ON CONTINGENT IDENTITY

Accepting the temporary identity thesis also commits one to accept the thesis that identity is contingent, a contention which is controversial, but which I will not discuss at any length in this thesis. Here I'll make a short remark about how one can defend the contingent identity thesis against the objection that it doesn't respect Leibniz's Law in a way that parallels the defence of temporary identity against that line of criticism.

According to the contingent identity thesis, identical individuals are possibly not identical:

Contingent Identity Thesis $(\exists x)(\exists y)[x = y \ \& \ \diamond \text{ not } (x = y)]$ ¹¹⁰

This thesis is entailed by the temporary identity thesis, but just as one might reasonably think that A is identical to itself at all times at which A exists, one might reasonably think that A must be identical to itself at all possible

¹⁰⁹ McIntyre (1978) has some interesting criticism of Perry's account which I find to be accommodated in the temporary identity thesis I have argued for allying myself with Gallois 1998. It seems to me to be a weakness about McIntyre's view that she has to recognise some third entity named 'Brown' before fission which is characterised as 'a specifiable portion of a personal set' which is not itself a person (cf. p. 457).

¹¹⁰ See Gallois (1998, p. 142) for this rendering of the principle.

worlds at which it exists. So it is reasonable to take A to have the modal property 'is necessarily identical to A.'

In the discussion of temporary identity, we saw that Leibniz's Law should be relativised to times in order not to beg any questions against the temporary identity thesis. In a similar way, we should relativise Leibniz's Law to worlds in order not to beg any questions against the contingent identity thesis:

M-Leibniz's Law $(x)(y)(w)[\text{in } w: x = y \rightarrow (\text{in } w: Fx \rightarrow \text{in } w: Fy)]$ ¹¹¹

Now the contingent identity thesis can't be said to violate M-Leibniz's Law for reasons similar to those presented with respect to the Leibniz's Law argument against temporary identity.

Let's say that the following statements are true:

- At w : $A = B$.
- At w : A has the property of being necessarily identical with A.

From this it follows that:

- At w : B has the property of being necessarily identical with A.

However, a proponent of the contingent identity thesis can say that if we accept the contingency of identity, substitution of identicals is only legitimate when relativised to the world in which the identity holds, just as substitutivity of identicals is only legitimate relative to the time at which the identity holds. And this means that just as:

¹¹¹ See Gallois (1998, p. 145) for this rendering of the principle.

$(t)(t^*)[(t: A = B \ \& \ t^*: A \text{ is } F) \rightarrow (t^*: B \text{ is } F)] \rightarrow (t = t^*),$

we should maintain that on the contingent identity thesis:

$(w)(w^*)[(w: A = B \ \& \ w^*: A \text{ is } F) \rightarrow (w^*: B \text{ is } F)] \rightarrow (w = w^*).$

So there is no violation of the very plausible claim that at all possible worlds at which A exists, A is identical with itself.

6.10 CONCLUSION

In this chapter I have discussed the prospect of adopting a description of fission, which invokes the claim that identity is temporary as opposed to eternal and necessary. I find this to be an attractive view for proponents of the psychological criterion working within an endurantist framework, because it allows one to maintain that multiplication of the process, which according to the psychological criterion secures a person's survival, doesn't entail that the person ceases to exist. I think the temporary identity thesis provides an albeit revisionary but consistent description of the case, which takes seriously that there is one person to begin with and two distinct persons in the end, and that fission is not a case in which a person ceases to exist. I claim that fission is a case in which the person existing before fission, Brown, can *at that time* truly say of himself that he will exist after fission, since his psychological life is continued after fission. And I maintain that both persons existing after fission, Lefty and Righty, can *at that time* truly say of themselves that they existed before fission, since their psychological life stretches back to before fission.

The temporary identity thesis is a controversial thesis and I have not presented a complete defence for it. To do so would require an assessment of its potential for handling a series of puzzles about persistence and change, and a more careful consideration of the denial of the necessity of identity

than I am able to provide here. Still I think temporary identity has been underestimated as a way of dealing with the problem of fission in the discussion of personal identity and as I think will become clear in chapter 10, adopting the temporary identity thesis also allows for interesting ways of dealing with the other objections raised against the psychological approach to personal persistence.

7. Introducing the Problem of Multiple-Kind Fission

In the last two chapters I have defended the claim that persons have psychological persistence conditions against the objection that the psychological criterion allows for fission scenarios, which arguably commits the proponent of that criterion to inconsistent claims.

In chapter 5 I criticised two responses to the fission problem: the closest continuer and the multiple occupancy views. On the closest continuer view, fission is a case in which the original person ceases to exist and is replaced by two distinct persons coming into existence. I argued that we should reject this response mainly due to its violation of the only x and y principle.

On the multiple occupancy view, the response to fission is to describe it as a case in which there were really two distinct persons all along. I argued that there are a number of unfortunate consequences of this view: one was that what we ordinarily count as one person could turn out to be any number of numerically distinct persons, and I also noted that it seems odd to argue in favour of the claim that diachronic psychological unity constitutes the life of a single person, and then deny that synchronic psychological unity constitutes a single person to *defend* diachronic identity.

In chapter 6 I defended the view that we should accept the temporary identity thesis, if we want to take seriously the view that psychological continuity constitutes identity. I showed how the temporary identity thesis is a coherent thesis, which allows us to maintain that each of the fission offshoots, Lefty and Righty, existed before the fission due to their psychological unity with the pre-fission person Brown.

In the final chapters of this thesis I will consider another problem which confronts a proponent of the psychological criterion. When it is claimed that a person's identity over time consists in psychological continuity, it seems there are both ordinary and extraordinary scenarios in which the life of a person, a person's history, comes apart from the history of the human organisms, which paradigm persons are plausibly thought to be identical with. Persons such as you and I are *human* persons, as opposed to robotic or angelic persons, in that we belong to the species *homo sapiens* forming part of biological taxonomy.

The multiple-kind fission problem is parallel to the standard fission problem discussed in the previous chapters in that the cases are cases in which we would seem committed to say that one individual becomes two distinct individuals, except that in the case where the histories of a human organism and a person diverge, the significant kinds of continuity relations in play are different. In 'multiple-kind' fission we have a human person whose psychological continuer is not identical with his biological or organic continuer.

Take a human person called 'Holmes.' To begin with it is, I think, fair to assume that Holmes is both a person and a human organism. All persons we know of are organisms of the species *homo sapiens*, and I think it is fair to assume that, within a materialist framework, the view that some persons are not human organisms will be the controversial one having to be argued for, not the other way around. So Holmes is, as we will ordinarily say, a *human person*, satisfying both the concept of a person and of a human organism. Holmes has all the psychological properties it takes to satisfy the concept of a person, he is rational and self-conscious etc. And he also has the biological properties characteristic of human organisms in that his internal physical structure is such that he manifests all the characteristic activities of a biological organism of the human kind.

As it happens Holmes is about to undergo a BST procedure this evening. The result of the BST is that tomorrow there will be a *bionic* non-human person, call him 'Hannibal', who is psychologically continuous with Holmes; and there will be a mindless human animal, call it 'Hanimal', which is biologically alive and biologically (and thus spatiotemporally) continuous with Holmes.

I have already explained why I think it is worth defending the view that Holmes is identical with Hannibal when defending my reasons for accepting the psychological criterion. However in the case at hand there is an organism, Hanimal, which is Holmes's unique biological continuer. Hanimal is digesting the food Holmes ate before the BST, Hanimal's immune system is fighting to overcome the flu that Holmes was complaining about yesterday, and so on. Basically the relation between Hanimal and Holmes is much like the relation between a mindless (and thus non-personal) human vegetable resulting from the changes happening to a human person, who for some reason or other ceases to realise personhood. Just as I think it is correct to think that Hannibal is Holmes and was around before the multiple-kind fission due to their psychological relatedness, I find it hard to accept that Hanimal isn't Holmes and that Hanimal wasn't around before the BST took place. Given Hanimal's biological continuity relation to Holmes, I find it hard to accept that Hanimal wasn't around before the BST, i.e. I am unhappy to reject the claim that Hanimal is Holmes.

The problem of multiple-kind fission may be presented as a commitment to accept the truth of the following claims:

- Holmes = Holmes.
- Hanimal = Holmes. They are related by biological continuity.
- Hannibal = Holmes. They are related by psychological continuity

From these claims it follows that:

- Hanimal = Hannibal.

But surely we will say that:

- Not: Hanimal = Hannibal. They occupy different locations, are composed of different matter, and are neither biologically or psychologically unified.

When discussing the 'standard' fission of Brown in the previous chapters, I considered the following replies: Brown ceases to exist and Lefty and Righty come into existence (closest continuer view); Lefty and Righty were there all along and 'Brown' refers ambiguously to both before fission (multiple occupancy view); Lefty and Righty were identical before fission and are distinct after fission (temporary identity thesis).

I am not aware that anyone has suggested that in the multiple-kind fission case, Holmes ceases to exist and Hanimal and Hannibal come into existence. Someone accepting the psychological criterion, even one involving a uniqueness requirement, will think that Hannibal is identical with Holmes and thus Holmes survives. Rather the challenge is to explain what to say about Hanimal. Should Hanimal be argued not to be identical with Holmes and thus to come into existence at the time of the multiple-kind fission? Or is there a way to allow that Hanimal was there all along? If the proponent of the psychological criterion can't account for the relation between Hanimal and Holmes in a satisfactory way, the possibility of multiple-kind fission will count against that criterion.

It may be suggested that Hanimal isn't identical with Holmes. The view that Hanimal doesn't exist before the multiple-kind fission and isn't identical with Holmes, I will call 'personalism.' According to personalism, Hanimal is not identical with Holmes because he isn't psychologically continuous with him. But then Hanimal is a numerically different human organism from the one we pick out before Holmes undergoes the BST.

Let me note about the name 'personalism' that I have chosen this name for the view in order to contrast it with 'animalism.' Animalism characteristically holds that the psychological criterion is not a criterion for the identity of any animals, and that the biological criterion states the persistence conditions for all animals, including those which are our paradigm examples of persons; and this is so even if there might be persons with other persistence conditions than biological ones. Personalism, on the other hand, characteristically holds that the biological criterion is not a criterion of identity for any persons, and that the psychological criterion states the identity criteria for all persons, including our paradigm persons which are also human animals. Of course there probably are organisms, even human ones, with non-psychological persistence conditions, but the personal ones aren't among them. In chapter 8 I discuss the prospect of personalism as an alternative to animalism.

The personalist view entails that possibly a human organism like Hanimal doesn't have a past despite being uniquely biologically continuous with a past human organism. And it entails that possibly an individual that is a human organism is identical to a future individual, which is not a human organism and which is not its biological continuer, despite the presence of a biological continuer. So biological continuity is neither necessary nor sufficient for the persistence of human organisms. Still, for *some* individuals that are human organisms, psychological continuity is a necessary and sufficient condition for their persistence, even if there are human organisms

such as Hanimal the persistence of which can't involve any sort of psychological relations whatsoever.

The main challenge for the personalist is, I think, to explain why we should not take the relation of biological continuity to constitute Hanimal's identity with Holmes. Even if we think that we have good arguments for the psychological view of personal identity, according to which Holmes is identical with Hannibal, we are still left with the puzzling claim that Hanimal isn't identical to Holmes. This is especially pertinent given that it seems very natural to think that human organisms *sometimes* persist in virtue of mere biological continuity. Hanimal presumably persists in virtue of mere biological continuity after he comes into existence, i.e. after the multiple-kind fission, and human vegetables in general, foetuses, anencephalic babies etc., all seem to persist in virtue of the continuity of their biological life.

Personalism would parallel an approach to the standard fission case according to which one and only one of the fission offshoots was identical with Brown. I challenged those responses to show what could possibly give one reason to think that one of Brown's offshoots was him while the other wasn't. A similar challenge is presented to the personalist: Why is it that Hanimal's biological continuity with Holmes, who is himself a biological organism, doesn't constitute their identity? And if Hanimal's biological continuity doesn't make Hanimal identical with Holmes, can it then be maintained that Hanimal's identity is constituted by biological continuity at all? And if Hanimal's persistence isn't a matter of biological continuity, then what can it plausibly be said to be?

Another view, similar to personalism in that it holds that in multiple-kind fission only one of the offshoots is identical with Holmes, is 'animalism' according to which Hanimal is, and Hannibal isn't, identical with Holmes. Animalists think that if Holmes is really a human organism, then Holmes is

identical with a future individual if and only if that individual is biologically continuous with him. Hence the animalist is committed to the claim that possibly a person like Hannibal doesn't have a past, despite being psychologically continuous with a past person such as Holmes. And possibly a person like Holmes is identical to a future individual which is not a person and which is not psychologically continuous with him, despite the existence of a person who is psychologically continuous with him. Animalists thus face hard questions about the relation between Hannibal and Holmes similar to those facing personalists with respect to Hanimal and Holmes.

Faced with the multiple-kind fission scenario, one might also think that we should take it to be a case in which Hanimal and Hannibal are there all along both occupying the location allegedly occupied by Holmes. And by the name 'Holmes' we refer ambiguously to two non-identical individuals of different kinds and with different persistence conditions: a person and a human organism.

In real life so-called vegetable cases, a human person such as Holmes loses the capacity to realise a mental life and has no psychological continuer. If we call the biological continuer of Holmes 'Hanimal*', then it seems natural to think that Hanimal* is identical with Holmes, since Holmes is Hanimal*'s biological predecessor. Hanimal* is a human organism and such individuals plausibly have biological persistence conditions. And, more exotically, we may imagine a case in which a BST is performed on Holmes and his organism destroyed, resulting in a bionic psychological continuer, Hannibal*, who is Holmes's unique psychological continuer and thus identical with him.

Now assuming that we think that in the two mono-cases Holmes survives as Hanimal* and Hannibal* respectively, the multiple-kind fission case may be seen as a case in which both of the series of events securing Holmes's

survival in these two cases co-occur. And keen to maintain that Hanimal* and Hannibal* of two other possible cases are none other than Hanimal and Hannibal, we claim that they both exist before Holmes's fission into distinct individuals of different kinds. 'Holmes' we might suggest is ambiguous. The view that Hanimal and Hannibal are both there all along in multiple-kind fission but not identical, I will refer to as 'coincidentalism.'¹¹²

Coincidentalism maintains that Hanimal and Hannibal are non-identical before fission, that they are located in the same place at the same time composed of the same physical parts with the same internal microstructure, and that they are of different kinds with different persistence conditions.¹¹³ In chapter 8 I discuss the coincidentalist response in more detail. In particular I criticise it for being forced to acknowledge too many token persons and other token properties before Holmes's fission. I then consider Baker's notion of constitution and present my worries. Baker's constitution view takes the numerical non-identity of Hanimal and Holmes at t2 to entail that they can't be numerically identical at t1, but to avoid the problems arising from taking them to be two numerically distinct individual at t1, she suggests that they stand in the relation of constitution, which allows us to count them as one though they are not numerically one. In this way the necessity of identity isn't compromised.

I criticise Baker's view for driving a wedge in between numerical identity and 'counting as one' by allowing that individuals count as one without being one and the same. Secondly I complain that a similar line of reasoning

¹¹² My discussion of coincidentalism will focus on Baker's constitution view, but while Baker's notion of constitution and 'constitutionalism' entails coincidentalism, one might be a coincidentalist without accepting Baker's notion of constitution. In this introduction I introduce coincidentalism in general.

¹¹³ Note that coincidentalism here is described as a relation between two distinct individuals, not between an individual and the mass of matter of which it is made, nor as a relation between a whole and the parts composing it. The case at hand has to do with how to deal with the possibility of two individuals of different kinds being composed of the same matter at the same time and thus sharing the same location.

may be used to argue that we should find a way to count Hanimal and Hannibal as two at t2, without this entailing that they are numerically non-identical at that time, if we think that they are numerically identical at t1 (i.e. not numerically non-identical at t1), and want to defend the eternality and necessity of numerical non-identity. Either way we will give up on the eternality of either numerical identity or non-identity, and have to accept ways of counting as one what is not numerically one or as many what is not numerically a plurality.

Finally there is the solution consisting in pointing out that the apparently inconsistent identity claims are only *apparently* inconsistent. If we point out the times at which they are said to be true, we avoid the inconsistency. That's what the temporary identity thesis points out. So indexing the identity statements to times we get:

- t1: Holmes = Holmes
- t1: Hanimal = Holmes
- t1: Hannibal = Holmes

But from this we don't get the unacceptable claim that Hanimal is identical with Hannibal. Instead we get the temporally indexed identity claim:

- t1: Hanimal = Hannibal.

But, assuming that identity statements aren't temporally rigid, this doesn't conflict with the temporally indexed version of the non-identity claim:

- t2: Not (Hanimal = Hannibal).

Accepting the temporary identity thesis allows the proponent of the psychological criterion to reconcile it with the plausible view that paradigm

persons are human organisms, without having to deny that Hanimal has a past life before Holmes's multiple-kind fission. As a matter of fact, the temporary identity theorist can accommodate pretty much everything we want to say about the case. He can make sense of the natural description according to which we'd say at t2 that Hanimal and Hannibal used to be identical, but aren't identical anymore. We can maintain that they both have a history going back before the multiple-kind fission due to their continuity relations with Holmes, and before the fission we can say that Holmes *becomes* a bionic person and that Holmes *becomes* a mindless human organism, without saying that after the fission Holmes is both Hanimal and Hannibal.

Now consider this set of claims:

- t1: Hanimal = Hannibal.
- t2: Hanimal is in the hospital.
- t2: Hannibal is in the cinema.

Let's say that I am Hannibal. At t1 I can truly say, 'at t2 I will be a person going to the cinema' and that 'I will *always* be a person', and that 'I exist in the future if and only if there is someone who is psychologically continuous with me.' This is what both personalists and constitutionalists want to say.

But I might also truly say at t1 that 'I will be in the hospital tomorrow', and 'I am a human animal *at all times* at which I exist', and 'I exist at some other time if and only if there is an individual at that time which is biologically continuous with me.' These are the sorts of claims both animalists and constitutionalists would like to be true when uttered by me. On the temporary identity thesis I can utter all these claims truly at t1.

What I can't say, and what isn't entailed by the three claims, is that at t2 I will both be in the cinema and in the hospital, and be both a mindless human

organism and a bionic person. I think the case for the temporary identity thesis will turn out to be even stronger when considering the multiple-kind fission problem.

Finally I think it is also worth noting that adopting the temporary identity thesis may open up interesting ways of approaching the questions about the alleged circularity of a psychological criterion and about what matters in survival.

In the standard fission case Parfit argues that we have to accept that Brown isn't identical with either of the fission offshoots, but that they are related to him by the relation that matters in survival. Hence, Parfit concludes, identity isn't what matters in survival. On the temporary identity view we can maintain that identity is what matters in survival. The question about who matters is relative to a time. We want to say that before fission Lefty and Righty matter to Brown. The reason for this is that before fission they are both identical with Brown. And after fission Brown matters to them (as our past life normally matters to us), since after fission it is true to say that they were identical with Brown *before* fission. And we can say all this, without being committed to the claim that after fission they are identical with Brown. So when asked what future individuals matter to me, I can say: Those individuals with whom I am now identical. And if asked, what past persons matter to you?', I can answer that those persons with whom I was identical matter to me. What is clear is that what matters is *identity*, but identity is temporary. So we can say that what matters in survival is temporary identity.

Similar considerations apply to the discussion of circularity and memory. We can say that I only remember the past of those persons with whom I am identical. Take an 'inverse' fission case, a fusion of Lefty and Righty resulting in Greene existing at t3. At t3 Greene is identical to Lefty and identical to Righty, hence at t3 Greene can truly be said to remember their past, i.e. what

they experienced at t_2 , despite the fact that at t_2 Lefty and Righty are not identical. At t_3 Greene has a genuine memory of the experiences had by Lefty and Righty at t_2 , since at t_3 he is identical with them. Importantly this doesn't entail that there is a time, t_2 , at which Greene experiences both what Lefty experienced at that time (waking up in room 102) and what Righty experienced at that time (waking up in 104), just as there isn't a time when Brown is both in room 102 and in room 104, though there is a time at which it is true that Brown will be in 102 and that Brown will be in 104. If we appreciate that identity is temporary we need not appeal to the notion of quasi-memory. I only remember the past of persons with whom I am identical.

In the remaining chapters I will consider ways for a proponent of the claim that Hannibal is identical with Holmes to deal with the problem of multiple-kind fission. The problem is how to account for Hannibal's relation to Holmes. I conclude that the temporary identity thesis succeeds in making best sense of this case. In particular, I think it is worth pointing out, how it allows us to reconcile two very plausible claims, which by many are argued to be incompatible, namely the claim that we are human organisms and have psychological persistence conditions. I also think I show that personalism is a view which deserves more credit than it is traditionally given and so to speak cancels out animalism. I will not discuss the animalist view as such, since it is a view which rejects the identity of Hannibal and Holmes, but in chapter 8 I present a case for personalism as an alternative to animalism. However I think that since my discussion shows that there are ways to deal with the multiple-kind fission problem, animalists can't assume that just because paradigm persons are human animals, they must have nonpsychological persistence conditions.

8. Personalism and Animalism

8.1 INTRODUCTION

In this chapter I will consider whether there is any prospect in the personalist suggestion that we should simply reject that Hanimal is identical to Holmes despite their biological continuity. I begin by focusing on the alleged implausibility of this suggestion and compare it with its animalist competitor, which states that in the multiple-kind fission case, we should reject that the psychological continuity between Hannibal and Holmes constitutes their identity. The animalist thinks he can prove the psychological criterion wrong by pointing out that paradigm persons are identical with human organisms, and that it is implausible bordering to the absurd to claim that something which is a human organism has psychological persistence conditions. The absurdity is supposed to be reflected by the fact that it would leave us forced to describe the multiple-kind fission scenario as one in which we deny the identity of Hanimal and Holmes, while confirming the identity of Holmes and Hannibal.¹¹⁴

I think the discussion shows that both the animalist and the personalist description of the multiple-kind fission case suffer from the weakness that they don't accommodate the plausible claim that both of Holmes's continuers can truly be said to exist where Holmes is at t1. Both Holmes's psychological continuer and his biological continuer have an equally good claim on being Holmes and neither continuity relation can comfortably be said to stand as metaphysically *dominant*. This is a small victory for the personalist since his view, that persons are identical to human animals and have psychological

¹¹⁴ For an animalist like Olson this entails that it ceases to be a relation which can constitute the persistence of anything at all, even in the absence of a competing psychological continuity relation. Burke (1994b) suggests that a relation such as biological continuity may still constitute identity despite there being cases where there is a co-occurring continuity relation which has priority over biological continuity. I'll come back to questions about dominance in a section 8.5.

persistence conditions, is typically taken to be too implausible even to consider within an endurantist framework.

Personalism is shown to be a way for proponents of the psychological criterion to stand their ground without adopting revisionist metaphysical claims such as fourdimensionalism, kind-relative identity, or temporary identity.¹¹⁵ However I don't think that personalism achieves more than a stand-off with animalism. My conclusion is thus that we shouldn't give metaphysical priority to either biological or psychological continuity. Instead we face the task of accounting for how both Hanimal and Hannibal can be said to exist before Holmes's fission.

In the sections to come I consider animalist criticism of the personalist claim that we are human animals with psychological persistence conditions. I focus on arguments from Olson and van Inwagen. The discussion is to a large extent focused on the question whether personalism is a view which has any plausibility at all. I think I show that it is, and thus I motivate the view that multiple-kind fission is a case in which we have two equally good candidates for being Holmes in Hanimal and Hannibal, and that the metaphysical tie between animalism and personalism should make us look for other ways of dealing with multiple-kind fission such as coincidentalism and temporary identity.

8.2 PERSONALISM AND THE PROBLEM CASES

My discussion of personalism will involve reference to three problem cases. The problem cases can be represented in this diagram:

¹¹⁵ Thomson (1997, p. 209) thinks that 'a hybrid view' which combines the psychological criterion with the claim that 'persons are their bodies' must accept a temporal parts ontology. Shoemaker (1984, p. 113) seems to think that the only way of making sense of the problem cases besides adopting his coincidentalist position is to accept relative identity. Olson (1997, 159 ff.) argues against temporal parts and kind-relative identity solutions. I assume endurantism and absolute identity in this thesis.

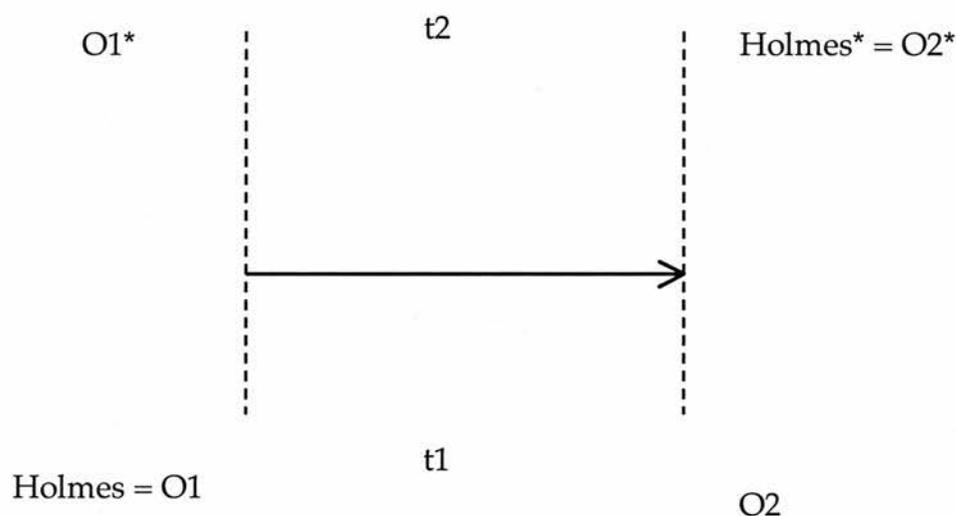


Figure 1

The BST-Case

Let 'O1' and 'O2' be names of distinct human organisms, and 'Holmes' be the name of a person. At time t1 Holmes is identical to human organism O1, the very organism Holmes points to when Holmes points to himself. Then there is a brain-state transfer between Holmes and a human organism O2. This means that at a later time t2, there is a person named 'Holmes*', who is identical to a human organism named 'O2*.' According to the psychological criterion, Holmes* is identical with Holmes. Still it is certainly natural to think that the organism O2* is identical to O2 and not to O1. This case is parallel to the multiple-kind fission case described in the introduction except that in that case, Hannibal is identical to a bionic individual at t2, not a human organism.

The Vegetable Case.

The punctured line from Holmes to O1* may be seen as representing a case in which Holmes at some point between t1 and t2 suffers brain damage resulting in a human vegetable at O1*. On the psychological criterion Holmes and O1* are not identical. Yet it seems natural to think that O1* and O1 are identical.

The Foetus Case

Another scenario is the Foetus Case represented in the diagram by the punctured line between O2 and O2*, which may be seen as representing a case in which a human organism develops from e.g. the foetal state, at which it doesn't have any psychological properties at all, to its adult stage involving a full-blown psychological life. On the psychological criterion Holmes* and O2 are not identical, since there is no relation of psychological continuity between them. Yet it seems natural to think that O2* and O2 are identical.

On the view I call personalism, the person Holmes is identical with a human organism O1, and thus O1 has psychological persistence conditions. Since O1 and O1* are not psychologically continuous, they are distinct. And since O2*-O2 are not psychologically continuous they are distinct. Yet, the line traceable between O1 and O2* represents an identity, since they are related by psychological continuity.

To begin with I will try to alleviate the immediate implausibility that the personalist description of the cases tends to be saddled with. The personalist description of the Vegetable Case and the Foetus Case entails that human organisms O1 and O1* (the Vegetable Case) are not identical despite the presence of spatiotemporal and organic continuity between them, and the same goes for O2 and O2* in the representation of the Foetus Case. This may be hard to believe.

What is immediately hard to accept about the BST-Case is that O1 and O2* can be identical given the absence of any spatiotemporal continuity at all. How can an organism located at place p1 at time t1 end up at another place p2 at a later time t2 merely on the basis of a transfer of information? A major challenge to proponents of personalism is to alleviate its immediate implausibility.

I think an important element in thinking that there is something very implausible about the way personalism describes the puzzle cases has to do with an unwarranted weight being put on the fact that the individuals in question are *human organisms*. If appropriate attention is given to the fact that O1 and O2* are *not just* organisms like worms or trees, but also *persons* who have rich mental lives, memories of their past and plans for their future, it is, I submit, not wholly implausible to think that O1 is identical with O2* as the personalist claims. As a matter of fact, I would suspect that many will find it rather hard *not* to think that this is what we should say about the case, as indeed the grip of the transplant intuition shows. We may change the names 'O1' and 'O2*' to *personal names*, such as Alf (O1), and Bob (O2*), thereby emphasizing that we are dealing with persons at the times in question, i.e. beings with a complex mental life. When it turns out that at t2 Bob is psychologically continuous with Alf at t1, it comes quite natural to think that they are identical persons. This may remove some of the implausibility reactions generated by focusing singularly on the fact that we are dealing with mere organisms.

With respect to the suggestion that O1-O1* and O2-O2* don't represent the continued existence of a human organism, this may admittedly seem implausible given the spatiotemporal and biological continuity which appears to hold between them. It is hard not to think that there is a single animal life represented by each of the lines.

In response to this objection, I think two features of the case should be pointed out. Firstly, due attention should be given to the fact that there is a psychological life which ends in the one case and begins in the other, with all the difference in properties that that entails. Considering the radical change in psychological properties which takes place between t_1 and t_2 in the cases described, I suppose most people would at least hesitate to reject the personalist description out of hand.

Secondly, the personalist is not committed to saying that there isn't a sense in which there is the continuation of the same biological being. He can invoke the distinction made by Olson (cf. my discussion in chapter 3) between practical and numerical identity and say that 'biologically speaking', or 'from the point of view of the biologist', there is 'practically the same' entity existing from t_1 to t_2 in the two cases. However, biologists are in no position to claim that the way biological science describes the material world is authoritative when it comes to metaphysical questions about numerical identity and persistence.

In general I think that a lot can be said in favour of personalism since it combines two individually quite attractive views, namely the view that we are enduring material beings, human organisms, and that we have psychological persistence conditions. Whether one is an animalist, coincidentalists, or personalist, one finds those two claims individually very plausible. Animalists reject the psychological criterion, coincidentalists reject that paradigm persons are identical to human organisms; in this company I think personalism is an attractive but somewhat neglected view.

It is important to underline that personalism challenges the animalist assumption that any organism must be reidentified on the basis of purely spatiotemporal and biological continuity. When the animalist claims that

persons are human animals, he must also accept, as it were, that most human animals are persons. We must be careful not to ignore a very significant aspect of the identity of persons and animals accepted by animalists, namely that persons may not have the persistence conditions associated with animals. Once it is accepted that human organisms are persons, there is an onus on the animalist to show that this doesn't entail that some animals have psychological persistence conditions. It is, I think, unfair but very easy to think that the onus is on the personalist to show that his view is acceptable. But animalism doesn't automatically come out of the truth of the claim that persons are animals. If there is doubt about whether the personalist description of the problem cases is plausible, this doubt should reflect back on the animalist view too. After all, the basis for that doubt is the claim they have in common: paradigm persons are identical to human animals.

Personalism purportedly describes the puzzle cases in highly controversial ways, but as it happens animalism must describe the puzzle cases in ways which are not easy to accept either. However, unlike the personalist, the animalist can't say that his position is the result of combining two individually plausible theses. The animalist view that we have nonpsychological persistence conditions doesn't at the outset appear attractive, as indeed animalists note themselves.

Eric Olson explicitly discusses the view that we are human animals with psychological persistence conditions (Olson 1997, p. 109 ff.). According to Olson personalism leads to redescriptions of the puzzle cases which are 'novel and absurd' (1997, p. 111). Olson's discussion is both enlightening and revealing and thus worth considering in more detail.

8.3 ANIMALISM AND PERSONALISM

Olson points out that the personalist's way of describing e.g. the Vegetable Case (the Holmes-O1* line in the diagram) is 'not the way a physiologist

would describe it' (1997, p.113). From a physiological point of view, the Vegetable case is a case in which there's 'continuity of life-sustaining biological functions and this makes it natural to suppose that a single animal is present throughout.' An animal loses a capacity which entails a 'profound alteration' but no substantial changes are brought about. Nothing goes out of existence and nothing comes into existence anymore than this happens when you change your hair colour.

The debate about identity over time is, as Olson stresses, about numerical identity. Thus it is about what it takes for persons to *cease* to exist. I think that in order to avoid confusion, it will be helpful to distinguish three ways in which the issue at hand invites us to talk of 'death.' To say that someone has died is in everyday language to say that a certain biological life has ended. Call this 'biological death.' But when faced with e.g. vegetable cases we often describe them by saying 'there's a sense in which Uncle Bob is dead now' to express that *the person* called 'Uncle Bob' is no longer. Call this 'psychological death'.

Now the metaphysical question at issue has to do with whether Uncle Bob is still around despite the fact that we think there is a sense in which he is dead, when all that is left is a human vegetable. The personalist has to make a case for the claim that it isn't obvious that biological death has any better claim to be what 'metaphysically ceasing to exist' comes to for human persons than 'psychological death.' However I do not think that reflection on the issue places the personalist in a spot tighter than the spot the animalist is in. The personalist can justify the non-identity of a human person and a human vegetable in a way parallel to the way in which the animalist justifies the non-identity of living animals and corpses.

When an organism dies in the biological sense (and in non-violent ways), it is succeeded by a spatiotemporally continuous corpse. According to the

animalists such as Olson, living animals don't continue to exist as corpses. Hence, when the health authorities are fishing dead swans out of the lakes in order to check for bird-flu virus, the dead swans are not strictly speaking identical with the swans that might have been infected with the flu, even if investigating the dead swans can provide evidence for whether some past swan was infected, because the corpse and the living swan don't stand in the appropriate biological continuity relations. The corpse is not a continuation of the swans biological life. Similar considerations apply to the case of human organisms and corpses.

According to an animalist like Olson (my token animalist in this thesis) the reason why the living organism doesn't survive as a corpse is basically that the changes going on in an animal when it dies are 'quite dramatic.' Olson writes: 'All that frenetic, highly organised, and extremely complex biochemical activity that was going on throughout the organism comes to a rather sudden end' (Olson 1997, p. 151).

Whatever the merit of such a response, it seems to me that it can be emulated by the personalist. It may be that there are radical and dramatic changes taking place in an organism when it biologically dies, but if we take seriously the animalist claim that it is *the organism* which is the person, and thus has all the psychological properties characteristic of personhood; if it is the organism that has memories, hopes and plans for the future and so on, and it is the organism which is the subject of a highly organised and extremely complex *psychological* activity, then the irrevocable cessation of such activity may be taken to be relevantly dramatic too. At least if we approach the subject unprejudiced. When the animalist accepts that animals are persons, he also attributes to them all the psychological properties that goes with being a person, and thus can't ignore the enormous change that *the animal* undergoes by the complete loss of such properties. It is, to use Olson's own phrase, 'very dramatic.'

When Brezhnev was no longer biologically alive his corpse was put in the Kremlin for people to see. According to the animalist, the corpse did not exist when Brezhnev was alive, and Brezhnev did not survive as a corpse, despite the spatiotemporal continuity between the corpse and the living organism Brezhnev. The claim is that *metaphysically* the spatiotemporal continuity isn't significant on its own; the metaphysically significant change with respect to a person like Brezhnev is the cessation of biological functioning. But why should a *metaphysician* take *biological* relations and changes to be decisive for whether there is numerical identity between Brezhnev and the corpse on display? Why not think it's the absence of psychological continuity which is metaphysically significant? Such change is, I contend, quite dramatic, and secondly, if spatiotemporal continuity *per se* isn't metaphysically significant with respect to identity, then it shouldn't be thought to be particularly implausible to claim that a person ceases to exist despite the presence of a spatiotemporal, albeit biological, continuer. I'm not even sure that biologists themselves take their views to involve any opinion with respect to the metaphysics of the entities they investigate. I suspect you can find biologists disagreeing about when Brezhnev metaphysically ceases to exist.

Remember that when Olson attempts to reject the Transplant Intuition, he brings in the notion of *practical* identity, which he uses to describe the identity we take there to be between the brain donor and the brain recipient in the transplant scenario. But the practical interest we have in tracing psychological continuers, Olson suggests, doesn't ground such continuity as being of metaphysical importance. Speaking on behalf of the personalist my counter question to Olson is this: 'Why should biological continuity be thought of as anything more than a *practical* notion which picks out relations investigated by a special science, in this case biology?' It is not clear why continuity under a *biological* concept is privileged with respect to numerical

identity over time and not merely a notion of what Olson calls practical identity (cf. Olson 1997, p. 65 ff.) which is of interest to biologists.

If one is a coincidentalists one might wish to present the animalist with a 'corpse problem' and ask why organisms don't share their matter with 'corpses-to-be', which have different persistence conditions? However this is not my concern. I grant the animalist his claim that an organism doesn't coincide with some other thing, a corpse-to-be, and that an organism doesn't metaphysically survive as a corpse. The present defence of personalism suggests that the personalist can utilise considerations analogous to those presented by the animalist. When a human animal, or some other organism, e.g. a dolphin or a cat, has a complex psychological life, when it is psychologically alive and not just biologically alive, the personalist claims that psychological death is metaphysical death.¹¹⁶

To sum up: The animalist supports his criticism of personalism by pointing out that biologists wouldn't think that an organism goes out of existence because it loses mental capacities. The entities and continuants which are investigated in biology are those which display organic life. What biologists are interested in is *biological continuers* or *biological lives*, and according to biologists these entities don't cease to exist by losing or gaining psychological properties. But the personalist may reply that there is no reason to take biology to be an authority when it comes to metaphysical questions.

What I have pointed out is that the animalist is in a bad position to claim that personalism is particularly marred by implausible descriptions of puzzle cases such as the Vegetable Case. Personalism allows a human organism to be superseded by *another* human organism in a way so seamless that it is hard to accept that there is a substantial as opposed to mere qualitative

¹¹⁶ I come back to the question whether nonpersonal animals such as dolphins and cats have psychological persistence conditions in the last section of this chapter.

change. I have merely pointed out that a similar problem faces the animalist's claim that you cease to exist when your biological life ceases to be continued. If the animalist acknowledges and argues that there is a substantial change brought about by the cessation of biological functioning, it seems that there may be an analogous argument in favour of thinking the same about psychological functioning.

From the personalist description of the Vegetable Case and the Foetus Case it is clear that personalism allows that some human organisms may cease to exist despite the presence of a spatiotemporal and biological continuer. What about the BST-Case?

The personalist will say that in the BST-Case organism O1 is identical with some future human organism, O2*, which is not spatiotemporally and biologically continuous with it. So we get the claim that some human organisms will be identical to a future individual, which isn't a spatiotemporal and biological continuer but a purely psychological continuer; and importantly this is the case despite the presence of a distinct, but purely spatiotemporal and biological continuer. This sort of claim may not be easy to swallow in case we focus on the fact that O1 and O2* are both *organisms*, as are O2 and O1*.

As in the discussion of the personalist response to the Vegetable Case I think it is vital not to ignore the fact that when we ask, 'What happens to the human organism O1, when the BST takes place?', we are not asking about a *mere* organism. The organism is, according to both the animalist and the personalist, a being with a complex mental life including memories of the past, plans for the future and so on; and the possession of all these properties, when properly appreciated, tends to make us think of a human person as quite different from mere organisms such as trees or worms.

When the animalist argues that the personalist description of the BST-Case is unacceptable, the animalist argument seems to be based on two claims entailed by personalism:

(P1) Possibly a human organism metaphysically ceases to exist despite the presence of a (unique) spatiotemporal and biological continuer.

(P2) Possibly a human organism is identical to a future individual which isn't a spatiotemporal or biological continuer despite the presence of a (unique) spatiotemporal and biological continuer.¹¹⁷

Now if we, along with animalists, think these two contentions are highly implausible, we should remind ourselves that they are the result of taking seriously two individually very plausible views, one of which the animalist is very keen to accept too; the claim that persons are human animals and the claim that persons have psychological persistence conditions. The possibilities stated in (P1) and (P2) are consequences of these very plausible claims, and our reluctance to accept (P1) and (P2) doesn't in itself show them to be incoherent or absurd. What they do make explicit is what someone agreeing with the animalist claim that persons are identical with human organisms must accept, if he also accepts that the psychological criterion is correct, a criterion which even animalists admit they find attractive.

The ammunition used against personalism is basically that it is wildly implausible because of what it entails will happen to human *organisms* in the problem cases. I have already made a case for the response that in the Vegetable Case (and I take the Foetus Case to be analogous to that case), the personalist can point out that reasoning analogous to the one employed by the animalist to meet points of criticism may be used to support the

¹¹⁷ I add the uniqueness clause here because I think that all adherents of continuity accounts of identity will presumably take satisfaction of this requirement to secure identity, even if some think it isn't necessary. I will come back to these two principles later in the chapter.

personalist's case. Let me finally consider whether the animalist description of the BST-Case comes across as more plausible than the personalist description.

Elaborating on the BST-Case thinking of it now as an *exchange* of the brain-states of two persons, we can give the organisms O1, O1*, O2, and O2* the personal names 'Alf', 'Alf*', 'Bob', and 'Bob*'. Let's assume that Alf is a poet and at t1 he is occupied rehearsing a recent poem he's written; Bob on the other hand is a keen chess player and at t1 he is calculating what move to make in a chess game. Then the BST procedure takes place, and a short moment later, at t2, the situation is this: Alf* is reaching a decision about what move to make in a chess game he takes himself to have been playing all evening, and Bob* takes himself to finish rehearsing a poem he wrote earlier in the evening.

It is a consequence of the animalist's description of the BST procedure that Bob* now falsely thinks he's finishing rehearsing a poem he began rehearsing earlier in the evening and written by himself. And Alf* falsely thinks he's just finished calculating a chess move in a game he's been playing during the evening. And this is so despite the fact that there is an extraordinary, but fully transparent causal link between the mental states of Alf* and Bob on the one hand, and between Bob* and Alf on the other.

The implausibility of the animalist description of course parallels the alleged implausibility of the personalist description. While the personalist has to reject that unified biological lives (O1-O1* and O2-O2*) always involve a single persisting organism, the animalist will have to reject the plausible claim that unified psychological lives always involve a single persisting person.

Now the animalist might point out that the mere fact that his view entails that a person's mental life may be completely incoherent from one time to another need not entail that we are not dealing with the numerically same person at those times. The animalist may argue that in real life cases of personality disorders, Alzheimer's syndrome etc. there occurs a lack of coherence in the mental life of what we still think is a single person from one time to another.

But I think the BST Case is relevantly different from ordinary cases in which we may think of a single person as the subject of a highly disintegrated psychological life. In the BST Case we have two persons (Alf* and Bob*) at t₂ and we have two persons (Alf and Bob) at t₁; given that we get a diachronically and causally unified psychological life from joining the mental state of Bob* with Alf as opposed to a completely disintegrated one by joining Bob*'s mental state with Bob's, I think it comes across as highly unsatisfactory if one insists that the psychologically unified life isn't the life of a single person. There would be something rather uncharitable about suggesting that the case is one in which two persons display a completely incoherent psychological life, when there is an alternative which allows for treating them as two quite normal and psychologically integrated persons, whose mental states are reasonably coherent.

The personalist allows us to maintain this more charitable and I think plausible description of the BST Case and thus advocates these two principles:

(P1) Possibly a human organism ceases to exist despite the presence of a (unique) spatiotemporal and biological continuer.

(P2) Possibly a human organism will be identical to a future individual which isn't a spatiotemporal or biological continuer despite the presence of a (unique) spatiotemporal and biological continuer.

Now the animalist wants to reject (P1) and (P2) in favour of:

(A1) Possibly a person ceases to exist despite the presence of a (unique) psychological continuer.

(A2) Possibly a person will be identical to a future individual which isn't a psychological continuer despite the presence of a (unique) psychological continuer.

It has been shown that it isn't comparatively less plausible to accept the personalist principles than to deny them in favour of the animalist principles. Denying the personalist principles hardly gives one an unproblematic and particularly plausible way of describing the problem cases. As soon as we draw appropriate attention to the huge difference between an organism which is truly said to be a person and thus is said to be *the subject of mental properties*, and a human organism which doesn't have those properties, it becomes significantly less obvious that such difference in properties shouldn't count in favour of acknowledging an ontologically significant difference between the two, despite their spatiotemporal continuity under the concept of organism. Personalism, I submit, is a viable alternative to animalism and it is, I think, fair to say, that we've reached a stalemate between animalism and personalism.¹¹⁸

¹¹⁸ Olson seems to acknowledge that personalism might achieve a stalemate (1997, p. 121).

8.4 IS PERSONALISM COMMITTED TO A METAPHYSICAL IMPOSSIBILITY?

In a recent article (van Inwagen 1997), van Inwagen targets the view that persons are human organisms and have psychological persistence conditions. His claim is that combining those views leads one to accept a metaphysical impossibility, namely that human organisms which are distinct at one time may be identical at another time. But this violates the necessity of identity.

Van Inwagen argues (p. 310 ff.) that a description of the BST-Case as a case in which a person moves from one place to another faces the following *reductio*:

- (1) At t1: you = human organism O1.
- (2) At t1: O1 is at place p1.
- (3) At t1: you are at place p1 (from 1, 2 & Leibniz's Law)
- (4) At t2: you = human organism O2*.
- (5) Not, at t1: O2* is at place p1
- (6) Not, at t1: you are at place p1 (from 4,5 & Leibniz's Law)
- (7) At t1: you are at place p1 & not, at t1: you are at place p1 (from 3 and 6).

I think there are two ways in which the proponent of the psychological criterion respecting endurantism and the eternity and necessity of identity can respond to van Inwagen's argument. One might follow the way of coincidentalism and simply take the sense in which persons are animals not to be the 'is' of identity. The coincidentalism will reject premises (1) and (4) and point out that persons are animals in a sense of 'are' which is not numerical identity.¹¹⁹ This is not a way to respond for someone who, like me,

¹¹⁹ One might point out that if a coincidentalism thinks that persons are identical with some material being, where this entails occupying a place, it would still seem to be vulnerable to the *reductio* stated in terms of identity with human animals. If we replace 'human animal' with 'X', where 'X' picks out a material being occupying a place, and a person is identical

agrees with both the animalist and the personalist that we are strictly speaking identical with human organisms.

The other way to respond is the personalist way. If we accept personalism, we will reject (5). We will point out that on the view under consideration, one according to which a human organism is identical to a person, one can't assume (5) without begging the question against us. In the BST scenario O2* is a *person* and thus to trace him back in time is to trace the line of psychological continuity, which we take to constitute the identity of persons. Assuming (5) is simply to assume that personalism is false, but that is what the argument is supposed to show in a non-question-begging way.

Importantly, personalism doesn't violate the necessity of identity. It doesn't entail that O1 and O2 are distinct at one time and identical at another (when they both exist), nor does it entail that O2* was both at p1 and p2 at time t1. Since personalism denies O2*'s identity with O2 by denying (5), we don't get (6) from accepting that (4) is true.

But van Inwagen's argument does bring out a challenge facing the personalist. While it might be reasonable to think that persons have psychological persistence conditions, we are still left with the plausible claim, that when a single physical and biological individual such as O2* stands in spatiotemporal and physical continuity relations (including causal relations) with O2, O2* is identical with O2 and not identical with O1. But then we will have to admit that the organism with which I am identical at t2, i.e. O2*, was not at place p1 at t1, and thus O2* is a distinct organism from O1. The insistence on the identity of O2* and O1 by the personalist doesn't make the physical continuity between O2* and O2 go away with no claim on being constitutive of their identity.

with X, then the argument seems to go through. For a coincidental (or constitutionalist) attempt to reply to van Inwagen's argument see Baker 2000, p. 142-145.

The general point which this brings to the surface is, I think, whether there can be said something to show that 'person continuity' (where this is understood as the continuity relation along which we trace a persisting person according to the psychological criterion) *dominates* animal continuity (the continuity relation suggested by the animalist for tracing an organism through time). Unless this is achieved I don't think personalism will have achieved more than a stand-off with animalism.

It is worth noting that both animalists and personalists must come up with some argument as to why animal persistence or psychological persistence is dominant with respect to human persons. It is not only the personalist who has a job justifying the claim that human persons have psychological persistence conditions. Animalists will be in a similar situation when they claim that human persons have biological and nonpsychological persistence conditions. In the next section I consider a proposal from Burke about how to decide whether one or another identity criterion applies to an object which satisfies sortals associated with incompatible persistence conditions on the face of it allowing the object to come apart from itself. I conclude that Burke's 'widest range of properties' criterion for sortal dominance doesn't cut it. I then consider Olson's 'locomotor argument' for the conclusion that 'person' is a bad candidate for being a sortal which determines the persistence conditions for the objects falling under it. I conclude that Olson's argument isn't satisfactory either.

8.5 PERSONALISM AND SORTAL DOMINANCE

Michael Burke (1994b) has argued that when faced with problem cases in which there appears to be a single object satisfying two sortals associated with incompatible persistence conditions, we should take its persistence conditions to be those associated with the *dominant* sortal. Thus when an object such as O2* satisfies both 'person' and 'human organism', and those

two sortals are plausibly taken to be associated with incompatible persistence conditions, as indeed the previous considerations show, one of the sortals will denote the kind the belonging to which determines the object's persistence conditions. That is to say, the object in question persists *as* an object of the kind denoted by its dominant sortal.

Burke writes:

Let's say that sortal *F* is object *o*'s *dominant* sortal just in case *o* has the persistence conditions it has in consequence of satisfying *F* (1994b, p. 605). On my account, an object that has the persistence conditions it has in consequence of satisfying *F* can persist only so long as it does not undergo changes that would cause it to satisfy a sortal that dominates *F* (1994b, p. 608). (...) In general how do we identify an object's dominant sortal? (...) An object's dominant sortal is the sortal that tells the object's sort, the sortal that tells what the object is. (...) Of the sortals satisfied by an object, the one that tells the object's sort, is the one whose satisfaction entails possession of the widest range of properties.(1994b, p. 610).

Burke admits that while his WRP criterion of dominance is perhaps vague, it gives agreeable rulings with respect to a set of representative cases which we may think involve a single object which satisfies sortals associated with incompatible persistence conditions. Some of the cases Burke considers are pieces of copper which are statues; hunks of metal which are typewriters; and hunks of cells (Burke's own expression) which are trees. The general idea is that satisfying sortals such as 'tree' and 'statue' entails having properties from various general categories of properties: physical, chemical, biological, psychological, functional, and aesthetic properties are the ones Burke mentions himself. In order to determine whether e.g. 'piece of copper' or 'statue' dominates when both apply to a single object, we are asked to consider whether falling under one or the other of the sortals in question entails possession of the widest range of properties. And, Burke remarks, 'What counts, on my criterion, is the entailed range of properties, not the range of specific properties entailed' (1994b, p. 610).

Consider the standard case in which 'Goliath' is the name of a statue, which, on the view under consideration, is identical with a lump of copper. There is a possible divergence in the history of a statue and a piece of copper. If we melt the copper the statue ceases to exist, while the lump survives. To determine whether Goliath is dominantly a statue or dominantly a lump of copper, and thus whether it has the persistence conditions associated with one or the other of those sortals, we consider what range of properties is entailed by being a statue and by being a piece of copper. Whereas being a piece of copper entails physical and chemical properties, being a statue not only entails physical and chemical properties but also aesthetic properties. Hence Goliath is dominantly a statue and has the persistence conditions of a statue. Or so one might think along with Burke.

Considering five representative cases Burke writes: 'Admittedly, the criterion is somewhat vague. But its vagueness seldom prevents a clear-cut decision' (1994b, p. 610). Now I don't think Burke's criterion does the job it's supposed to do. In what follows I'll briefly present my reasons for being unhappy with the WRP criterion, some of which are noted by Burke himself.

Firstly, as Burke notes, there are cases which create problems such as cases of 'found art' and amateurish art. Take a case of amateurish art. Let's say that I am given a lump of gold and shape it into an, albeit poorly made, statue of Beethoven. Now my statue is very poorly made and its aesthetic value and properties are of a standard which makes it very tempting to think of the statue as dominantly a very valuable piece of gold, rather than an poorly crafted work of art. On the other hand, if a master sculptor has used a piece of gold to make a statue of Beethoven which is an aesthetic masterpiece, the fact that it's also a piece of gold seems to be overshadowed by its being a statue. Now, in the case of my Beethoven statue, Burke suggests that we might think my statue of Beethoven is dominantly a piece of gold despite the

fact that 'piece of gold' entails a narrower range of properties than does 'statue,' and he admits that in the other case we may take 'statue' to dominate. However he dismisses the cases as 'atypical' (1994b, p. 614). Still it leaves one worrying about the feasibility of his criterion.

While cases of amateurish art and valuable materials make one suspect that the WRP criterion doesn't always fit our verdicts about dominance, there are also cases about which there is no clear-cut decision to be given on the basis of the WRP criterion.¹²⁰ It seems quite possible for objects to satisfy more than one artefact sortal. Take the case, mentioned by Rea (2000, p. 184), in which a statue is also a pillar. In this case the WRP criterion won't give us a clear decision about the persistence conditions of the object, since there seems to be 'no reason to think that, say, the range of properties entailed by being a statue exceeds or is exceeded by the range of properties entailed by being a pillar' (Rea 2000, p. 184). It is clear that both 'pillar' and 'statue' entail physical and chemical properties, and while 'pillar' entails certain functional properties, 'statue' entails certain aesthetic properties. In general it seems a strong contention of the WRP criterion to assume that no two sortals entailing equally wide ranges of properties can be satisfied by the same object, which is what I think is the general point one can make on the basis of Rea's remarks).

Faced with that problem, it might be tempting to claim that in cases where there isn't a clear decision to be made, we can recognise a third sort of thing which the object in question belongs to (cf. Rea 2000, p. 185). Maybe sortals which are cosatisfiable and entail having ranges of properties of the same width combine to make up a third sortal such as 'statue-pillar.'¹²¹ If one is willing to swallow a few camels, this reply might have something going for it. It would seem to imply that using a statue as a pillar would either entail

¹²⁰ This is a point made by Rea (2000, p. 183 ff.).

¹²¹ This is not a response that Burke presents himself.

that an object goes from being dominantly a statue to being dominantly a statue-pillar, which on Burke's account would entail that the object ceased to exist. Or, if one wants to avoid that conclusion, it may be claimed that the object was *always* a statue-pillar, despite the fact that neither the creator of the statue or anyone else had any recognition of this when the statue was created, maybe hundreds of years ago. Taking this reply will presumably entail that not even 'statue-pillar' can be known to dominate now, since some other even more elaborate sortal, e.g. 'statue-pillar-road block' might turn out to be satisfied by the object.

Even if we allow for this sort of response, i.e. accept sortals such as 'statue-pillar', to cases in which an object satisfies two or more artefact sortals whose satisfaction entails the same range of properties, the WRP criterion will be vulnerable to a second objection (cf. Rea 2000, 186). Rea considers a case in which a human being is also a chess piece (I suppose this is the case e.g. when a chess board is painted on a square and we dress up human persons as chess pieces and they perform the moves of the chess piece in the game). In this case it seems that 'the agreeable ruling' would be to think of the human chess piece not dominantly as a 'human chess piece', but as a person or a human organism. But if one allows for the 'statue-pillar' response to the undecided cases, this seems to force one to acknowledge 'human chess piece' as a sortal and that sortal will, contrary to what we think, be dominant when a thing is both human and a chess piece.

When it comes to the question about whether 'person' or 'human organism' is dominant, it seems hard to decide the question either way on the basis of the WRP criterion. 'Person' entails (insofar as we accept materialism) that persons have physical and chemical properties as well as psychological properties. 'Human organism' doesn't entail psychological properties, but

then it entails biological properties which is not entailed by being a person.¹²² So we don't seem to have come closer to a decision about whether human persons are dominantly *human* or dominantly *personal*. Furthermore, even in case it turned out that there was a decision to be made on the basis of the WRP criterion with respect to 'person' and 'human organism', the criterion is I think generally unsatisfactory in that there clearly will be cases about which it will seem to give the wrong answer, and cases in which there is no decision to be made since neither of two competing sortals entails the widest range of properties. Hence I am sceptical that Burke's WRP criterion supports a general dominance account, let alone a particular answer to whether 'person' dominates 'human animal' or not.

In the remaining sections of this chapter I will turn to discuss Olson's argument for the conclusion that 'human animal' is a better candidate for being the sortal which determines a human person's persistence conditions than 'person.'

8.6 OLSON'S LOCOMOTOR ARGUMENT

Olson, referring back to Wiggins (1980), introduces a theory of 'substance concepts' according to which:

(...) [E]very particular object falls under some kind or concept that tells us, in a special sense, what the object is, and not merely what it does or where it is located or some other accidental feature of it. And that concept determines persistence conditions that necessarily apply to all (and perhaps only) things of that kind. Concepts of this sort are substance concepts. (...) However the theory of substance concepts does not by itself tell us what substance concept you and I fall under (Olson 1997, p. 28).¹²³

¹²² I think it is worth noting that 'fully functional human organism' does plausibly entail psychological properties.

¹²³ Here Olson follows Wiggins (1980) taking the 'What is it?' question to provide a criterion for when a sortal is a substance sortal. I am not going to debate whether we should at all

If a concept *C* denotes a kind *K* and every individual belonging to that kind necessarily belongs to that kind, then the concept is a substance concept denoting a substantial kind. When we ask with respect to some object, in a 'special sense', a metaphysical sense, *what it is*, the correct answer will tell us, which substance concept it falls under. Thus on the theory adopted by Olson, the criterion for being a substance concept is that a concept answers the *what is it* question with respect to some possible object. I suppose 'unicorn' might be a substance concept if unicorns might possibly exist and unicorns is *what they are*.

Let me note here that there appears to be something a bit misleading about Olson's characterisation of how to understand the *what is it* question. He writes that when a concept tells us what an object is, i.e. what it is essentially as opposed to accidentally, this will not merely be telling us 'what it does or where it is located.' But there seems no reason to assume that telling us what, e.g. a cat or an oak tree *does* will not be an essential feature of it. I don't see why the capacity to perform some activity or function can't be an essential

accept the notion of a substance sortal, but grant Olson the idea (Olson notes that he assumes the Wigginsian theory, 1997, p. 28). Wiggins describes the notion of a sortal like this:

If somebody claims of something named or unnamed that it moves, or runs or is white, he is liable to be asked the question by which Aristotle sought to define the category of substance: What is it that moves (or runs or is white...)? Perhaps the man who makes the claim that something moves does not need to know the answer to this question, and one may envisage circumstances in which he can know that it moves without knowing what the thing is. Yet it seems certain that, for each compliant of a predicate like 'moves', 'runs' or 'white', there exists a known or unknown named or nameable kind to which the item belongs and by reference to which the 'what is it' question could be answered. Everything that exists is a this such.

If a man reports that the thing that runs is the same as the thing that is white, then his judgement has no chance of being true unless at least two preconditions are satisfied: (a) there exists some known or unknown answer to the question 'same what?'; and (b) this answer affords some principle by which entities of this particular kind – some kind containing things that are such as to run or be white – may be traced through space and time and reidentified as one and the same.

feature of some kind of thing. Nor do I see that we can rule out that to some possible objects their location is essential.

On the theory adopted by Olson, there is an intimate connection between something's being of a substantial kind and its having certain persistence conditions. The persistence conditions of an object are determined by *what it is*. Assume that 'tiger' is a substance concept and that tigers thus have their persistence conditions determined by being tigers; in that case I will say that tigers persist *as such*, i.e. as tigers. Notably the formulation that the kind to which something essentially belong determines its persistence conditions seems to suggest, that in order to find out what persistence conditions e.g. tigers or persons have, we must first be in a position to answer the *what is it* question with respect to those individuals. Indeed focusing on this question is part of the strategy of animalism. But, as I explained in chapter 2, our best way to determine what something essentially is may well be to consider what it takes for it to persist. In that chapter I discussed Olson's claim that focusing too much on persistence has misled proponents of the psychological criterion, since that criterion goes badly with plausible answers to the *what are we* question.

Again I want to point out that even if the persistence conditions of an object are determined by its belonging to a particular substantial kind, we may do well to let our views about persistence conditions guide our answers to what substantial kind an object belongs to. In the case at hand, if it turns out that we have good reason to think that some human organisms have psychological persistence conditions, it will be somewhat dogmatic to rule out that there is a substantial kind all of whose members have those persistence conditions. As Olson is aware of himself, his theory of substance concepts doesn't by itself tell us what we are. So in order not to come across as dogmatic, the animalist must give us some argument for why we shouldn't take 'person' to be a substance concept given that we think that

beings of the personal kind have their persistence conditions determined by belonging to that kind, not due to being e.g. human animals.

Coming back to my discussion of personalism and animalism, I submitted that the question seems to be whether we should take the persistence conditions of human persons to be the nonpsychological conditions associated with the sortal 'human organism', or the psychological conditions associated with the sortal 'person.' Personalist and animalist agree that paradigm persons are human organisms. The question is whether they have their persistence conditions determined by falling under the concept 'person' or the concept 'human animal.' I.e. the question is 'what substance concept you and I fall under.'

Burke offered a way of deciding this which I found to be unsatisfactory. Olson has presented an argument for why *nothing* can have its persistence conditions determined by being a person, arguing that 'person' is a bad candidate for being a substance concept (Olson 1997, p. 32). Olson's argument is different from Burke's in the sense that it aims at ruling out 'person' as supplying the persistence conditions for anything at all. If Olson is right, being a person has no more impact on something's persistence conditions than does being a philosopher. Burke (1994b) has an alternative view of the relation between substance sortals and persistence conditions, which entails that though a substance sortal is required to possibly determine the persistence conditions for some things falling under it, it need not do so for all things falling under it; it might be dominated by other sortals. E.g. 'piece of copper' is a sortal determining the persistence conditions for pieces of copper which don't satisfy a sortal, e.g. 'statue', which dominates 'piece of copper.' Olson's view of substance sortals doesn't allow them to be

dominated at all. I will leave aside these differences here.¹²⁴ If Olson is right, our paradigm examples of persons won't have the persistence conditions of the psychological criterion and that would undermine my case for accepting it.

Olson writes:

'To say that something is a person is to tell us something about what it can do, but not to say what it is. To say that something is a person is to say that it can think in a certain way – that it is rational, that it is ordinarily conscious and aware of itself as tracing a path through time and space, that it is morally accountable for its actions, or the like. It doesn't tell us *what* it is that can think in that way. We might still ask, Is the thing that can think a biological organism? A Cartesian ego or Leibnizian monad? An angel? A machine made of metal and silicon? What sort of thing is it that has those special psychological properties?' (1997, p. 32).

To illustrate and support his point, Olson considers the claim that 'locomotor' is a substance sortal and that there is a criterion of locomotor identity (1997, p. 32 ff). This is the criterion of numerical identity for all and only locomotors, a locomotor being something which can move about under its own steam (as do human animals, tortoises, white blood cells, nuclear submarines, cars etc.). An alleged criterion of identity for all and only locomotors may be this:

Criterion of Locomotor Identity A locomotor persists if and only if its capacity for locomotion is preserved - if and only if there is 'locomotive continuity.'¹²⁵

¹²⁴ Burke develops his alternative account of substance sortals in Burke 1994b. It is worth noting that Burke's account would allow for there to be human organisms which have biological persistence conditions even if there are human organisms which don't. Olson thinks one has to give up 'human organism' as determining persistence conditions for anything if it doesn't do so for all objects falling under it.

¹²⁵ Olson 1997, p. 32.

Take a locomotor like my car. If it loses its capacity to move, it will not be a locomotor any more. And if the thing which is my car stops being a locomotor, the thing which was a locomotor, the car with the engine in it, will cease to exist. Hence, the removal of the engine from my car will, given that the car is a locomotor, result in the end of one object and the generation of a numerically different one. The car without engine would be one thing, the car with the engine another thing. If we took the engine of my car and put it in a boat, we'd get the same locomotor first driving on land and later sailing on the local lake. And in case my car has an engine which can be split into two distinct functioning engines there might be a fission of my locomotor if the two engine parts come apart and function independently. This is all analogous to the typical cases of brain transplants, brain division etc.

The objects which qualify as locomotors may have very different internal physical structures. Both tortoises and nuclear submarine are locomotors, but the internal physical structure of tortoises and nuclear submarines comes across as being quite different.¹²⁶ There is a sense in which the similarity between locomotors is quite 'superficial' (I'll come back to the superficiality of similarities and differences in the next paragraph). Furthermore, on the locomotor criterion the mobile tortoise would cease to exist if it became completely paralysed, and something essentially different, a immobile tortoise, would come into existence. The mobile tortoise, being a locomotor, has the persistence conditions of a locomotor, and thus will cease to exist, when those conditions aren't met. Paralysis is such a case. So on the locomotor criterion the pre- and the post-paralysis tortoises are numerically different objects. But, again, the difference between a mobile tortoise and an immobile tortoise come across as being 'superficial.'

¹²⁶ There is clearly a sense in which this is true, but later I will suggest that physical structures, which are in this clear sense very different, are also in an important sense very similar, when they both realise e.g. mental properties.

Olson writes: 'The fact that two things are locomotors seems to be, in an important sense that is hard to define, a *superficial* similarity, whereas the mere fact that one thing is a locomotor and another is not is a superficial difference. It is a mere similarity or difference in one particular ability, which need not reflect any further similarity or difference' (1997, p. 34). Call this 'the superficiality worry.'

Now clearly we don't think 'locomotor' is a good candidate for being a substance sortal determining the persistence conditions of the things falling under it. Olson finds, that a part of the reason why a concept such as 'locomotor' is not a good candidate for being a substance concept is, that it is a concept which is characterised *functionally* in that its application conditions have to do with what an individual can *do* and not with its internal physical structure. So it is partly due to the fact that we are dealing with a functional concept that we get 'the superficiality worry' just presented. When it comes to the physical structure of locomotors the concept of locomotion is extremely liberal. And being liberal with respect to physical structure makes it hard to see how 'locomotor' could be an acceptable answer to the *what is it* question.

The locomotor argument now becomes this: We have seen that at least part of the reason why it seems problematic to think that locomotors persist *as such*, i.e. that there is a criterion of locomotor identity, is that locomotors may be only superficially similar, and this is partly due to the fact that it's a functional concept with no restrictions on the physical structure of the entities to which it applies. Locomotion, we might say, is 'multiply realizable' in the sense that it is realizable in multiple and different physical structures.

Olson now goes on to compare 'locomotor' with 'person.' Olson writes: 'Personhood, like locomotion, is merely a capacity or ability of a thing, and

different kinds of people may have no more intrinsic similarity than do different kinds of locomotors' (1997, p. 35). The concept of a person also has a functional characterisation in that persons are individuals which, in short, are capable of rational, self-conscious thought, and the possibility that very different physical structures realise personhood is widely acknowledged by proponents of the Lockean concept of personhood and the psychological criterion. It is thus possible that there is only a superficial similarity between persons, i.e. between bionic persons and human persons, and there will also be cases where there's a superficial difference between persons and some non-persons, e.g. nonpersonal human organisms and human persons. In this latter case, members of a biological species, with all the similarities entailed by this, come out as being of essentially different kinds of things. This is hard to accept. It seems that a personalist who rejects 'locomotor' as a substance concept has his work cut out for him given that the considerations canvassed about to thinking that 'locomotor' is a substance concept will presumably also apply to the functionally characterised Lockean concept of a person. Thus Olson challenges those who hold that 'person' is a substance concept to show how 'person' could be a substance concept even though 'locomotor' could not (Olson 1997, p. 36).

8.7 RESPONSE TO OLSON'S LOCOMOTOR ARGUMENT

In this section I'll present what I think are important differences between 'locomotor' and 'person' which provide reasons for rejecting Olson's locomotor argument.

One major difference between locomotors and persons, which I think is worth pointing out, is that the notion of a locomotor has little if any practical and theoretical importance. I am not aware of any practical interests with respect to which the notion of a locomotor is recognised or useful. Nor does locomotor appear to be a theoretically valuable concept, where this involves being relevant in explaining and predicting phenomena, 'locomotive

phenomena,' in a systematic way. Neither our scientific practices and interests nor our practices and interests in general have use for the notion of locomotors as such. The notion of a locomotor is thus not a natural kind concept, neither in the limited sense of being a concept useful for and recognised by the natural sciences, nor is it natural in the general sense of being a natural way for us to group together individuals for explanatory purposes in general. The notion simply has little general explanatory value.

Another important point to note is that there has been no temptation to see anything as persisting as a locomotor; perhaps partly for the reasons that we don't find locomotors to be a practically and theoretically important group of entities to single out, we have no inclination to think that things persist as locomotors, i.e. that there is a criterion of identity for locomotors as such. Contemplating the adventures locomotors can undergo, no unified account of the necessary and sufficient conditions for their continued existence announces itself as plausible.

Now I think it is fair to say that the concept of a person differs quite significantly from the concept of a locomotor with respect to practical and theoretical significance. The explanatory value of the concept of a person in dealing with many ordinary phenomena in a systematic way is, I think, incomparable with the almost non-existing explanatory value of 'locomotor.' The gamut of disciplines and areas of knowledge in which the concept of a person is of major explanatory importance is rather extensive: History, literature, social sciences, psychology, political theory, ethics, jurisprudence, fine arts, performance arts, religion, anthropology, pre-scientific 'daily life,' and so on.

And, furthermore, while nobody ever found it to be plausible to think that locomotors have persistence conditions as such, many have found it not only natural, but also quite plausible, to think that persons have persistence

conditions as such, not according to whatever else they might truly be said to be: robots, organisms, angels etc.

I suspect Olson and others might claim that 'person' is not a narrow natural kind concept (where 'narrow' indicates that it is a concept employed by one of the natural sciences), and that narrow natural kind concepts are the obvious candidates for being substance concepts. This will favour 'human organism' over 'person', since 'human organism' is recognised by biological taxonomy. However, while narrow natural kind concepts may be good candidates for substance concepts, it is not clear that concepts, which are of major importance from the perspective of other practices and interests, including social science and the Humanities, are bad candidates. It would be scientific to claim that the only metaphysically significant kinds are those recognised by and useful for natural science, and furthermore, insofar as we find it natural to think of persons as persisting as such, this counts in favour of recognizing 'person' as a substance sortal. (I will consider whether 'organism' is plausibly thought to be associated with a criterion of identity in our ordinary ontology in the next section). I contend that the fact that 'person' isn't a narrow natural kind concept doesn't rule it out as a substance sortal.

I guess some might claim that the fact that persons, unlike locomotors, are actually quite similar in physical structure makes the possibility of multiple realizability play less of a role in our thoughts about personhood. If we were actually confronted with e.g. bionic persons (and they with us), we (and they) might not be any more inclined to think that the same answer to the *what is it* question would be correct when asked with respect to both them and us. Olson seems to think so. Indeed it is common for people to think of persons and human animals, or 'men' as Locke called them, as the same kind of thing. But it is my impression that, when prompted to answer questions about the persistence conditions of persons, they will find there to be a

conflict between the persistence of persons and human organisms. So the idea that persons persist as such has a strong grip on us, as Olson also notes with respect to the Transplant Intuition discussed in chapter 2. The locomotor argument does nothing to diminish that grip.

Coming back to Olson's locomotor argument, I don't buy into the premise that functional similarities are superficial since they may be multiply realised in different physical structures. Nor do I think that similarity of internal physical structure is 'deep' *per se*. After all, physical beings will in a sense all be very similar in their physical structure insofar as they're all composed of atoms interacting in certain ways. What seems to make them importantly different is, at least partly, the sort of things they can *do*.

Take Olson's favourite substance concept 'human organism' or 'organism.' Now organisms are largely characterised by what they can do, what activities they are able to perform. If one goes through Olson's characterisation of what an organism is (1997, pp. 126-131) there's a lot of *activities* which something must have the capacity to perform if it is to qualify as being a human organism. Here's one characterisation of what it is to be an organism: 'An organism adjusts its activities to take advantage of the changing conditions in its surroundings. When the ambient temperature drops, a warm-blooded animal's metabolism increases and blood flow is directed away from its surface to conserve heat' (1997, p. 128). Other general capacities central to being an organism are the ability to metabolise and reproduce.

Finally I think Olson's suggestion that the lack of intrinsic similarity counts against taking a set of entities to be of a single kind stands in need of clarification. I guess Olson thinks that biological taxa qualify as picking out natural kinds, but presumably organisms may and do have what may be

considered to be little intrinsic similarity, unless one focuses on their functional character.

Now there are many important issues one might take up in this discussion which I don't have space for here. It might be that the functional similarity delineating the class of locomotors is superficial, and the physical similarity between certain non-locomotors and certain locomotors is 'deep,' but I don't think the point can be generalized. As a matter of fact, it seems plausible that delineating a group of individuals on the basis of common functional properties may be of great explanatory importance and underpin acknowledging them as being of essentially the same kind persisting as something of that kind, despite what Olson would call very different physical structures. Physical structures may be importantly very similar with respect to what capacities they are able to realise, and while locomotion as such seems of little relevance, realization of organic structure, metabolism etc. is admittedly of great interest and explanatory value. And so, I contend, is personhood.

As I have already suggested there is no particular reason to think that biologists are authoritative with respect to the metaphysical questions about how we persist and what we *really* essentially are. And as I have just pointed out, I think there are good reasons to think that being a person is a property which it is implausible to leave aside in the metaphysical debate as being of the same metaphysical status as the property of being a locomotor. As a matter of fact, I think there is a case for claiming that there are such significant differences between organisms that it is unlikely that anything at all has the persistence conditions defended by animalists. I'll consider that claim in the next section. For now I'll conclude that Olson's locomotor argument shouldn't worry proponents of the psychological criterion and the view that Lockean personhood characterises a substantial kind.

8.8 DOES BIOLOGICAL DEATH ENTAIL CEASING TO EXIST FOR ANY OBJECTS RECOGNISED BY OUR ORDINARY ONTOLOGY?

I want to end my discussion of animalism and personalism by asking whether it is at all obvious that any of the things we register in our ordinary ontology and which have the property of being biologically alive (where this is not a property entailing any kind of psychological properties) are such that continuity of biological life is necessary for their continued existence?¹²⁷

It seems to me that we can roughly divide the biological realm into three groups of entities: Personal organisms, sentient nonpersonal organisms and nonsentient organisms.¹²⁸ Being a sentient being I take to be characterised by having a mental life of some complexity where this doesn't entail having self-conscious thoughts.¹²⁹ Personal organisms are exemplified by beings such as you and I; as our exemplar of a sentient nonpersonal organism we might take Fido my pet dog, and as an example of a nonsentient organism we can take the oak tree outside my window 'Oakie.' Now I have already made a case for the view that it is by no means obvious that personal organisms can't survive biological death, i.e. the loss of biological life. I have defended the view that it is neither a necessary nor a sufficient condition for the persistence of a personal organism that its biological life continues.

What about nonpersonal sentient organisms such as dogs and dolphins? It seems clear to me that the transplant intuition carries over to cases involving merely sentient beings. A consideration in favour of the claim that nonpersonal sentient beings such as dogs and dolphins don't have biological continuity as a necessary condition for their survival is the transplant

¹²⁷ This question is also raised by Unger in Unger 2000.

¹²⁸ I do not pretend this to be in any way an authoritative taxonomy. It is simply a way for me to present a line of argument.

¹²⁹ This is admittedly a rather loose description of 'sentient being', but I think that all I need here is to point out that by 'sentient being' I mean beings which are subjects of mentality broadly speaking, and those which aren't are nonsentient.

consideration applied to the case of my pet dog Fido. Presumably the mental properties of Fido are realised by Fido's brain states. In case we take Fido's brain and transplant it into the empty head of Fido's brainless canine twin, we presumably end up with a canine creature, call it 'Gido', which recognises 'his masters voice', and when coming home after the operation with Gido, Gido manifests the habitual and familiar behaviour of Fido, e.g. it knows where to find food etc.¹³⁰ If we accept the transplant intuition with respect to the Fido/Gido case, we don't think that biological continuity is either necessary or sufficient for dog identity over time. Generalizing, I think it is fair to assume, that we take sentient beings in general, and not just persons, to go with the continuation of their mentality.

Take finally organic nonsentient beings such as Oakie the oak tree in my garden. It is by no means uncontroversial that Oakie ceases to exist when it ceases to be biologically alive. One reason for thinking that it does survive is that we certainly often talk that way. We talk about dead trees as trees, the very same ones, that were once alive. It is not obviously false that I can say of the tree I climbed as a child that it was alive then, but now it is dead, despite the absence of continuity of biological life. In general one might suspect that the metaphysically important distinction to make between material objects is not so much between the organic and inorganic ones but between those which are sentient and those which aren't.

Now as I have already said, I am by no means clear about whether I agree with Olson and other animalists that there are beings such as trees and oysters for which the continuation of biological life is essential, but I think that biological continuity is a very good candidate for what it takes for them to persist. What I have pointed out here is that just as the animalist claims that none of the material beings acknowledged by our ordinary ontology can

¹³⁰ See Carter 1990 for a different response to this case.

plausibly be said to have psychological persistence conditions, a pretty good case can be made for the claim that while all sentient beings, personal as well as nonpersonal ones, may plausibly be said to have psychological and nonbiological persistence conditions, it is not at all obvious that there are any beings which have the persistence conditions of the biological criterion. Thus it is not clear to me that accepting materialism and endurantism in any way entails that there are beings for which the biological criterion of identity over time is true.

8.9 CONCLUSION

In this chapter I have considered *personalism*, a view according to which our paradigm examples of persons are identical with particular human organisms and yet have psychological persistence conditions. Insofar as one is interested in the prospect of reconciling the view that we are identical to human organisms or animals with the psychological criterion, one rejects both the coincidental and the animalist lines of reasoning. The former denies the numerical identity of persons and human animals; the second denies the psychological criterion.

Personalism faces an immediate problem since the histories of human organisms and persons seem to come apart in various cases such as the Fetus and Vegetable Cases and the BST Case. I considered Olson's objection to personalism that its description of these cases is implausible bordering to the absurd. I suggested that the immediate implausibility of personalism may be alleviated by pointing out that the general form of the considerations favouring the animalist view may be applied to the personalist's case and show personalism to be a genuine alternative.

I then went on to consider van Inwagen's claim that personalism is committed to a metaphysical impossibility. I argued that van Inwagen's argument can be met by the personalist, but it highlights that personalists

still need to provide some account of why personal identity *dominates* animal identity when the two come apart as they do in the problem cases. Burke's WRP criterion for sortal dominance was presented and criticised for not yielding clear-cut decisions in a whole range of cases.

Finally I have criticised Olson's locomotor argument against the plausibility of taking 'person' to be a substance concept. The fact that it, like 'locomotor,' is functional and allows for multiple realizability doesn't seem to entail that it, like 'locomotor,' is a bad candidate for picking out a substantial kind. On the other hand, I think there's reason to be sceptical about the plausibility of the animalist claim that 'human organism' and 'organism' somehow pick out a relevantly homogenous kind of beings for it to be a substantial kind.

My overall conclusion about personalism is that while it can be shown that it is a viable alternative to animalism, it hasn't been shown that when we have a human person it is the psychological criterion which tells us how that person persists and not the nonpsychological biological criterion advocated by the animalist. Whether to accept personalism or animalism hasn't been decided in either position's favour. I take it that this is a little victory for the personalist, since that view has hardly been taken seriously as a way of defending the psychological criterion, whereas animalism has enjoyed increasing popularity in recent years. However, my overall conclusion is that neither the biological nor the psychological criterion has a claim to dominate in the multiple-kind fission case, but that both are plausibly taken to secure the existence before Holmes fission of both Hanimal and Hannibal. They are, so to speak, equally good continuers of Holmes and thus both have a claim to have been there all along.

In the final two chapters I will consider ways of maintaining that in cases of multiple-kind fission, the distinct and non-identical fission offshoots can

truly be said to exist all along. I begin by taking up the suggestion that they coincide before the fission.

9. The Constitution View

9.1 INTRODUCTION

The starting point for the present chapter is that there is a time, t_2 , at which we can identify two distinct non-identical material individuals, Hanimal and Hannibal, and tracing them back in time to t_1 along paths of continuity plausibly thought to constitute their persistence, we end up with what appears to be a single material individual, Holmes. In chapter 8 I argued that Hanimal and Hannibal's continuity relations to Holmes give them an equal claim on being Holmes at t_1 , and I submitted that it will be implausible to take either one of them to have a better metaphysical claim on being Holmes than the other.

In this chapter I will consider the claim that we should conclude that both Hanimal and Hannibal are located where Holmes is located at t_1 , but acknowledge that since they are numerically non-identical at t_2 , they are numerically non-identical at t_1 as well, assuming the eternality and necessity of identity. I begin by introducing a view I call 'coincidentalism' and present what I refer to as 'too many'-style arguments against it. I then consider Lynne R. Baker's recent attempt to develop a notion of constitution which avoids the problems facing coincidentalism, and end by canvassing my main concern about her constitution view.

9.2 COINCIDENTALISM AND 'TOO MANY'-STYLE OBJECTIONS

At t_2 Hanimal is a mindless human vegetable (thus not a person) and its persistence is plausibly constituted by biological continuity. Hanimal is a *biological continuant* and essentially a human organism. At t_2 Hannibal is made up of bionic materials and his persistence is constituted by psychological continuity. Hannibal is a *psychological continuant* and essentially a person. At t_1 there is what we'd usually take to be a single material individual, Holmes, who is both a biological predecessor of

Hanimal and a psychological predecessor of Hannibal. If we don't want to accept that identity is temporary, and thus allow for the possibility that the case should literally be described as one in which Hannibal is identical to Hanimal at one time and not at another, we might be tempted to suggest, that where we find Holmes we find *both* Hanimal and Hannibal. Hanimal and Hannibal may be said to be co-located at t_1 without being identical at t_1 . Thus there is no single thing which 'Holmes' denotes. I will call this view 'coincidentalism.' A standard definition of coincidence is this:

Coincidence Material individuals x and y coincide at time t only if (1) x is numerically non-identical to y ,¹³¹ (2) x and y are not of the same kind, and (3) the place wholly occupied at t by x is identical to the place occupied by y at t .¹³²

Coincidence of material individuals is not to be thought of as coincidence of a mass of matter and the individual it makes up, or the parts of an individual and the individual they compose. How to think of the relation between *the gold* making up a coin and 'the gold atoms' composing a coin is an interesting question, but here I am considering cases where the coincidence is of e.g. a lump of gold and a coin, where these are distinct individuals falling under count nouns such as 'lump of gold' and 'coin.'

Say that material individuals are *perfect* physical duplicates when they have the exact same intrinsic physical structure *and* are composed of the very same microphysical parts *and* occupy the same place at the same time. This is the

¹³¹ Later we shall see how Baker suggests that we should disambiguate the non-identity claim and distinguish between two ways of being non-identical, 'constitution' and 'numerical non-identity.' This is why I explicitly prefix 'numerical' to 'non-identity.' The constitution view thus differs from coincidentalism in working with a notion of 'non-identity' hovering between numerical identity and numerical non-identity. More on that later. For the time being 'non-identity' means 'numerical non-identity.'

¹³² See e.g. Burke (1992). I will assume that when material individuals are said to occupy the same place at the same time they are composed of the same microphysical parts.

case when we have coincidence. At t_1 Hanimal and Hannibal are perfect physical duplicates in this sense.

When presented with the coincidence view that we should take the possibility of Holmes's multiple-kind fission to show that Hanimal and Hannibal are perfect physical duplicates at t_1 and yet non-identical, one might naturally ask how they so to speak manage to be non-identical when they are perfect physical duplicates of each other?

The obvious reply is that they manage to be non-identical because there is *another* time at which they both exist and are non-identical; assuming the necessity and eternality of identity, it simply follows from the fact that at t_2 they both exist, are non-identical and stand in identity constituting relations to Holmes at t_1 , that they are both located where we take a single individual, Holmes, to be located at that time. What grounds their non-identity at t_1 is thus contemplation of different possible futures for the individual ordinarily picked out as Holmes, and acknowledging such different possible futures, we realise that 'Holmes' is an ambiguous name, which names continuants of different kinds. By 'Holmes' we might mean, 'the individual which may end up as a human vegetable and which was once a foetus,' or 'the individual which may gradually change into a bionic entity as long as its psychological life is continued.' It becomes relevant to ask whether by 'Holmes' one means Hanimal or Hannibal.

When I discussed the multiple occupancy thesis, I complained that it was not determined by *present* facts, whether there was one or many persons sitting in my chair. Whether there would be a fission in the future or not would make a difference as to how many persons were sitting in my chair at now. This consequence together with the worry that in other possible worlds where e.g. Brown dies just before fission there would only have been one person all along, made me submit that the multiple occupancy thesis is

unacceptable. Furthermore I found it to be a mystery how to distinguish one person from another before fission, since in such same-kind fission scenarios, not even modal and historical properties would distinguish them.

My main worry about the coincidence approach to the multiple-kind fission case is that if we accept coincidence of perfect physical duplicates, which are different kinds of things with different persistence conditions, then it also commits us to accept coincidence of things of the same kind with the same persistence conditions. As a matter of fact, if coincidence has anything going for it, I suspect it will come across as more plausible when taken to be coincidence of *the same* kind of things given the fact that coinciding things are perfect physical duplicates. Let me present my worry.

Take the place where Holmes is located at t1. If we analyse 'Holmes' as ambiguously naming a human animal, Hanimal, and a person, Hannibal, coinciding as perfect duplicates in that location, it is reasonable to ask why a perfect physical duplicate of a person like Hannibal isn't also a person: 'a psychological continuant kind of thing'? And why isn't a perfect physical duplicate of Hanimal a biological continuant kind? These are questions I think faces adherents of coincidence.¹³³

In general it is, I think, very plausible to hold that perfect physical duplicates have many of their properties in common. If Hanimal has hay fever and there is a perfect physical duplicate of it, then that duplicate will have hay fever too. And, assuming that mental properties supervene on intrinsic physical properties, if Hannibal thinks, 'I have to mow the lawn soon,' then a perfect physical duplicate of him would seem to think so too. That is certainly a view which proponents of the psychological criterion accepting the Lockean concept of personhood will seem to accept. On the Lockean

¹³³ This line of criticism against coincidence is given a powerful presentation in Olson (1997, p. 97 ff.)

concept of a person, any being which has rational and self-conscious thought is a person, and the claim of the psychological criterion is that such beings have certain psychological persistence conditions. So it would seem as if perfect physical duplicates of a person will also be persons having all the mental properties essential to having personhood, but if they are persons, they must also have the psychological persistence conditions of persons. This is problematic if the perfect physical duplicate is taken to be a human organism with nonpsychological persistence conditions, since we then seem to have to say that not all persons have psychological persistence conditions, or else that the fact that a human organism has the intrinsic properties characteristic of Lockean personhood isn't sufficient for it being a person, since it must also have certain relational properties. But personhood is supposed to be an intrinsic property of an individual.

Now ordinarily when we say that two non-identical individuals both have hay fever, we take there to be two distinct hay fevers being had, one for each individual. And if you think 'I have to mow the lawn tomorrow' and I think that 'I have to mow the lawn tomorrow,' then there will have to be two thoughts being thought, one by you another by me. Access to one's first-person thoughts is the privilege of one and only one person, oneself. Thus if you and I are identical there is a single token thought being thought and one hay fever being had when one of us thinks or sneezes. If we are not identical but two distinct persons, then there will be two token thoughts and two token hay fevers.

Say that Hannibal and Hanimal coincide at t_1 . They are then perfect physical duplicates at t_1 . But then they are non-identical at t_1 and thus have different tokens of the same type of thoughts and hay fevers etc. Alternatively there will be two non-identical individuals having a single hay fever and a single first-person thought, which is in my view even harder to believe. But now we seem to have two persons and two animals and two hay fevers and two first-

person thoughts that 'I have to mow the lawn tomorrow' occurring at t_1 , where we take there to be a single human person called 'Holmes' who has one hay fever and one first-person thought etc. These considerations in my view reveal that the coincidence suggestion has completely implausible consequences. The non-identity of coinciding individuals, which are nevertheless perfect physical duplicates, allows for 'too many'-style objections. The criticism can of course be stepped up a bit by considering possible worlds increasing the number of continuants coinciding where we find Holmes, thereby generating perhaps hundreds of hay fevers being had and first-person thoughts being thought by hundreds of human organisms and persons.

Here is the line of reasoning I ascribe to coincidentalism when confronted with the multiple-kind fission case: At t_2 we acknowledge that Hanimal and Hannibal are numerically non-identical. They are *two* separately existing individuals. Hanimal is essentially a human organism with biological persistence conditions and Hannibal is essentially a person with psychological persistence conditions. We can trace both Hanimal and Hannibal along the continuity paths constitutive for their persistence to the location where we locate Holmes. Numerical identity and non-identity is an eternal relation, so if Hanimal and Hannibal are numerically non-identical at t_2 , they are numerically non-identical at all times at which they exist.

We now get the following coincidentalists conclusions:

Co-location Where we locate Holmes at t_1 , we find both Hanimal and Hannibal.

Non-Identity Where we locate Holmes at t_1 , there are really two numerically non-identical individuals.

Kind-Multiplicity Where we locate Holmes at t_1 there are really two different kinds of individuals; something which is essentially a person with psychological persistence conditions *and* something which is essentially a human organism with biological persistence conditions.

Ambiguity The name 'Holmes' is ambiguous.

Here is my challenge to the coincidentalists: At t_1 , when Hanimal and Hannibal are claimed to be numerically non-identical, they are perfect intrinsic physical duplicates. But then it seems that at t_1 we have to ascribe the properties realised by the intrinsic physical structure of the one to the other and *vice versa*. Thus if Hanimal has hay fever at t_1 so does Hannibal, and if Hannibal thinks 'I should mow the lawn tomorrow', so does Hanimal. And furthermore, assuming that Hanimal is a human organism due to his internal physical structure, so is Hannibal, and assuming Hannibal's personhood to be realised by his intrinsic physical structure, then Hanimal also has personhood.

Let's say that the following set of properties is realised by the internal physical structure shared by Hanimal and Hannibal at t_1 :

- The property of having a hay fever.
- The property of thinking, 'I should mow the lawn tomorrow.'
- The property of being a human organism.
- The property of being a person.

Now when two numerically non-identical individuals x and y have a hay fever or think a thought, they don't have the same token hay fever or think the same token thought. And if numerically non-identical individuals x and y both have the property of being a person, then we have two numerically non-identical persons. The same goes for human organisms x and y .

Assuming that non-identical individuals do not share the same token properties, the claim that Hanimal and Hannibal coincide at t1 appears to entail that:

- There are two token hay fevers where Holmes is.
- There are two token 'I should mow the lawn tomorrow' thoughts.
- There are two human organisms, one of which has psychological persistence conditions and isn't essentially an organism.
- There are two persons one of which has biological persistence conditions and isn't essentially a person.

But these are unhappy consequences of the coincidental response to the multiple-kind fission scenario. The claim that Hanimal and Hannibal are non-identical at t1 entails a massive multiplication of token properties of the same type, including the properties of being a person and being a human organism. The result of the coincidental description is that when Holmes has a hay fever and begins to sneeze thinking 'I should mow the lawn tomorrow,' there is a plurality of numerically non-identical individuals sneezing a plurality of sneezes due to the instantiation of a plurality of hay fevers and I-thoughts. And on top of that it also turns out that at t1, we have an individual, Hanimal, which is a person and which does not have the persistence conditions we think persons have. And we have a human organism, Hannibal, who is a human organism and does not have the persistence conditions we take human organisms to have.

But all this would seem to undermine the central assumptions which generate the problem of multiple kind fission in the first place, namely that beings which are persons have certain psychological persistence conditions, and that human organisms have biological persistence conditions.

Take Hanimal: at t1 Hanimal apparently has all the qualitative properties characteristic of personhood, but if we allow him to *count* as a person, i.e. to be a person *numerically speaking*, we seem to end up with a person at t1 who isn't essentially a person and who doesn't have the persistence conditions of a person. And similar problems arise with respect to Hannibal's status as a human organism at t1. As will become clear I think we should take Hannibal and Hanimal to be identical at t1 and distinct at t2.

As we shall see, Baker has presented what she calls 'the constitution view' by which she tries to mitigate the implausible consequences of the non-identity of coinciding individuals by characterising coincidence as a relation of 'constitution.' I will come to Baker's constitution view shortly. Let me first briefly discuss Shoemaker's suggestion that there aren't two persons when Hanimal and Hannibal coincide, since human animals simply can't have mental properties essential to personhood at all due to their persistence conditions.¹³⁴

9.3 SHOEMAKER'S VIEW: ANIMALS CAN'T THINK

Shoemaker (1984, pp. 92-97, 1999) has argued on the basis of a functionalist account of mental properties that when a human organism coincides with a person this doesn't entail that both will have mental properties. A subject of mental properties must have certain persistence conditions which make it fit for being the subject of the successor states which are relevant for their individuation and, as the problem case reveals, Hanimal might not be the subject of the relevant successor states of Hannibal's mental properties at t1.

¹³⁴ Another way of dealing with the 'too many'-style problems is suggested by Noonan 1998 and 2003, but I leave Noonan's view, which revises the traditional semantics of the first-person pronoun, aside. On Noonan's suggestion, the coinciding human animal thinks 'I-thoughts' on behalf of the coinciding person, not on behalf of itself. The problem of there being too many persons thinking my thoughts features prominently in Olson's criticism of coincidentalism (1997, p. 97 ff.) as 'the problem of the thinking animal.'

Shoemaker's argument is interesting but I don't have space to consider it in detail here. Let me briefly mention my reasons for finding it unattractive.

Firstly, it seems hard to accept that there aren't any organisms of any kind which have mental properties. Not only are we are tempted to think that our coinciding human animals can have mental properties, but also that our pets have mental properties, yet, on Shoemaker's proposal, no organism can be the subject of mental properties. What is true of many animals is that they coincide with something else, even something material, that has those properties.

Secondly, it is also hard to believe that there may be a difference in the properties realised by perfect physical duplicates given the assumption that the mental is realised by the physical. Now Shoemaker accepts functionalism about the mental and thus multiple realizability of mental states, but invoking his functionalist account of the nature and individuation of properties in general, he argues, that on such an account, there are physical properties, the properties realizing mental states, 'realiser properties', which are not had by both of the coinciding individuals. So there are physical properties not had by coinciding material individuals. Maybe this argument will work if one accepts the functionalistic account of properties in general, but I won't pursue the issue of functionalism in general.

Thirdly, Shoemaker says that on the functionalist account of mental properties we find the same person where we find appropriate successor states. I.e. there is nothing more to personal identity than having a mental state which is an appropriate successor state of a mental state had by a person at some past time. Hence, if P1 has M1 at t1, and P2 has M2 at t2, and M2 is an appropriate successor state of M1, then P2 is the same person as P1. Appropriate successor states are, as Shoemaker emphasises (1984, p. 93, 1999), *states of the same person*.

I find that there is something that doesn't add up with regard to Shoemaker's functionalist account of personal identity and his approach to the standard fission case: Let's say that Brown forms an intention to run a marathon. This is a mental state and thus, according to the functionalist, it is individuated by the relevant successor states it gives rise to. Now it seems quite clear to me that in the case of Brown's fission, both Lefty and Righty will have mental states functionally relevant for the individuation of Brown's intention to run a marathon. As a matter of fact, this is part of the reason why I am so keen to maintain that they exist before Brown's fission, and this seems to me to be the insight of Locke's original claim that sameness of consciousness is what makes for identity of person. Lefty and Righty have appropriate successor states to Brown's mental states. But then they should be taken to be *the same person* as Brown. (This is what I argued we can make sense of by adopting the temporary identity view.)

But this seems to conflict with Shoemaker's own description of ordinary fission discussed in chapter 5 according to which neither of Brown's two psychological continuers is identical with Brown. Despite the fact that the mental states of Lefty and Righty certainly seem to be appropriate successor states of those of Brown in the mono-cases, Shoemaker rejects the claim that they are identical with Brown.

If Shoemaker is right to reject that Lefty and Righty are identical with Brown, then he will also have to say that Lefty and Righty don't have appropriate successor states to Brown's states. But now it begins to look as if it isn't just states of the same person which are relevant, but also those of other persons. In the fission case there is arguably no intrinsic difference in Brown's relations to each of his offshoots compared to the mono-cases. What seems to rule out e.g. Righty's mental states as appropriate in the standard fission case is the fact that there is some *other* person, who has certain mental states, but

then someone *else's* mental states *are* relevant for determining whether Righty's mental states qualify as appropriate successor states or not.

So much for Shoemaker's view. I don't pretend to have given a full-blown rejection of it, but I don't find it to be the most promising way to go. If we can, we should allow for Hanimal to have mental properties when coinciding with a person.

9.4 BAKER'S IDEA OF CONSTITUTION

Recently Lynne R. Baker (2000) has developed and defended what she calls 'the constitution view' as a way of cashing out the intimate relation between non-identical but coinciding individuals. I will focus my attention on her view in the remaining sections of this chapter and see if it provides the conceptual machinery for dealing with my criticism of coincidentalism.¹³⁵

Baker introduces the need for a notion of constitution thus:

'We need constitution to be similar to identity in order to account for the fact that if x constitutes y , then x and y are spatially coincident and share many properties; but we also need constitution to differ from identity in order to account for the fact that if x constitutes y , then x and y are of different kinds and can survive different sorts of changes. (...) I want to make sense of constitution as a third category, intermediate between identity and separate existence' (2000, p. 28-29)

More specifically the constitution relation is a *contingent* relation between non-identical *individual things*, not between things and their parts or between things and the masses of matter that make them up (2000, pp. 33-34).

¹³⁵ Let me note here that 'coincidentalism' and 'the constitution view' are not the same view since they operate with different ways of describing the relation between Hanimal and Hannibal. On the constitution view the relation is one of *constitution* which is different from *coincidence* in that it doesn't take Hanimal and Hannibal to be *numerically* non-identical, but to stand in another relation of non-identity: constitution. I present the notion of constitution in this section.

Baker describes the constitution relation specifically in this way:

Where *being an F* and *being a G* are distinct primary-kind properties, it is possible that an F exists without there being any spatially coincident G. However, if a F is in G-favourable circumstances, then there is a new entity, a G, that is spatially coincident with the F but not identical to it (2000, p. 42).

Take the case in which a piece of granite is taken from a quarry loaded on a truck and put in a square and thereby comes to constitute a monument. In that case it is possible that there is a time at which the piece of granite doesn't coincide with anything, but necessarily, when in 'monument-favourable circumstances,' the piece of granite comes to coincide with a monument, a thing which didn't exist before. If these conditions are satisfied the relation between the piece of granite and the monument is the relation of constitution. It is important to note here that the piece of granite comes to constitute a monument without any intrinsic change. The changes which result in the generation of a monument, a new additional individual with different essential properties and persistence conditions than the piece of granite, are relational.¹³⁶

In the case at hand Hanimal and Hannibal are of distinct kinds; one is essentially a human organism, the other essentially a person. At time t1, Hanimal is in 'person-favourable circumstances' in that his intrinsic and environmental circumstances are conducive for the realization of personhood (cf. Baker 2000, p. 96). Hanimal constitutes Hannibal at t1 due to the fact that necessarily, when a human organism is in 'person-favourable circumstances'

¹³⁶ When Baker puts emphasis on the distinction between intrinsic and relational changes this is because she is keen to allow that for some kinds of things, it is essential to them that they have certain relational properties. In the case in which a human organism comes to or ceases to constitute a person, there will, I presume, have to be certain intrinsic changes. But the generation of a piece of rock on a beach in Australia which is exactly like a statue of Elvis Presley fashioned in Memphis, Tennessee, will not be a statue due to not being in 'statue-favourable circumstances.'

there will be a person coinciding with it. Still, for them to stand in the relation of constitution at t_1 they must not only coincide at that time. It must also be the case that possibly a human organism exists without coinciding with anything. Hanimal could exist at t_1 and not constitute a person, hence there is a person constituted by Hanimal at t_1 .

In general Baker's idea of constitution is intended as a way of accounting for the way in which many things of different kinds coincide without being identical. Examples abound: pieces of paper and dollar bills, pieces of metal and stop signs, DNA molecules and genes, flags and pieces of cloth, and so on (cf. 2000, p. 27). What these cases have in common is that when an object of a certain kind enters certain circumstances, a new thing of a different kind with e.g. different causal properties is generated, even in the absence of intrinsic change in 'the old thing'. But this doesn't mean that 'the old thing' ceases to exist. The piece of paper comes to constitute a dollar bill when entering a dollar economy. The same idea is suggested with respect to human organisms and persons: When human organisms are in intrinsic and environmental circumstances conducive for the development of a first-person perspective or personhood, they come to constitute a person.

So what happens to say a piece of marble or a human organism when it comes to constitute a statue or a person respectively? According to Baker, 'The identity of the constituting thing is submerged in the identity of what it constitutes. As long as x constitutes y , y encompasses or subsumes x ' (2000, p. 33). And she writes: 'If x constitutes y at a certain place and time, then there is a unified individual at that place at that time, and the identity of that individual is determined by y ' (2000, p. 33). What this means is that when x constitutes y , the identity of the unified individual is that of y , and yet x may exist independently and separately from y at other times than those at which x constitutes y (cf. 2000, p. 33). It is the constituted thing, e.g. the statue, which has *ontological priority* in that it has causal powers which the piece of

marble wouldn't have had if it didn't constitute a statue at that time (cf. 2000, p. 33).

So Baker's project is more ambitious than 'just' claiming that at t_1 Hanimal and Hannibal coincide. She wants their coincidence to be understood as constitution, a relation such that 'when x constitutes y , there is a unitary thing – y , as constituted by x – which is a single thing' (2000, p. 46).¹³⁷

So much for the central elements in Baker's notion of constitution. An important motivation for developing a notion of constitution 'intermediate between identity and separate existence' is to make room for coinciding individuals to share properties such as having a hay fever and thinking that the lawn has to be mowed. The idea of constitution is accompanied by a distinction between having a property derivatively and nonderivatively which is supposed to make sense of this idea without invoking the notion of identity.

9.5 DERIVATIVE AND NONDERIVATIVE PROPERTIES

On the basis of the relation of constitution, Baker draws a distinction between having a property derivatively and having it nonderivatively. A nonderivative property is a property an individual has independently of standing in any constitution relations, whereas a derivative property is a property an individual has in virtue of standing in a constitution relation to an individual which has that property underivatively.

Baker writes: '[At time t] x has H derivatively if and only if x 's having H at t depends wholly on x 's being constitutionally related to something that has H

¹³⁷ I should note here that Baker does line up a more technical definition of constitution (2000, pp. 42-43), but I think the more informal presentation just given will do, and that engaging in the formal definition will introduce more issues than necessary for assessing Baker's suggestion about how to deal with the relation between Hanimal and Hannibal before the multiple-kind fission.

at *t* independently of its being constitutionally related to *x'* (2000, p. 47, see also p. 51). And also she points out that: 'The fact that *y* has a property at *t* derivatively just *is* the fact that at *t* *y* is constitutionally related to some *x* that has the property at *t* independently of being constitutionally related to *y* at *t'*' (2000, p. 55).

To have a property 'independently of constitution relations' is defined thus (2000, p. 49):

x has H at *t* independently of *x*'s constitution relations to *y* at *t* =df

(a) *x* has H at *t*; and

(b) Either

(1) (i) *x* constitutes *y* at *t*, and

(ii) *x*'s having H at *t* (in the given background) does not entail that *x* constitutes anything at *t*.

or (2) (i) *y* constitutes *x* at *t*, and

(ii) *x*'s having H at *t* (in the given background) does not entail that *x* is constituted by something that could have had H at *t* without constituting anything at *t*.

From this we can see that a human organism has the property of being a human organism independently of its constitution relations to a person, that a person has the property of being a person independently of its constitution relation to a human organism, and that persons are human organisms, and human organisms are persons, derivatively and not independently of standing in any constitution relations.¹³⁸

¹³⁸ In Baker's own words: you and I are animals derivatively. The organisms that constitute us are animals independently of their constitution relations to us, and it is not the case that you or I is an animal independently of our constitution relations to the animals that in fact constitute us. Conversely, as is equally clear from the definitions, the organisms that constitute us are organisms nonderivatively, and those organisms have first-person perspectives derivatively (2000, pp. 97-98).

The distinction between derivative and nonderivative properties allows for individuals standing in the relation of constitution to share many of their properties. Baker writes:

(...) With the idea of having properties derivatively, I need not deny that the animal that constitutes me is a person; indeed I insist on it. That animal is a person derivatively: It is a person in virtue of constituting something that is a person.' (...) My point is that when an animal has a first-person perspective or the capacity for one, a new entity comes into being: a person. The person is not a duplicate of the animal; rather the person is constituted (in the sense defined) by the animal. There are not two animals; there are not two persons. There is one person constituted by an animal' (2000, p. 98).

On the constitution account an individual a can truly be said to have the property of being a person if and only if a has the property of being a person either derivatively or nonderivatively. In general: '(...) for any property G , if x has G at t , then $\exists y$ (y has G at t nonderivatively and either $x=y$ or x is constitutionally related to y at t)' (2000, p. 54). Baker now suggests that her notion of constitution allows us to read the 'is' of constitution thus: "If ' a is (an) F ' should be read as ' a constitutes (is constituted by) something that is (an) F ,' then a has the property of being (an) F derivatively" (2000, p. 54).

In the present context the important feature of constitution is that it allows that when x constitutes y , there are *some* properties of x which x possibly has without standing in any constitution relation, and y can correctly be said to have those properties derivatively at the times y is constituted by x . We don't need to multiply properties just because we allow that x and y are non-identical when they coincide. As a matter of fact, when at t y has a property derivatively, this just *is* the fact that y is constitutionally related to some x which has that property at t independently of being constitutionally related to y at t .

Let's suppose that at t1 Hanimal has a blood alcohol level of 1.2 percent. This is a property Hanimal has independently of the fact that Hanimal constitutes Hannibal at that time.¹³⁹ *Hannibal-as constituted-by-Hanimal* will then have the derivative property of having a blood alcohol level of 1,2 percent at t1 (2000, p. 55). Constitution is a unifying relationship between non-identical individuals which generates *a unified thing*, and the ontologically significant constituted object *encompasses* the constituting object with the result that there are some properties which the ontologically significant object can be said to have derivatively, but it nonetheless has those properties.

However not all properties can be had derivatively. What Baker calls 'alethic', 'constitution/identity/existence', 'rooted outside the times they are had', and 'hybrid' properties all fall outside the realm of the derivative/nonderivative distinction (cf. 2000, p. 48-49). Thus the property of being a person is importantly different from *being identical to* a person. For *x* to be *identical to* a person, there must be a *y* which is a person nonderivatively and $x = y$. But *x* can have the property of being a person either by being identical to a *y* who is nonderivatively a person, or by standing in a relation of constitution to a *y* who is underderivatively a person. It is only if *x* is identical to a *y* which is nonderivatively a person that *x* has the property of being a person underderivatively.

If all this is correct, what we refer to as 'Holmes' can be said to be a 'single constituted thing': *Hannibal-constituted-by-Hanimal*, a person-constituted-by-a-human organism, a unified individual. After the multiple-kind fission, Hannibal and Hanimal both still exist, but they don't stand in the relation of

¹³⁹ The independence here need not be causal. It might be that there is a causal link between Hanimal's having the blood alcohol level it has and the fact that it constitutes a person, e.g. it is due to constituting a person who went to a party that Hanimal now has the blood alcohol level it has. But it could have it now without constituting a person (some might say it does...).

constitution any more. The virtue of constitution when it comes to explaining the relation between Hanimal and Hannibal is that it is a temporary relation, and that it allegedly provides the conceptual tools for dealing with the hard questions asked about unwanted multiplication of properties when non-identical individuals coincide. In the next section I will criticise Baker's constitution view.

9.6 COUNTING BY CONSTITUTION – CRITICISM OF THE CONSTITUTION VIEW

As should be clear from my presentation of Baker's constitution view, that view is an attempt to flesh out a relation which is able to accommodate the non-identity of Hanimal and Hannibal at t_1 without entailing the 'too many'-style problems challenging the coincidental proposal. The constitution view is supposed to provide a basis for *allowing* co-location of non-identical individuals of different kinds with different persistence conditions, which are qualitatively identical at the time of coincidence, while *avoiding* multiplication of numerically non-identical individuals with numerically non-identical token properties of the same type including kind properties and persistence conditions.

In order to avoid there being *too many* persons and *too many* token properties instantiated where we locate Holmes, the suggestion is that since Hanimal isn't essentially a person and doesn't have the persistence conditions of a person, it should not really *count* as a person. *Numerically speaking* Hanimal is not an additional person at t_1 . Instead it may be said to have that property derivatively at that time. And similarly, Hannibal may be said to have the property of being a human organism derivatively without being an additional human organism. And when Hanimal at t_1 appears to be in a state of having a first-person thought, that state doesn't *count* as a token first-person thought in addition to the token thought entertained by Hannibal. Nor would Hannibal's hay fever count as a hay fever *in addition to* Hanimal's

hay fever. Hannibal would have a hay fever derivatively during the time he is constitutionally related to Hanimal.

On the constitution view individuals count as two only when they are *numerically* non-identical, as is the case with Hannibal and Hanimal at t2. But at t1 they are not numerically non-identical, they are 'constitutionally non-identical.' This entails that they are qualitatively alike, but not in the fashion of *duplicates* with the unwanted duplication of properties leading to the 'too many'-style worries. We know from the very setup of the multiple-kind fission case that there is something which is essentially a human organism located where Holmes is at t1 and also that something which is essentially a person located where Holmes is at t1, given the identity-constituting continuity relation assumed to hold between Holmes and Hanimal and Hannibal respectively.

Baker's suggestion is that instead of having to admit an unwanted addition of individuals and properties at t1, we may take co-located individuals of different substantial kinds to stand in a relation of constitution, and that properties may be had in different ways at a given time depending on what sort of thing something essentially is. This is cashed out in terms of the notions of having properties *derivatively* and *noderivatively*.

It is important to make clear that having a property derivatively is no less having it than having it nonderivatively. Thus suppose Hannibal is arrested for drunk-driving. The fact that he derivatively has a blood-alcohol level above what is legal doesn't make him less of a drunk driver. He can't defend himself by saying, 'It isn't really me who has too high a blood-alcohol level to drive, it is the human animal constituting me. I am only derivatively drunk, it is really the animal constituting me which is drunk.' And the human organism can't avoid being charged with the offence by (derivatively) pointing out that it is only derivatively a person who has responsibility for

his actions, and that it is really not *it*, but the person it constitutes, who should be charged.

While I think the constitution view does provide a way to steer clear of the 'too many'-style problems, I am unhappy about it in two respects. One is that constitution allows for counting objects as one when they, numerically speaking, are not. What counts as two at one time may count as one at another. This is because what is numerically non-identical at one time may not be numerically non-identical at another time, and hence not count as a plurality at that time. But it seems to me an odd way of defending the necessity and eternality of identity to give up on the eternality and necessity of numerical non-identity. Baker seems to think that we should give priority to maintaining that since Hanimal and Hannibal are definitely not identical at t_2 , there is no time at which they can be identical. But why not argue instead that since Hanimal and Hannibal are definitely one and the same at t_1 , there is no time at which they can be numerically non-identical, and thus we should reject the claim that at t_2 they are correctly said to be numerically non-identical and numerically two? This leads me to my second point.

There is I think an overlooked alternative to the constitution view's way of describing the multiple-kind fission case which is grounded on the same rationale as the constitution view, namely that we should preserve the necessity and eternality of numerical non-identity. Hence if there is a time at which Hanimal and Hannibal aren't numerically non-identical, then there can be no time at which they are numerically non-identical. And if there is a time at which it seems appropriate to *count* them as *two* though they were at some other time numerically identical, they should be counted as two by some other relation than numerical non-identity.

Say that Hanimal and Hannibal are numerically identical at t_1 (i.e. they are not numerically non-identical at t_1). This entails that there is no time at

which they exist and are numerically non-identical. Hence there is no time at which they exist and numerically count as two in virtue of standing in a relation of numerical non-identity. Nevertheless there is a time at which it seems appropriate to count them as two and not to say that they are numerically identical, namely at t_2 when they are 'separately existing.' This is parallel to the temptation to say that at t_1 they shouldn't be counted as two due to their intimate relation.

Now as a parallel to constitution one might suggest that Hanimal and Hannibal instead of standing in the relation of numerical non-identity at t_2 stand in a relation of 'xidentity' which hovers between numerical identity and numerical non-identity. When individuals are 'xidentical' they can be counted as two despite not being numerically non-identical, but they are not numerically identical either since that entails counting as one.

So as soon as we go down the constitution track, we invite the view that we can say that at t_2 Hanimal and Hannibal *count as two* counting by some relation X which holds between them at t_2 , and thus we can say that they are numerically identical at t_1 and not numerically non-identical at t_2 , due to them standing in a relation X which hovers between numerical identity and numerical non-identity. Instead of suggesting that there are different ways for things to be non-identical, namely numerical non-identity and constitution, there are different ways for things to be identical, namely numerical identity and relation X, neither of which entails numerical non-identity but one of which allows us to count two instead of one.

My complaint is thus not that there is something incoherent about constitution as such. My complaint is that the rationale behind constitution doesn't in itself justify the view that the multiple-kind fission case should be described the way Baker thinks, namely as a case in which we should take Hanimal and Hannibal to be two and numerically non-identical at t_2 and

therefore deny that they are identical at t_1 defending the eternity and necessity of identity. One might as well apply the same line of reasoning and come up with the reverse result, namely that since Hanimal and Hannibal are not numerically non-identical at t_1 , they must not be numerically non-identical at t_2 .

I don't see the rationale in treating numerical identity and numerical non-identity asymmetrically in respect of eternity, and still, if one does accept Baker's asymmetrical treatment then I suppose one might as well opt for the eternity of numerical non-identity as opposed to the eternity of numerical identity. The central point I want to make is that either way seems a desperate way to try to accommodate the eternity and necessity of numerical identity or the eternity and necessity of numerical non-identity respectively, without being in a position to maintain the eternity and necessity of them both.

Furthermore, we get the result that numerical identity and numerical non-identity become separated from counting as one and counting as many. We either get the result that we can count what is by all accounts not numerically identical as one due to some other relation such as constitution, or count as two what are not numerically non-identical individuals, but 'xidentical' individuals.

All this seems to me to be highly unsatisfactory, especially as a way of maintaining the eternity and necessity of numerical identity or numerical non-identity. I don't see why we should be interested in maintaining the eternity and necessity of the one by giving up the eternity and necessity of the other.

The worry I have raised here is similar to the worry I raised with respect to the multiple occupancy view in chapter 5, when I complained that it is

unclear why we should take fission to be a case in which there are *two* persons occupying the same place where Brown is just because there are two persons after his fission. Someone defending the importance of psychological continuity for personal identity might wonder why we shouldn't accept that there is one person before Brown's fission and thus also one person after fission, despite a temptation to count two non-identical persons after the fission, if it is important to maintain the eternity and necessity of identity. It seems arbitrary to think that the distinctness of the post-fission individuals has a better claim on being relevant for what facts to take to obtain pre-fission, than the other way around. In this way Baker's constitution view is vulnerable to problems similar to those presented to the multiple occupancy thesis.

Let me end my criticism by noting that the constitution view doesn't help the proponent of the psychological criterion much when it comes to the standard fission cases. Distinct individuals of the same kind can't by its very definition stand in the relation of constitution. So even in case it could explain the relation between coinciding individuals of distinct kinds with different persistence conditions, this would not help in putative cases of same kind coincidence.

9.7 CONCLUSION

In this chapter I've criticised the coincidental approach to the multiple-kind fission case for being subject to 'too many'-style objections. I then presented Baker's notion of constitution, and while I think that notion does manage to provide a coherent way of describing the relation of Hanimal and Hannibal at t_1 when they are co-located, I think it is generally unsatisfactory as a way of maintaining the eternity and necessity of identity and entails an unwanted discrepancy between counting as one and being numerically one.

In my view there is a much more appealing way of dealing with the multiple-kind fission scenario than adopting the constitution view which allows us to maintain that we count by numerical identity and numerical non-identity, and which has already proved successful with respect to the standard fission cases. In chapter 6 I argued that temporary identity is a perfectly coherent way of dealing with the standard fission case. In the next chapter I will apply it to the multiple-kind fission case and argue that it does quite well in that respect too. When in multiple-kind fission it appears as if individuals which are not identical at one time are identical at another this is because that's what they are: identity is temporary.

10. The Temporary Identity Solution

10.1 INTRODUCTION

I have submitted that in the multiple-kind fission scenario, both Hanimal and Hannibal stand in continuity relations to Holmes which are plausibly taken to constitute their identity with him. In chapter 8 I provided an elaborate discussion of whether either of the continuity relations can be argued to dominate the other and concluded that we have good reason to assume that both Hanimal and Hannibal exist at t_1 . I don't find it plausible to reject that both Hanimal and Hannibal exist at t_1 . But now the question is how they relate at t_1 , if they both exist at the same location at that time composed of the same matter? In chapter 9 I challenged the coincidental suggestion that they coincide presenting it with 'too many'-style problems. I then presented and criticised Baker's recent attempt to mitigate the problems of coincidence by developing a notion of constitution and a distinction between having properties derivatively and nonderivatively. The central notion of constitution as a non-numerical kind of non-identity was criticised for driving a wedge in between counting as one and being numerically one.

In this final chapter I briefly rehearse central elements in the way the temporary identity thesis allows us to deal with the multiple-kind fission case in a way parallel to the standard fission case. I then show how I think it may deal with both the circularity objection discussed in chapter 4 and support the view that identity matters, which was discussed in chapter 3.

Since I have already given a detailed defence for the general thesis that temporary identity is coherent against various objections that it violates the transitivity of identity and Leibniz's Law, I will not engage in discussing *whether* it is an alternative that can be appealed to. I will assume that it is and merely outline how accepting it equips us with a rather elegant way of

dealing with not only the multiple-kind fission problem, but also the other problems for the view that persons have psychological persistence conditions raised in this thesis.

10.2 TEMPORARY IDENTITY AND MULTIPLE-KIND FISSION

The problem posed by the multiple-kind fission scenario is structurally analogous to the standard fission problem discussed in chapters 5 and 6. In the standard fission case, a single person has two psychological continuers, Lefty and Righty, each with an equally good claim on being the original person Brown. The difference between the standard fission case and the multiple-kind fission case is that the 'equally good continuers' are equally good though related to the original individual by different kinds of continuity. However, just as we are keen to maintain that Lefty and Righty are there all along in the standard fission case, I have argued that Hanimal and Hannibal may justifiably be claimed to have been around all along in the multiple-kind fission case. In this I agree with the coincidentalists, but the coincidentalists take the claim that Hanimal and Hannibal exist both at t_1 and t_2 to entail that there is a relation of diachronic identity which holds between them at t_2 and Holmes at t_1 . This is what I think raises problems about how to deal with their relation at t_1 .

If we want to maintain that Hanimal and Hannibal are both identical with Holmes, we commit ourselves to the truth of the following set of inconsistent claims:

- (i) Holmes = Holmes
- (ii) Hanimal = Holmes
- (iii) Hannibal = Holmes

Identity is symmetrical and transitive so we get:

(iv) Hanimal = Hannibal.

But surely we want to deny (iv) and accept:

(v) Not (Hanimal = Hannibal).

Adopting the temporary identity thesis allows us to maintain not just the plausible view that after Holmes's fission, we have two distinct non-identical individuals both of which 'were there all along,' but also that before the fission there is a single individual, Holmes, which Hanimal and Hannibal are identical to before the fission.

As in the standard fission case, the temporary identity thesis allows us to say that the inconsistency is merely apparent and due to a lack of appreciation of the fact that the identity statements are ambiguous, when it is not made clear at what time they are claimed to be true. Indexing the identity statements to times we get:

- t1: Holmes = Holmes.
- t1: Hanimal = Holmes.
- t1: Hannibal = Holmes.
- t1: Hanimal = Hannibal.

But this doesn't conflict with what is also clearly true, namely that:

- Not (at t2: Hanimal = Hannibal).

If we accept the temporary identity thesis we also index the transitivity of identity to times thus accepting:

T-Transitivity (x)(y)(z)(t) [(t: x = y & t: y = z) → t: x = z].

Some might argue that the temporary identity solution is untenable even given T-Transitivity, presenting a version of the transitivity complaint (cf. chapter 6 section 6.1):

- (a) The animal called 'Hanimal' at t2 is identical with the animal called 'Hanimal' at t1; *and*
- (b) The person called 'Hannibal' at t2 is identical with the person called 'Hannibal' at t1; *and*
- (c) The animal called 'Hanimal' at t1 is identical with the person called 'Hannibal' at t1.

However from those three premises we seem to get the conclusions that:

- (d) The animal called 'Hanimal' at t2 is identical with the person called 'Hannibal' at t1 (from a, c, and transitivity).
- (e) The animal called 'Hanimal' at t2 is identical with the person called 'Hannibal' at t2 (from b, d, and transitivity).
- (f) Hanimal is identical with Hannibal at t2.

But, again, the time of the identities is not made explicit and doing so undercuts the *reductio*. Here's how it works with the appropriate temporal indexing:

- (A) At t1: the animal called 'Hanimal' at t2 is identical with the animal called 'Hanimal' at t1; *and*
- (B) At t1: the person called 'Hannibal' at t2 is identical with the person called 'Hannibal' at t1; *and*
- (C) At t1: the animal called 'Hanimal' at t1 is identical with the person called 'Hannibal' at t1.

However from those three premises we get the conclusions that:

- (D) At t1: the animal called 'Hanimal' at t2 is identical with the person called 'Hannibal' at t1 (from A, C, and transitivity).
- (E) At t1: the animal called 'Hanimal' at t2 is identical with the person called 'Hannibal' at t2 (from B, D, and transitivity).
- (F) At t1: Hanimal is identical with Hannibal at t2.

But surely (F) isn't the claim that Hanimal is identical with Hannibal at t2. So we don't get the inconsistency if we make explicit the times of the identity statements.

When we adopt the temporary identity thesis we also accept a temporal version of Leibniz's Law (cf. chapter 6, sect. 6.2):

T-Leibniz's Law (x)(y)(F)(t)[t: x = y → (t: Fx → t: Fy)].

This brings me to what I think is an important consequence of accepting the temporary identity thesis. It is clear that we should accept that:

- (1) t1: Hanimal = Hannibal.
- (2) Not (t2: Hanimal = Hannibal).

Now let's say that Hanimal is in the hospital and Hannibal is in the cinema at t2. Then we will have to accept that:

- (3) t2: Hanimal is in the hospital.
- (4) Not (t2: Hannibal is in the hospital).

From (3) we get:

(5) t1: Hanimal is in the hospital at t2.

But together with T-Leibniz's Law (5) entails that:

(6) t1: Hannibal is in the hospital at t2.

But it seems natural to conclude from (6) that :

(7) t2: Hannibal is in the hospital.

Now (7) contradicts (4) which states that: Not (t2: Hannibal is in the hospital). So the acceptance of the T-Leibniz's Law seems to lead to a *reductio*, since we'd then have to accept (1) at t1: Hanimal is identical with Hannibal; (3) Hanimal is in the hospital at t2; but (4) states that Hannibal is not in the hospital at t2.

When I discussed how to respond to the T-Leibniz Law Argument (chapter 6, sect. 6.3) I suggested that we accept a principle about the necessary and sufficient conditions for something to have a property of being 'F at t*' at some other time t, which grounds the moves from (3) to (5) and from (5) to (6):

E-Principle $(x)(t)(t^*)(F)[\text{at } t: Fxt^* \leftrightarrow (\exists y)(\text{at } t: x=y \ \& \ Fyt^*)]$.

According to this principle at time t x has the property of being F at t^* if and only if there is something which is identical to x at t and which is F at t^* . At t1 Hannibal is identical with Hanimal and Hanimal is in the hospital at t2, hence at t1 Hannibal has the property of being in the hospital at t2, which is what (6) states. However, if we accept the E-Principle (7) doesn't follow from (6). We can't conclude from the fact that at time t1 Hannibal is identical to something which is in the hospital at another time t2 and thus given T-

Leibniz's Law at t1 has the property of being in the hospital at t2, that Hannibal is in the hospital at t2, since on the E-Principle it is sufficient for Hannibal to have that property that he is identical something at t1 which is in the hospital at t2. But that something need not be himself, as indeed it isn't on the temporary identity description of the case.

If one wants to get (7) to follow from (6), then one would have to accept an alternative and stronger principle than the E-Principle which might be thought to be correct:

A-Principle $(x)(t)(t^*)(F)[\text{at } t: t^*:Fx \leftrightarrow (y)(t: x=y \rightarrow t^*: Fy)]$

According to the *A-Principle*, Hannibal can have the property of being in the hospital at t2 at t1 if and only if *everything* identical with Hannibal at t1 is in the hospital at t2. This is of course incompatible with the temporary identity description of multiple-kind fission and would thus beg the question against that view.

On the other hand, it doesn't work in support of the T-Leibniz's Law argument either, since it would rule out the move from (3) to (5); (3) says that at t2: Hanimal is in the hospital and from this we are supposed to get (5) which says that at t1: Hanimal is in the hospital at t2. But assuming for the sake of *reductio* the temporary identity description of the case, we don't get (5) from (3) on the A-Principle. In order for (5) to follow from (3) on the A-Principle, *everything* identical with Hanimal at t1 must be in the hospital at t2. But on the temporary identity description of the case, Hannibal isn't, and thus the result of accepting the A-Principle, without begging the question against the temporary identity description, is that it will be incorrect to claim that at t1 *Hanimal* has the property of being in the hospital at t2, since not everything identical with Hanimal at t1 is in the hospital at t2. However, at t2 everything identical with Hanimal is in the hospital at t2, so Hanimal will

change its time-indexed property of being in the hospital at t_2 during its existence.

Thus someone accepting the A-Principle will not be in a position to claim superiority by pointing out that on the E-Principle a thing can change its time-indexed properties. Assuming the temporary identity description of the multiple-kind fission case the A-Principle will also have this implication. I don't take the A-Principle to be plausible, but find the E-Principle to be correct and thus I think we should block the *reductio's* move from (6) to (7).

Denying the move from (6) to (7) entails that a thing can change its time-indexed properties from one time to another (cf. chapter 6, sect. 6.4). At t_1 Hannibal has the time-indexed property of being in the hospital at t_2 , and at t_2 he doesn't. Thus the temporary identity thesis involves rejecting the following principle:

Transmission $(x)(t)(t^*) \{ \text{at } t: \text{Ex} \ \& \ \text{at } t^*: \text{Ex} \rightarrow [\text{at } t^*: \text{Fxt}^* \leftrightarrow (\text{at } t: \text{Fxt}^*)] \}$ ¹⁴⁰

This principle states that if some thing exists at different times t and t^* , then it has a property F -at- t at t if and only if it has the property F -at- t at the other time t^* . *Transmission* expresses the general metaphysical thesis that things don't change their time-indexed properties.

In my discussion of standard fission I pointed out that the temporary identity thesis rejects *Transmission* when acknowledging the truth of

(6) at t_1 : Hannibal is in the hospital at t_2 ,

and the falsity of

¹⁴⁰ Cf. Gallois 1998, p. 96. 'Ex' should be taken to mean 'x exists.'

(7) at t2: Hannibal is in the hospital (at t2).

Whether we accept the E-Principle or the A-Principle as stating the truth-conditions for ascriptions of time-indexed properties, we must accept that things can change their time-indexed properties over time if we accept the temporary identity thesis. On the E-Principle it is enough for Hannibal to have the property of being in the hospital at t2 at t1 that there is something with which he is identical at t1, which is in the hospital at t2. Since there's nothing identical with Hannibal at t2 which has the property of being in the hospital at t2, Hannibal change his time-indexed properties.

But this is also the case if we accept the A-Principle instead of the E-Principle. On the A-Principle Hannibal won't have the property of being in the cinema at t2 at t1, since not everything identical with Hannibal at t1 is in the cinema at t2. But at t2 everything identical with Hannibal does have the property of being in the cinema at t2, and thus he has changed his time-indexed properties.

Now an important question with respect to the temporary identity thesis's description of the multiple-kind fission case is *why* Holmes isn't identical to either of his offshoots at t2 (cf. chapter 6, sect. 6.5). After all, a guiding concern for someone who, like me, is attracted to the temporary identity thesis is that it doesn't entail that Hannibal and Hanimal come into existence when Holmes's fission occurs, but that it is a case in which they are identical to Holmes before Holmes's fission. But the following question then arises: Why isn't Holmes identical to, say, Hannibal at t2, when he is identical to Hannibal at t1 and Hannibal exists at t2?

The answer is the same as it is with respect to the other properties Hannibal doesn't have at t2 despite having them at t2 at t1. At t1 Holmes has the

property of being identical with Hannibal at t2, since Holmes is identical to something at t1 which is identical to Hannibal at t2. This is what the E-Principle says. But in order for Holmes to be identical with Hannibal at t2, *everything* identical with Holmes at t1 must be identical with Hannibal at t2 (as required by the A*-Principle¹⁴¹), and this is not so, since Hanimal is not identical with Hannibal at t2.

Now one might be tempted to reason that from:

- The thing which is Hannibal at t2 = the thing which is Hanimal at t1,
and
- The thing which is Hanimal at t1 = the thing which is Hanimal at t2,

we get that:

- The thing which is Hannibal at t2 = the thing which is Hanimal at t2.

But as I noted with respect to a similar line of reasoning (chapter 6, sect. 6.6), identity is not a relation which holds trans-temporally on the temporary identity view. What trans-temporal identity comes to given the temporary identity thesis is the following *non-symmetrical* relation:

Trans-Temporal Identity $(x)(y)[x T= y \leftrightarrow (\exists t)(\exists t^*)(z)(\text{at } t: z = x \rightarrow \text{at } t^*: z = y)]$.

It is clear that in multiple-kind fission it is true that everything identical with Hanimal at t2 is identical with Holmes at t1, but it is not the case that everything identical with Holmes at t1 is identical with Hanimal at t2.

¹⁴¹ A*-Principle $(x)(t)(t^*)(F)[(y)(\text{at } t: x=y \rightarrow \text{at } t^*: Fy) \rightarrow \text{at } t^*: Fx]$

Let me finally note that in the multiple-kind fission case, we don't face the sort of problems discussed in chapter 6, sect. 6.7. concerning the fixing of the reference of 'Brown' as 'the person in room 100 before fission.' The problem with respect to 'Brown' was that after fission there are to be two distinct individuals, Lefty and Righty, which can both be said to satisfy the description 'the person in room 100 before fission' fixing the reference of 'Brown':

(D1) After fission: the person in 100 before fission = Lefty.

(D2) After fission: the person in 100 before fission = Righty.

However, I invoked the following principle stating the truth conditions for statements like (D1) and (D2):

(D) (t)[At t: the Fx is G if and only if something is F, everything which is F is identical with x at t, and x is G at t].¹⁴²

According to (D), (D1) is true if and only if something satisfies 'the person named Brown in room 100 before the fission' and everything which satisfies the definite description after fission is identical with Lefty after fission. But in standard fission this is not the case, since after fission Righty also satisfies the description 'the person named Brown in room 100 before the fission.' So neither (D1) nor (D2) are true.

Let's say that we use 'the person in the bar at t1' to fix the reference of 'Holmes' at t1. Then:

(D3) At t2: the person I named 'Holmes' in the bar at t1 = Hannibal,

¹⁴² This is my rendering of Gallois's (D') 1998, p. 106. My discussion follows his on p. 105 ff.

is true if and only if Hannibal satisfies the definite description 'the person I named 'Holmes' in the bar at t1' and everything which satisfies the description at t2 is identical with Hannibal at t2.

Now this condition *is* satisfied in the multiple-kind fission case. So there is no problem in saying at t2 that Hannibal is the person we named 'Holmes' before fission. Had we picked out the referent of 'Holmes' at t1 by using a description which contained 'the human organism ...', it would not be true to take the referent of 'Holmes' at t1 to be identical with Hannibal at t2, just as at t2 Hanimal isn't the referent of 'Holmes' at t1.

This concludes my presentation of the central elements in the temporary identity view's way of describing the multiple-kind fission case. I end by pointing out how the temporary identity thesis in a special way allows us to say what personalists, animalists and constitutionalists want to say about the situation before Holmes's multiple-kind fission. And how it also provides some interesting and in my view quite attractive ways of dealing with the circularity problem and the claim that identity is what matters in survival.

10.3 TEMPORARY IDENTITY, ANIMALISM, PERSONALISM, AND CONSTITUTIONALISM

I think it is worth pointing out how someone who adopts the temporary identity view is able to accommodate both animalist and personalist key claims while avoiding the problems I think face constitutionalism (i.e. Baker's version of coincidentalism).

Let's say that I am a person, Hannibal, and that at t1 I am identical to a human organism called 'Hanimal'. This is what both personalist and animalist agree on. As it happens, Hanimal is a human vegetable in a hospital at t2 and Hannibal is in the cinema at t2. Now on the temporary identity view we can maintain the central personalist claim that I (Hannibal)

am a person at all times at which I exists. But it can also be admitted that there are times at which Hanimal exists without being a person, even if it is identical with me at t_1 . So the central personalist claims that I am a person at all times at which I exist, that I am currently identical to a human animal, and that I exist at some other time if and only if I am psychologically continuous with someone at that time, all come out true on the temporary identity view. What doesn't come out true is that my existence at other times entails standing in a relation of diachronic identity with someone existing at those times, since identity is a temporally indexed relation which holds between individuals at a time, not over time.

On the temporary identity view I can also be an animalist in the sense that I can maintain three central claims of animalism, namely that paradigm persons are identical to particular human organisms, that a human organism is a human organism at all times at which it exists, and that for a human organism to exist at other times it must be biologically continuous with a human organism existing at those times. On the temporary identity view all three claims can be maintained. What can't be maintained is that a human organism to exist at other times must be identical over time to an organism existing at those times. Again, this is because identity is not a diachronic relation on the temporary identity view.

These ways of accommodating personalist and animalist key claims may seem to entail a problematic consequence. It appears that temporary identity entails the possibility that there are times at which I am identical to something which is essentially (or eternally) a human organism with biological persistence conditions, and that there are other times at which I am not identical to something which is essentially a human organism with biological persistence conditions. But isn't it implausible to claim that I am essentially of some kind K with persistence conditions C at one time and not

at another? Can essential properties and persistence conditions really be temporary?

Here's how I think a temporary identity theorist should state the necessary and sufficient conditions for x 's having a property F essentially *at a time*:

Temporary Essence Principle $(x)(t)(F)[t: \text{Es.Fx} \leftrightarrow (\exists y)(t: x=y) \ \& \ (t^*) (\exists y t^* \rightarrow \text{Fyt}^*)]$

Thus we can say that: at t_1 Hannibal is a human animal essentially, if and only if Hannibal is identical with some y and at all times at which y exists y is a human animal.¹⁴³

At t_1 Hannibal is identical with Hanimal and at all times at which Hanimal exists it is a human animal. So at t_1 Hannibal has the property of being a human animal essentially. At t_2 things are different. There isn't anything identical to Hannibal at t_2 which is a human animal at t_2 , so at t_2 Hannibal doesn't have the property of being a human animal essentially.

Now let's say that at t_3 , Hannibal is (again) identical with Hanimal (assume the BST has been reversed. If Hannibal is identical to Hanimal at t_3 , then Hannibal is a human animal at t_3 , and Hanimal is a human animal at all times at which it exists. Hence at t_3 , Hannibal is essentially a human animal.

The Temporary Essence Principle thus gives an account of the truth conditions for ascribing an essential property to an object at a time which accommodates that essential properties can be had temporarily.

¹⁴³ I should like to note that my treatment of essential properties is tentative and that this is an agenda that I hope to come back to in the future.

There is an alternative way in which one might suggest that the temporary identity view should account for having an essential property temporarily. One might suggest that in order for x to have property F essentially at time t , *everything* identical with x must be F at all times at which it exists, we get the following principle:

Alternative Temporary Essence Principle $(x)(t)(F)[t: \text{Es.Fx} \leftrightarrow (y)(t: x=y \ \& \ (t^*)(\exists y t^* \rightarrow \text{Fy} t^*)]$

On the Alternative Principle *everything* identical with Hanimal at t_1 must be a human animal at all times at which it exists in order for Hanimal to have the property of being a human animal essentially at t_1 . Now Hannibal doesn't have that property at all times at which he exists, and thus there is something identical to Hanimal at t_1 which isn't a human animal at all times at which it exists (Hannibal), so Hanimal doesn't have the property of being a human animal essentially at t_1 . Nor can Hannibal be said to be a person essentially at t_1 due to there being times at which Hanimal exist and isn't a person.

Now I find the Temporary Essence Principle to be a more plausible principle than the Alternative Temporary Essence Principle. On the Temporary Essence Principle I can maintain that Hannibal is essentially a person at t_1 and that Hanimal is essentially a human animal at t_1 , and that is what I want to maintain given that in the multiple-kind fission case, there are no times at which I can say that Hannibal exists and is not identical to a person and that Hanimal exist and is not identical to a human animal. It is unsatisfactory to claim, as is entailed by the Alternative Temporary Essence Principle, that despite being a human animal at all times at which it exists, Hanimal is not essentially a human animal at t_1 due to being temporarily identical to something which is not always person. And similarly it is unsatisfactory to say that Hannibal is not essentially a person at t_1 despite being a person at all times at which he exists.

I think it is interesting to note that if we explicitly adopt a position according to which only the past and present exists, then at t1 Hannibal will essentially be a human animal even on the Alternative Principle, but the result is that come t3, neither Hannibal nor Hanimal is essentially a human animal. Let's say the following statements are true:

At t0: Hanimal exists & Hanimal = Hannibal.

At t0: Not Hannibal exists.

At t1: Hanimal = Hannibal.

At t2: Hanimal = Hanimal.

At t2: Hannibal = Bion & not (Bion = Hanimal).

At t3: Hannibal = Hanimal.

If we take 't' to pick out the present and take 't*' only to range over past and present it follows from the identity statements that:

- At t0: Hanimal is essentially an animal since everything identical to Hanimal at t0 is a human animal at every time it exists.
- At t1: Hannibal is essentially an animal since everything identical to Hannibal at t1 is a human animal at all times at which it exists.
- At t1: Hanimal is essentially an animal since everything identical to Hanimal at t1 is a human animal at all times at which it exists
- At t2: Hannibal is not essentially an animal since something identical to Hannibal at t2, namely Hannibal, isn't a human animal at all times at which it exists.
- At t2: Hanimal is essentially an animal since everything identical to Hanimal at t2 is a human animal at all times at which it exists.
- At t3: Hannibal is not essentially a human animal since something identical to Hannibal at t3, namely Hannibal, isn't a human animal at all times at which it exists.

- At t_3 : Hanimal isn't essentially a human animal since something identical to Hanimal at t_3 , namely Hannibal, isn't a human animal at all times at which it exists.

At a given time Hannibal is essentially a human organism if and only if he is identical to a human organism at all times past and present at which he exists. Now at t_1 it is true that Hannibal has been identical to a human organism at all times at which he exists. So he is essentially an organism at t_1 . But at t_3 , when we take Hannibal to be identical to Hanimal again, it is not the case that Hannibal is essentially a human organism, since there are past times at which he exists and isn't identical to a human organism. But now we also get the result, on the Alternative Principle, that Hanimal despite being identical to a human organism at all times at which it exists, isn't essentially an organism at t_3 , since it is identical with something which isn't always identical with a human organism when it exists.

As I have already pointed out, I think that the Temporary Essence Principle is more plausible than the Alternative Principle, but on either principle we allow for something to have a property essentially at one time at which it exists and not have it essentially at another time at which it exists.¹⁴⁴

Finally, unlike the constitution view, the temporary identity view doesn't involve counting Hannibal and Hanimal as one without them being numerically one. There is no need for a notion of constitution which, even if it has its virtues such as providing a basis for avoiding 'too many'-style objections, entails the unhappy consequence that what *counts* as one at some time t isn't *numerically* one at t , but perhaps constitutionally one at t .

¹⁴⁴ How the temporary identity thesis connects with competing positions in the philosophy of time is in my view interesting and I hope to consider it in more detail elsewhere.

Now where Baker's notion of constitution is a notion of a unifying synchronic relation which allows for two individuals x and y to exist in the same place at the same time without being numerically identical, the temporary identity view might be seen as a view which takes there to be a synchronic unifying relationship between x existing at t and y existing at another time t^* , and while this constitutes the existence of an individual continuous between those two times, there is not a diachronic relation of identity which holds between x at t and y at t^* . What this means is that persistence conditions are not conditions for the *identity* of some kind of thing, e.g. for persons, since identity is not a diachronic relation on the temporary identity view. When we say, as I have frequently done in this thesis, that for x to exist at some other time is for there to exist a y at that other time which is related to x by some continuity relation which purportedly constitutes the numerical identity of x with y , this way of putting the special question about persistence is misleading. If we accept temporary identity, since it assumes that things picked out at different times can stand in the relation of numerical identity and that genuinely to persist is for one to stand in such a relation to things at other times.

On the temporary identity view existing at different times is not a question about identity over time, but a question about who one is identical to at a time. So the claim is not that persistence doesn't involve identity, it is just that whether I persist depends not on whom I am identical to tomorrow, but on whom I am identical to today. And whether tomorrow I have a past depends not on whom I am identical to today, but on whom I am identical to tomorrow. I think this is a perfectly coherent and as a matter of fact quite attractive view to take on persistence and identity.

My basis for finding the temporary identity thesis attractive is based on the way it allows a supporter of the psychological criterion to deal with fission scenarios. The relation of psychological continuity constitutes identity, but

identity is a synchronic relation. Still the diachronic relation of psychological continuity can nevertheless be maintained to be a condition for the existence of an individual enduring person at different times which justifies the things that fission offshoots would naturally want to say about the prefission situation. I'll end by presenting how I think the temporary identity view provides attractive ways of dealing with the circularity problem and the question about whether identity matters in survival.

10.4 TEMPORARY IDENTITY AND THE CIRCULARITY PROBLEM

In chapter 4 I discussed the objection that the psychological criterion is viciously circular in that it analyses personal identity in terms of relations such as memory relations, which themselves can only be known to obtain if one is in a position to determine questions about the identity of the person apparently remembering and the person of the experience apparently remembered. Since my apparent memory of some past experience is only veridical if I am identical with the person whose experience I apparently remember, we simply can't determine whether my apparent memories are veridical independently of applying a criterion of personal identity, and thus it is circular to state such a criterion in terms involving reference to memory relations.

In chapter 4 I discussed a rather popular approach to the problem of circularity suggested e.g. by Shoemaker (1970, 1984). The popular way to avoid the objection is to analyse personal identity in terms of a relation, quasi-memory, which doesn't presuppose the identity of quasi-rememberer and the original subject of the quasi-remembered experience. Ordinary memories are the subset of quasi-memories which are quasi-memories of my *own* past, i.e. of experiences had by someone identical to me. In cases of person-fission such as the case discussed in chapters 5 and 6 where Brown fissions into Lefty and Righty, the fission offshoots both seem to stand in a relation to Brown which is *just like* memory, except that on some accounts

they aren't identical to Brown. We can maintain that their apparent memory state is veridical and not an illusion, if we take it to be a quasi-memory state, since such a state is veridical despite the fact that the quasi-rememberer isn't identical to the original subject. Similar considerations apply to e.g. Lefty and Righty's acting on an intention formed by Brown. In that case they act on a quasi-intention.

If we accept the temporary identity thesis there seems to be two ways to go: Either there aren't any apparent memories which are genuine and veridical, since there is no such thing as being identical over time. Or we can say that the importance of being identical when it comes to having a genuine memory is, that in order for me to genuinely remember some past experience, I must presently be identical to someone who *was* identical to the original subject of the remembered experience at the time of its occurrence. Here I will show how we can maintain that identity is a precondition for genuine memory when taken to be temporary identity.

Let's say that at t_3 we have a person named 'Green' and at t_3 Green is identical to Lefty and also identical to Righty, since at t_3 everything psychologically continuous with Lefty is identical to Green and everything psychologically continuous with Righty is identical to Green. Thus Green may be thought of as a fusion of Lefty and Righty. Alternatively one may say that at t_3 everything which has Lefty as a psychological predecessor and everything which has Righty as a psychological predecessor is identical to Green.

Now, according to my suggestion, for Green to truly remember what Lefty experienced at t_2 , he must be identical with a person who was identical with Lefty at t_2 . And for Green to remember what Righty experienced at t_2 , he must be identical with a person who was identical with Righty at t_2 . At t_3 Green is identical with Lefty and Green is identical with Righty, so at t_3 he

satisfies the conditions for having a veridical memory of what Lefty experienced at t_2 and of what Righty experienced at t_2 .

Adopting the temporary identity thesis thus allows us to point out that while psychological continuity is central to recalling the past from the inside in a genuine non-fictional, a fission offshoot like Lefty can truly say that 'I was identical with someone who experienced such and such.' On Shoemaker's view, a *mere* quasi-rememberer like Lefty is not in a position correctly to say so, since he did not exist when the experience quasi-remembered was had.

On the temporary identity thesis there is no identity over time to be presupposed by genuine memory which is a transtemporal relation. I might now stand in a genuine memory relation to some past person with whom I was not identical at the time the past person had the experience, and I may do so in virtue of being presently identical with someone who did exist at the time of the occurrence of the experience and was identical with its subject at that time. So if we adopt the notion of temporary identity, we need not appeal to the notion of quasi-memory in order to reply to the circularity objection. Ordinary memory presupposes identity, but identity is temporary and thus there is need for the notion of quasi-memory.

It is important to note that none of this entails that Green was in two places at t_2 having two distinct experiences (e.g. of waking up in room 102 and of waking up in room 104). Just as there won't be a time when Brown *will* be in both in room 102 and in room 104, there never *was* a time at which Green was in both room 102 and 104. At t_2 there is no individual who is identical to Green, since there is no individual at t_2 with which every psychological predecessor of Green is identical at that time.

What would a quasi-memory be on the temporary identity view? Well, it would be what I have, when at some time t I am identical to *someone* who *was*

identical to the original person who had the experience apparently remembered, while not being identical to *everyone* who was identical to the original person. This is the case at t_2 with respect to Lefty and Righty's apparent memories of Brown's life. And clearly, when formulated in this way on the basis of temporary identity, it would be wholly uncalled for to doubt the veridicality of quasi-memories such as those had by Lefty and Righty at t_2 . It would make the nature and epistemic value of a mental state which is apparently a memory wholly dependent on relational facts. The requirement would be that in order for an apparent memory of mine to be genuine, there must not exist *someone else* at the time in question who is also identical to someone who *was* identical to the person with the apparently remembered experience. This is not an attractive view to take on what is required for an apparent memory to be veridical. Hence it seems reasonable to think that quasi-memories under the temporary identity characterisation are veridical.

10.5 TEMPORARY IDENTITY AND WHAT MATTERS IN SURVIVAL

I'll end by pointing out how the temporary identity thesis allows us to maintain that identity *does* play a significant role in prudential concern i.e. that it 'matters in survival.' In chapter 3 Parfit's argument that identity isn't what matters in survival was discussed. According to Parfit, Brown's fission should be described as a case in which a person ceases to exist and is replaced by his fission offshoots; Brown isn't identical to either of his offshoots, but stands in the relation that matters in survival to them both.

I think that adopting the temporary identity thesis gives us an interesting way of acknowledging that identity does matter in survival. We agree that the relation between Brown and his two fission offshoots matters. Parfit argues that since there isn't identity, what matters is not the identity relation. However, if identity is a temporary relation, then the question about whether identity matters in survival should be reformulated to take this into account.

So the claim that identity matters in survival should be put as a question about whether or not the future of those who are identical to me *now* matters, not about whether those identical to me in the future matter to me now. It might be that I am now identical to someone with whom I am not identical next week, but given that I am identical with him now, my concern for my own future is *ipso facto* a concern for his future as well.

But now we can explain how identity does matter with respect to Brown's concern for Lefty and Righty. Brown is identical with Lefty and Righty before the fission, and hence their future matter to him before his fission. And after the fission, Brown's past life will matter to both Lefty and Righty the way our own past matters to us, because after the fission they both *were* identical to Brown even if they aren't identical to him anymore. So while it is true that identity over time doesn't matter in survival, the truth of this claim is due to the fact that there is no such relation as diachronic identity for persons picked out at different times to stand in. It is however mistaken to conclude that *identity* doesn't matter in survival at all. What matters in survival is *temporary* identity.

Thus Parfit's claim that what I am concerned about when ordinarily concerned about my own survival isn't personal identity, but a relation 'just like it' except that it needn't entail personal identity (what one might call 'quasi-prudential concern') is in a sense correct, just as the view that we can have quasi-memories which are 'just like memories' except that they don't entail identity, is in a sense correct. They are correct in the sense that the identity relation supposed to be entailed by the ordinary notions of prudential concern and memory isn't to be found. But they are mistaken in the claim that identity doesn't play a central role to the veridicality of memory and prudential concern. It is just that it is temporary identity which grounds claims to genuinely remember some past person's experiences and to be genuinely prudentially concerned about some future person's welfare.

To remember someone's past and be prudentially concerned about someone's future he needn't be identical to me in the past or in the future, but he must be identical to me *now*.

10.6 CONCLUSION

In this chapter I have presented how a temporary identity theorist may describe the multiple-kind fission case allowing for paradigm persons to be identical to human organisms even if human organisms and persons have different persistence conditions.

I began the last three chapters by discussing whether personalism or animalism can provide a case for taking either biological continuity or psychological continuity to be metaphysically dominant with respect to the persistence of paradigm persons like us who are plausibly taken to be identical to human organisms. I found that we have equally good reason to think that both candidate relations are constitutive of the persistence of individuals such as Hanimal and Hannibal respectively. This led me to consider another way of dealing with multiple-kind fission, coincidentalism, which suggests that there are really two numerically non-identical individuals of different kinds with different properties occupying the location I occupy now. I presented coincidentalism with some hard questions and then canvassed Baker's idea of 'constitution' as a relation which holds between individuals which aren't identical nor of the same kind, but nevertheless occupy the same location at a given time. I criticised the notion of constitution for selling out on the eternity of numerical non-identity and the idea that if something *counts* as one at a time, then it is *numerically* one at that time.

Finally I presented how the temporary identity view provides way of dealing with multiple-kind fission cases in a way parallel to the standard fission case and how it can handle having essential properties temporarily. I ended by

pointing out how it also allows for ways of dealing with the circularity objection and the claim that identity isn't what matters in survival discussed in chapters 3 and 4.

The claim that identity is temporary is a controversial view. I began this thesis not even taking it seriously as a way of dealing with the problems facing the psychological criterion. While I don't pretend to have given a complete defence of the temporary identity view, I submit that accepting temporary identity provides the most satisfactory and unified way for a proponent of the psychological approach to deal with the objections discussed in this thesis.

Bibliography

- Ayers, Michael. 1974: 'Individuals without Sortals', *Canadian Journal of Philosophy*, iv: 113-148.
- 1991: *Locke – Epistemology and Ontology*, London: Routledge.
- Baker, Lynne Rudder. 2000: *Persons and Bodies*, Cambridge: Cambridge University Press
- 2002: 'On Making Things Up: Constitution and its Critics', *Philosophical Topics* 30: 31-51.
- Burke, Michael. 1992: 'Copper Statues and Pieces of Copper: A Challenge to the Standard Account', *Analysis* 52: 12-17.
- 1994a: 'Dion and Theon: An Essentialist Solution to an Ancient Puzzle', *The Journal of Philosophy*, 91: 129-39.
- 1994b: '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*, 54: 591-624.
- 1997a: 'Coinciding Objects: Reply to Lowe and Denkel', *Analysis* 57: 11-18.
- 1997b: 'Persons and Bodies: How to Avoid the New Dualism', *American Philosophical Quarterly* 34: 457-67.
- 2004: 'Dion, Theon, and the Many-Thinkers Problem', *Analysis*, 64: 242-50.
- Butler, Joseph. 1975: 'Of Personal Identity', in Perry (ed.) 1975, pp. 99-105.
- Campbell, John. 1994: *Past, Space and Self*, Cambridge MA: MIT Press
- Carter, W. R. 1982: 'Do Zygotes become People?', *Mind*, 91: 77-95.
- 1990: 'Why Personal Identity is Animal Identity', *LOGOS* 11: 71-81.
- 1997: 'Dion's Left Foot (and the Price of Burkean Economy)', *Philosophy and Phenomenological Research*, 57: 371-379.
- 1999: 'Will I be a Dead Person?', *Philosophy and Phenomenological Research*, 54: 161-171.

- Cassam, Quassim. 1992: 'Reductionism and First-Person Thinking', in Charles and Lennon eds. 1992.
- Charles, D. and Lennon, K. (eds.) 1992: *Reduction, Explanation, and Realism*, Oxford: Oxford University Press.
- Chisholm, Roderick. 1976: *Person and Object*, London: Allen & Unwin.
- Dancy, Jonathan. (ed.) 1997: *Reading Parfit*, Oxford: Blackwells.
- Dickens, Charles. 1859: *A Tale of Two Cities*, London: Penguin Popular Classics.
- Evans, Gareth. 1982: *The Varieties of Reference*, Oxford: Oxford University Press.
- Fine, Kit. 2003: 'The Non-Identity of a Material Thing and its Matter', *Mind*, 112: 195-234.
- Garrett, Brian. 1998: *Personal Identity and Self-Consciousness*, London: Routledge.
- Gallois, Andre. 1998: *Occasions of Identity*, Oxford: Oxford University Press.
- Gibbard, Alan. 1975: 'Contingent Identity', in Rea (ed.) 1997, pp. 93-125.
- Gill, C. (ed). 1990: *The Person and the Human Mind*, Oxford: Oxford University Press.
- Grice, H. P. 1941: 'Personal Identity', in Perry (ed.) 1975, pp. 73-95.
- Hawley, K. 2001: *How Things Persist*, Oxford: Oxford University Press.
- Hershenov, David. 2005: 'Do Dead Bodies Pose a Problem for Biological Approaches to Personal Identity?', *Mind*, 114: 31-59.
- Hirsch, Eli. 1982: *The Concept of Identity*, Oxford: Oxford University Press.
- Hudson, Hud. 2001: *A Materialist Metaphysics of the Human Person*, Ithaca: Cornell University Press.
- Hume, David. 1978: *A Treatise of Human Nature*, Oxford: Oxford University Press.
- Johnston, Mark. 1987: 'Human Beings', *Journal of Philosophy*, 84: 59-83.
- 1989: 'Fission and the Facts', *Philosophical Perspectives*, 3: 369-397.
- 1997: 'Human Concerns without Superlative Selves', in Dancy (ed.) 1997, pp. 149-179.

- Laycock, Henry. 1972: 'Some Questions about Ontology', *The Philosophical Review* 81: 3-42.
- Locke, Don. 1971: *Memory*, London: Macmillan.
- Locke, John. 1996: *An Essay Concerning Human Understanding*, London: Everyman.
- Lewis, David. 1971: 'Counterparts of Persons and their Bodies', *Journal of Philosophy*, 68: 203-211.
- 1976: 'Survival and Identity', in Rorty (ed.) 1976 pp. 17-40.
- 1986: *On the Plurality of Worlds*, Oxford: Blackwell.
- Lowe 1989a: *Kinds of Being*, Oxford: Blackwell.
- 1989b: 'What is a Criterion of Identity', *The Philosophical Quarterly*, 39: 1-21.
- 1995a: 'Coinciding Objects: In Defence of the 'Standard Account'', *Analysis*, 55: 171-178.
- 1995b: *Locke on Human Understanding*, London: Routledge.
- 2002: *A Survey of Metaphysics*, Oxford: Oxford University Press.
- McIntyre J. L. 1978: 'The Role of Temporal Adverbs in Statements about Persons', *Nous*, 12: 443-461.
- Mackie J. L. 1976: *Problems from Locke*, Oxford: Oxford University Press.
- Mackie, Penelope. 1994: 'Sortal Concepts and Essential Properties', *Philosophical Quarterly*, 44: 311-33.
- McDowell, John. 1997: 'Reductionism and the First Person', in Dancy (ed.) 1997, pp. 230-250.
- Martin, R. & Barresi, J. (eds.). 2003: *Personal Identity*, Oxford: Blackwell.
- Martin, C. B. & Deutscher, Max. 1966: 'Remembering', *The Philosophical Review* 25.
- Munitz, Milton K. (ed.) 1971: *Identity and Individuation*, New York: New York University Press.
- Nagel, Thomas. 1978: *The Possibility of Altruism*, Princeton NJ: Princeton University Press.
- 1986: *The View from Nowhere*, Oxford: Oxford University Press.
- Noonan 1985: 'The closest continuer theory of identity', *Inquiry*, 28: 195-229.

- 2003: *Personal Identity*, Routledge.
- Nozick, Robert. 1981: *Philosophical Explanations*, Cambridge MA: Harvard University Press.
- Olson 1997: *The Human Animal*, Oxford: Oxford University Press.
- 2001: 'Material Coincidence and the Indiscernibility Problem', *Philosophical Quarterly* 51: 337-55.
- 2002: 'What does Functionalism Tell Us about Personal Identity?', *Noûs* 36: 682-98.
- 2003: 'An Argument for Animalism', in Martin. R. and Barressi, J. (eds.), *Personal Identity*, pp. 318-34.
- 2004: 'Animalism and the Corpse Problem', *Australasian Journal of Philosophy* 82: 265-74.
- Parfit, Derek. 1971: 'Personal Identity', *Philosophical Review*, 80: 3-27.
- 1987: *Reasons and Persons*, Oxford: Oxford University Press.
- 1999: 'Experiences, Subjects and Conceptual Schemes', *Philosophical Topics*, 26.
- Perry, John. 1972: 'Can the Self Divide?', *Journal of Philosophy*, 69: 463-88, reprinted in Perry 2002a, pp. 34-63.
- (ed.) 1975: *Personal Identity*, Berkeley: California University Press.
- 1976: 'The Importance of Being Identical', in Rorty (ed.) 1976, pp. 67-90.
- 2002a: *Identity, Personal Identity and the Self*, Indianapolis Ind: Hackett.
- 2002b: 'The Two Faces of Identity', in Perry 2002a, pp. 64-83.
- Prior, Arthur. 1957: 'Opposite Number', *Review of Metaphysics*, 11: 196-201.
- Quine, W. V. 1950: 'Identity, Ostention, and Hypostasis', in Quine 1980, pp. 65-79.
- 1980: *From a Logical Point of View*, Cambridge MA: Harvard University Press.
- Quinton, A. 1962: 'The Soul', in Perry 1975, pp. 53-72.
- Rea, Michael. (ed.) 1997: *Material Constitution*, Lanham Md: Rowman & Littlefield.

- Robinson, Denis. 1985: 'Can Amoebae Divide without Multiplying?', *Australasian Journal of Philosophy*, 66: 299-243.
- Rorty, Amelie. (ed.) 1976: *The Identities of Persons*, Berkeley: California University Press.
- Shoemaker & Swinburne 1984: *Personal Identity*, Oxford: Basil Blackwell.
- Shoemaker, Sydney. 1959: 'Memory and Personal Identity', reprinted in Perry (ed.) 1975, pp. 119-134.
- 1963: *Self-Knowledge and Self-Identity*, Ithaca: Cornell.
- 1970: 'Persons and their Pasts', reprinted in Shoemaker 2003, pp. 19-48.
- 1971: 'Wiggins on Identity', in Munitz (ed.) 1971, pp. 103-117.
- 1997a: 'Parfit on Identity', in Dancy (ed.) 1997, pp. 135-148.
- 1997b: 'Self and Substance', *Philosophical Perspectives*, 11: 283-204.
- 1999a: 'Self, Body and Coincidence', in Shoemaker 2003, pp. 381-406.
- 1999b: 'Eric Olson, The Human Animal', *Nous*, 33: 496-504.
- 2003: *Identity, Cause and Mind*, Oxford: Oxford University Press.
- Sider, Ted. 2001: *Fourdimensionalism*, Oxford: Oxford University Press.
- Snowdon, Paul. 1990: 'Persons, animals, and ourselves', in Gill, C. 1990, pp. 83-107.
- 1991: 'Personal Identity and Brain Transplants', in *Human Beings*, ed. D. Cockburn (ed.), Cambridge: Cambridge University Press.
- 1995: 'Persons, Animals and Bodies', in *The Body and the Self*, J. L. Bermudez, A. Marcel, and N. Eilan (eds.), MIT Press.
- 1996: 'Persons and Personal Identity', in *Essays for David Wiggins: Identity, Truth and Value*, eds. S. Lovibond and S. G. Williams, Oxford: Basil Blackwell.
- Sorenson, Roy. 1992: *Thought Experiments*, Oxford: Oxford University Press.
- Swinburne 1984: 'The Dualist Theory', in Shoemaker & Swinburne 1984, pp. 1-66.
- Thomson, Judith Jarvis. 1997: 'People and their Bodies', in Dancy 1997, pp. 202-229.

- Unger, Peter. 1990: *Identity, Consciousness and Value*, Oxford: Oxford University Press.
- 2000: 'The Survival of the Sentient', *Philosophical Perspectives*, 14.
- Van Inwagen, Peter. 1980: 'Philosophers and the Words "Human Body"', in *Time and Cause*, ed. P. van Inwagen, Reidel.
- 1981: 'The Doctrine of Arbitrary Undetached Parts', in Rea (ed.) 1997, pp. 191-208.
- 1990: *Material Beings*, Ithaca: Cornell.
- 1997: 'Materialism and the Psychological-Continuity Account of Personal Identity', *Philosophical Perspectives*, 11.
- Wasserman, Ryan. 2002: 'The Standard Objection to the Standard Account', *Philosophical Studies*, 111: 197-216.
- Velleman, J. David. 1996: 'Self to Self', in *The Philosophical Review*, 105: 39-76
- Wiggins, David. 1967: *Identity and Spatio-Temporal Continuity*, Oxford: Blackwell.
- 1980: *Sameness and Substance*, Oxford: Blackwell.
- Williams, Bernard. 1956: 'Personal Identity and Individuation', in Williams 1973.
- 1966: 'Imagination and the Self', in Williams 1973.
- 1970a: 'Are Persons Bodies?', in Williams 1973.
- 1970b: 'The Self and the Future', in Williams 1973.
- 1973: *Problems of the Self*, Cambridge.
- Yablo, Stephen. 1987: 'Identity, Essence and Indiscernability', *Journal of Philosophy*, 84: 293-314.
- Zammiello, Fred. 2001: 'The Logic of Burke's Sortal Essentialism', *Pacific Philosophical Quarterly*, 82: 71-86.
- Zimmerman, Dean. 1995: 'Theories of Masses and Problems of Constitution', *The Philosophical Review*, 104: 53-110.
- 1997: 'Coincident Objects: Could a 'Stuff Ontology' help?', *Analysis*, 57: 19-27.