USE OF THESESES

This copy is supplied for purposes of private study and research only. Passages from the thesis may not be copied or closely paraphrased without the written consent of the author.
Surviving Matters:
Pluralism and the Self

Caroline Jane West
Philosophy Program
Research School of Social Sciences
The Australian National University

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

Declaration

This dissertation is my own work, except where otherwise acknowledged.

Caroline West
Acknowledgments

This thesis has benefited greatly from many helpful and stimulating discussions with friends and colleagues. Without them it would be much the less. I must thank Nick Agar, Daniel Nolan, Graham Oppy, John O'Leary Hawthorne, Philip Gerrans, Brian Garrett, Moira Gatens, Denis Robinson and Kevin Mulligan. I am especially indebted to my advisers, David Braddon-Mitchell and Philip Pettit, and to my supervisor, Frank Jackson, for innumerable and invaluable comments and criticisms, encouragement and enthusiasm; and to Michael Smith who, although in an unofficial capacity, provided all of the same. Thanks too, to my mother, Katharine, for her unflagging interest, faith and support.

It is hard to imagine a better or more stimulating research environment than I have been lucky enough to have in the Philosophy Program at the Research School of Social Sciences, The Australian National University. The place and the people have made my time as a PhD student a fascinating, rewarding and memorable one.
## Contents

Acknowledgments

Introduction 1

Chapter 1 Survival and Practice-Independence 7

1.1 A brief survey of the personal identity debate 7
1.2 Practice-independence versus practice-dependence 14
1.3 Response-dependence 22
1.4 Practice-dependence again 24
1.5 Locke: An unexpected ally 39

Chapter 2 Varieties of Practice-Independence 35

2.1 Old-Style Cartesianism: The one-phase *a priori* analysis 38
2.2 The role of the method of cases 40
   2.2.1 Option 1: Elucidating ordinary beliefs about personal identity 41
   2.2.2 Option 2: Practice-independence as rigidification 48
2.3 ‘Survival’ as a rigid designator 56
2.4 Survival as a natural kind 61
2.5 The causal/historical theory 67
2.6 Retaxonomizing the personal identity debate 70

Chapter 3 Practice-Independence and the Contingency of our Person-Directed Practices 72

3.1 Practice-independence and subject’s person-directed practices: 75
   Attempts to close the conceptual gulf
3.2 Two senses of ‘what matters’ in survival 76
3.3 Whether identity matters distinguished from practice-dependence versus practice-independence 84
3.4 Exclusive and inclusive justifiers 92

Chapter 4 Towards a Pluralist Account of Survival 102

4.1 The viking ship 102
4.2 The Teletransporters and the Somataphiles: 104
   A challenge for practice-independent theorists
4.3 A different presentation of the challenge? 118
4.4 The lesson of the Teletransporters and the Somataphiles 124

Chapter 5 A Pluralist Account of Survival 126

5.1 Survival as a first-order property 128
5.2 Survival as a second-order property 145
5.3 Survival: a first- or second-order property? 148

Chapter 6 Temporal-Phase Pluralism 155

6.1 The problem of refiguration 156
6.2 Are we Protean People? 161
6.3 Temporal-Phase Pluralism 170
6.4 Refiguration revisited 185
6.5 Organizing our person-directed concerns to survive 189
6.6 Pluralism distinguished from relativism 193

Chapter 7 Some Possible Objections 199

7.1 The Paradox of Analysis 199
7.2 Charting a course between Scylla and Charybdis 202
7.3 Do our person-directed practices presuppose survival? 205
7.4 Identity and personal identity 210

Chapter 8 A Constrained Pluralism? 215

8.1 The Contacti 219
8.2 The Contacti* 226
8.3 The Contacti** 230
8.4 Two possible lessons 236
8.5 Is the Cartesian view possible? 238

Chapter 9 Pluralism in the Actual World 243

Afterword 263

Bibliography 267
INTRODUCTION

The concept of personal identity or survival is intimately entwined with a whole host of our self- and other-directed beliefs, concerns, and practices. We naturally and habitually believe and judge certain people to be the same persons over time; identify and reidentify them; trust them; love them; punish them; are angry with them; care about them; attribute them certain ongoing rights and obligations; ascribe to ourselves and others actions, beliefs, intentions, sensations, thoughts, feelings, memories, perceptions and physical traits; expect to remember directly our own experiences; have a special and intense interest in what befalls ourselves; feel remorse, embarrassment, guilt, regret, pride, shame; and so on almost ad infinitum. Call all these many self- and other-directed beliefs, concerns, attitudes, expectations and practices, our person-directed practices. For persons, and their continuing existence over time, are the object of all these beliefs, concerns, attitudes, expectations and practices. So all these many person-directed practices are organized around our concept of personal identity. And it is this special and distinctive role which the concept of personal identity plays in our social network and in our own concerns that would seem to mark it off from concepts like squareness, for example, which do not play this special kind of role in subjects’ practices and concerns. As J.M. Shorter, for example, writes,

Why is it important to determine the identity of a given person correctly? From this point of view the importance of the concept is obviously immense. Our views about how we ought (morally speaking and otherwise) to behave towards other people depend very largely upon who we think those people are. How I think it proper to treat the person before me depends upon whether that person is (the same person as) my old friend, my employer, my wife, the homicidal maniac who escaped from detention last week, my father, my mother, my son, the man who did me a good turn yesterday, the Prime Minister, the man who robbed the bank six months ago, and so on. There is, therefore, a close connexion between the way I treat people and the sentences containing personal identity expressions which I believe to express true propositions. For example if I think that “Smith is not the person who embezzled the funds” expresses a true proposition, I shall not think that Smith should be punished for embezzling the funds, for something he did not do. As long as I retain this connexion between my attitudes and the sentences to which I assent, any decision of a verbal character that I make in a puzzle case will involve the taking up
of certain attitudes. If I decide to call A and B the same person I will not debar myself from blaming A for what B did. If I call them different people, then I shall be so debarred.\footnote{Shorter (1970-1), p.166.}

Given how richly and intimately our concept of personal identity is bound up with our person-directed practices it’s \textit{prima facie} somewhat surprising that nearly all accounts of personal identity to date, and certainly all those which seem to have gained widespread currency, have appeared to assume that the concept of personal identity is in exactly the same conceptual boat as primary quality concepts like squareness: a concept which is only contingently associated with all these manifold person-directed beliefs, attitudes, concerns, expectations and practices; a concept which is wholly constituted by some or other independently determined set of facts served up by nature (or God) alone.\footnote{Although rarely understood in this way, one of the most famous philosophers of personal identity, John Locke, is a notable exception here (as I will explain in Chapter 1). Two contemporary exceptions are Steven White (1989) and Mark Johnston (1989c) although their accounts differ from the one I shall favour.} They have been, what I will call, \textit{practicedependent} accounts of personal identity.

In this thesis I set out to argue for a fundamentally different kind of account of personal identity—a \textit{practice-dependent} account—according to which all these person-directed practices are not merely accidentally associated with the concept of personal identity, as practice-independent theorists have it, but are, in a crucial sense, constitutive of it. Survival, for any community, in any world, involves there holding among person-stages some relation around which the community in question organizes their person-directed practices and concerns.

This sort of view might initially strike some as unlikely. As Eli Hirsch writes,

\begin{quote}
The ordinary distinction between “me” and “not-me”, between that which does and that which does not lie within the boundaries of a single self, seems at least on first reflection completely inevitable. It is difficult even to understand the suggestion that this distinction might be arbitrary, or that it might legitimately be redrawn in some other way. Here, if anywhere, a “conventionalist” attitude is likely to strike us as intuitively incredible.\footnote{Hirsch (1982), p.286.}
\end{quote}
However intuitively incredible at first sight it might seem to view personal identity as (like the identity of corporations, nations or artefacts) a conventional matter, by the end of this thesis, I hope, it will seem incredible to think that we might ever have thought otherwise. Of course, there are various explanations for why we might have thought otherwise. Perhaps we thought that our survival conditions were determined by God, the all-powerful creator of all things, and so scarcely something which might be determined by the cares and concerns of mere mortals. Or perhaps we thought that nature played God’s role as independent, all-powerful arbiter of when we live and die. Or perhaps we thought that something as important as the difference between life and death couldn’t depend on something as apparently comparatively lightweight and dangerously contingent as the way in which a community happens to organize their person-directed practices. Or perhaps, as the vast majority of discussions about personal identity would suggest, we have assumed much, without really having paid much thought to the question of in virtue of what a relation earns its right to the name ‘personal identity’ at all.

For, to date, discussion of the question of personal identity—the question of what makes a person identified at one time the numerically same person as a person identified at another—has focussed almost exclusively on the question of what relation it is to which the term ‘personal identity’ necessarily refers. To this one question, philosophers have offered a bewildering array of competing answers. Some insist that the privileged relation to which ‘personal identity’ necessarily refers is sameness of soul; others that it is some other irreducible, further fact; others that it is continuity of a person’s brain and/or body; others that it is continuity of a person’s psychology; and yet others that it is some or other subtle mixture of the above. And so, notoriously, the personal identity debate has raged on with philosophers designing more and more sophisticated puzzle cases designed to support their preferred answer to the question of what relation it is to which ‘personal identity’ necessarily refers, and engaging in more and more elaborate and desperate attempts to explain away the appeal of other competing accounts.

Strangely enough in itself, but all the stranger in light of this substantial and ongoing disagreement about exactly what relation it is to
Introduction

which ‘personal identity’ necessarily refers, scarcely anyone has thought explicitly to raise and examine the higher-order question of in virtue of what any of these relations (physical continuity, psychological continuity, sameness of soul...) earn their right to the name ‘personal identity’ in the first place. Philosophers have debated long and hard about whether ‘personal identity’ refers to physical continuity, or whether it refers rather to psychological continuity, or whether it refers instead to sameness of soul, or to yet some other relation among person-stages. But rarely has there been even a whisper about in virtue of what physical continuity, or psychological continuity, or the soul, get to count as candidate relations for personal identity in the first place.

However, although philosophers have been strangely silent on this important higher-order conceptual question, lurking surreptitiously but influentially in the background to the heated limelight debate seems to be an implicit (and, I think, mistaken) assumption about the right answer to this higher-order question. Underlying and silently shaping this debate about what the referent of ‘personal identity’ is, is a certain view about the higher-order question of in virtue of what a relation earns its right to the name ‘personal identity’; a view which has been neither explicitly acknowledged nor defended. For, whatever exactly philosophers have insisted is the privileged relation to which ‘personal identity’ necessarily refers, the implicit common assumption has been that, for any community, in any world, if there is survival in that world there is that privileged relation, and if there is that privileged relation there is survival; and that is so, even in worlds where that relation is utterly inert as regards the person-directed practices and concerns of the community involved. This assumption is practice-independence. And the major aim of chapters 1 to 4 is, first, to out this assumption in all of its possible guises, and then, having outed it, to show how and why it is wrong for the case of personal identity.

To this deconstructive end, chapter 1 is devoted to explaining exactly what the higher-order question is, how it relates to the issue between practice-dependence and practice-independence, and how all this relates to the case of personal identity. Chapter 2 takes a closer look at the various forms practice-independence might take. Despite their methodological differences, what all these forms share (and what makes them practice-
independent) is the assumption that a community’s person-directed practices are no necessary condition for a relation counting as the relation of personal identity for a community. For philosophers of personal identity to date, a community’s person-directed practices are only contingently associated with the concept. This has meant, *inter alia*, that philosophers of personal identity have been faced with the additional task of explaining why it is that personal identity matters to us in the way that it is does. In chapter 3, I discuss the question of what matters in survival and how practice-independent accounts have attempted to close the conceptual gulf left between the question of personal identity and the question of what matters in personal identity. Having so set the scene, in chapter 4, I argue that practice-independence about personal identity is mistaken.

My aim for the remainder of the thesis is more constructive. The considerations that show that practice-independence is the wrong answer to the higher-order question also go to show that practice-dependence is the right answer. Like practice-independence, however, practice-dependence can take a number of more particular forms. In chapter 5, I discuss some of these forms and argue for the one I favour. I call this view pluralism (as distinct from relativism). Briefly and roughly, according to pluralism, survival for a community is whatever (first-order) relation among person-stages it is that plays the survival role in the person-directed practices and concerns of that community. According to pluralism, a relation earns it right to count as the relation of survival for a community in virtue of the fact that that is the relation around which a community organize their person-directed practices and concerns.

Practice-dependence seems to make for a radical new possibility in survival. If there is nothing served up by nature or God which constrains us to think of personal identity in a particular way if we are not to be seriously mistaken—if what relation is personal identity is up to us in important ways—then it seems as if we have the power to change our survival conditions by reorganizing our person-directed practices and concerns in terms of some new relation. In chapter 6, I explain the sense in which practice-dependence makes for this possibility, and the details of the pluralist account of personal identity that I favour. I call this more particular sort of pluralist view, temporal-phase pluralism. In chapter 7, I briefly discuss some
possible objections to this account and to this general way of analyzing personal identity. If pluralism frees us to reorganize our person-directed practice in terms of some new relation, the question arises: exactly what relations are we free to organize around; what is the range of possible natures that survival could have? Chapter 8 explores this question.

The implications of pluralism, however, are not confined to what we say about personal identity in merely possible worlds. Pluralism, I think, has substantial and quite radical implications for what we should say about personal identity in the actual world as well. These implications are the subject matter of chapter 9. Moreover, the recognition that the issue of identity and the issue of social practices are more intimately entwined than we had originally thought has implications that go well beyond the case of personal identity alone. It has, I believe, implications for ongoing debates on issues such as nationalism, ethnicity and gendered identity, amongst others. Analytic philosophy has found few allies in discussions of these issues outside analytic philosophy. Pluralism is a view which emerges from a problem which has long preoccupied analytic philosophers, but it has, I believe, much to offer non-analytic discussions of these issues. I sketch the very bare bones of this potential contribution in a brief, concluding afterword.
1

SURVIVAL AND PRACTICE-INDEPENDENCE

1.1 A brief survey of the personal identity debate.

The question of personal identity is the question of what are the logically necessary and sufficient conditions for a person identified at one time being the same person as a person identified at another.\(^1\) The question of personal identity is thus not the (epistemic or evidential) question of how do we know or tell that someone is the same person over time. It is the question of what personal identity consists in: what it is to be the same person over time.

It is important not to confuse the question of personal identity with the different, albeit related, question of what are the necessary and sufficient conditions for something's being a person. The question of personal identity is not the question of what it takes for something (a foetus, a chimpanzee, a Martian...) to be a person, but rather what makes something already uncontroversially a person at an earlier time the same person as something uncontroversially a person at a later time. When we talk about personal identity we are assuming that the thing identified at the earlier time and the thing identified at the later time are both already, uncontroversially persons. Of course, the two questions, though different, are related. Any account of personal identity will presuppose some basic things about the nature of persons. Psychological continuity accounts, for example, presuppose that persons are the kind of things which are capable of having memories, intentions, beliefs and so on, whilst physical continuity accounts presuppose that persons are physical or otherwise embodied beings. And, of course, both physical and psychological continuity accounts presuppose that persons are a causally efficacious kind. The account of personal identity I shall argue for presupposes

\(^1\)Derek Parfit puts a different spin on the issue, which we can set aside for the moment. According to Parfit, the central issue is what are the logically necessary and sufficient conditions for a person's survival over time; where 'survival' leaves it open that the properties or relations in which personal identity is taken to consist might take a branching form. Survival is just personal identity, without the requirement of uniqueness. The different sorts of analyses of personal identity that I will survey in this section also hold good as analyses of survival with this simple modification: an analysis of survival does not require a clause to rule out branching.
merely the orthodoxy: that persons are self-conscious beings with the capacity for reflective mental life; exemplars of which are possessed of cognitive faculties like the ability to form beliefs and make judgements, and capable of affective or reactive attitudes like desire, concern, love and anger. In discussing personal identity, then, we are assuming that the thing at the earlier time and the thing at the later time are both already persons (in that they meet something like the above criteria), and we are asking what makes those two persons at different times the same person.

There are two sorts of identity over time: numerical identity and qualitative identity. Sometimes when people talk about personal identity they are talking about qualitative identity. As, for example, when people say things of the sort 'She's a different person since having the baby'. They do not mean that having the baby literally killed her, rather they mean her qualitative identity—the sort of person she is, the beliefs, desires and behavioural dispositions that she has—are very different since she had the baby. However, the question of personal identity (or, at any rate, the question with which philosophers of personal identity have been concerned) is the question, not of qualitative identity—what makes a person the same sort of person that they are at or over time—but rather the question of numerical identity—what makes them one and the same person over time. Although a person may be a very different sort of person after having a baby, they are still the numerically same person. We do not regard having a baby as death. Of course, although different, these two questions are not unrelated. Too much qualitative change can destroy numerical identity. For example, if I chopped up a table for firewood, not only would it cease to be the same sort of thing, plausibly, it would cease to be a table altogether. The table would cease to exist. Similarly, some accounts of personal identity have it that too much change in the physical and/or psychological make-up of a person amounts to that person ceasing to exist. After these changes we fail to have numerically the same person as we had before.

Historically, the division between competing positions on personal identity has been, broadly speaking, a division between two kinds of accounts: what Derek Parfit respectively calls non-reductionist accounts and reductionist accounts. The Cartesian view of personal identity, according to which personal identity consists in the continued existence of the same immaterial entity or ego, is probably the most famous example
Chapter 1  Survival and Practice-Independence

of a non-reductionist account. In fact, according to Parfit, it is the only coherent kind of non-reductionist account. What non-reductionists distinctively claim is that personal identity consists in some further fact, over and above, and not reducible to, facts about physical and/or psychological connections and continuities among person-stages. In contrast, reductionists claim that personal identity just consists in, and is reducible to, the holding of certain more particular facts (facts which can be impersonally described); most notably, facts about physical and/or psychological connections and continuities among person-stages, when these hold in a non-branching form.

Here, talk of 'person-stages' should be thought of as a convenient, shorthand way of talking of how a person is at some particular time. Instead of saying 'Is the person who is Ronald Reagan, the actor in 1950, the same person as the person who is Ronald Reagan the forgetful, ex-president in 1995?' we can say 'are the two person-stages stages of the same person?' Understood this way, talk of 'person-stages' is agnostic on the metaphysical issue as between three-dimensionalism and four-dimensionalism. Four-dimensionalists will gloss the notion of a 'person-stage' in terms of temporal parts: a person-stage is a certain temporal part of a person. Three-dimensionalists will think of a person-stage as a person fully present at a time. For the time being, we can set aside these controversial metaphysical issues. Talk of person-stages has no bearing on the issue.

Most philosophers nowadays believe that the non-reductionist view of persons as Cartesian egos or souls is, if not incoherent, then certainly as a matter of fact false. For, in this world at any rate, there are no immaterial entities such as Cartesian egos or souls; and, as Parfit notes, it is hard to imagine what other sort of entity could be the non-reducible locus of mental life. So in recent times, the personal identity debate has chiefly boiled down to a debate between proponents of two different kinds of reductionist views: physical continuity theorists and psychological continuity theorists. What physical continuity theorists maintain is that physical continuity of some sort—typically either continuity of the animal, or continuity of enough of a person’s body, or continuity of enough of a person’s brain—is both necessary and sufficient for personal identity. According to physical continuity theorists, a person Y today is one and the same person as a person X at some past time if, and
only if, (i) Y is the same animal as X, and (usually) (ii) this continuity holds in a one-one or non-branching form. Or, a person Y today is one and the same person as a person X at some past time if, and only if, (i) Y has enough of X's body, and (usually) (ii) this continuity holds in a one-one or non-branching form. Or, as Parfit formulates the physical continuity thesis, a person Y today is one and the same person as a person X at some past time if, and only if, (i) enough of X's brain continued to exist for it to be the brain of a living person, and is now Y's brain, and (usually) (ii) this continuity holds in a one-one or non-branching form.

Of course, physical continuity theorists do not maintain that a person has to retain exactly the same cell for cell physical make-up over time in order to qualify as being the same person over time, for molecules in our bodies are constantly being replaced by new ones. Rather, what physical continuity theorists require is merely that the arrangement of physical matter constituting the person’s brain or body at the later time has resulted from that constituting the person’s brain or body at the earlier time by a series of more or less gradual replacements.

In contrast, for psychological continuity theorists, such physical continuity is either not necessary, or else insufficient (or neither or both) for personal identity. According to psychological continuity theorists, psychological continuity—continuity of memories, intentions, beliefs, desires, and other character traits—is a necessary condition for personal identity over time. For psychological continuity theorists, then, a person X now is the same person as a person Y at some time in the future if, and only if, (i) there is sufficient psychological continuity between X now and Y at that time in the future, (ii) this continuity has the right kind of cause, and (usually) (iii) this continuity takes a one-one or non-branching form. It is important to note that four-dimensionalists such as Lewis (1976) and Perry (1976) do not require the non-branching clause, (iii). (I will say a bit about why they do not require it in Chapter 6.)

Moreover, within the psychological continuity camp, there has been much further debate about what counts as the ‘right kind of cause’. Some psychological continuity theorists, such as Thomas Nagel, maintain that the ‘right kind of cause’ of the furthering of a person’s psychology into the future must be its normal cause i.e. continuity of the same functioning brain. Nagel writes, “What I am is whatever is in fact the seat of the person TN’s experiences and his capacity to identify and reidentify
himself and his mental states, in memory, experience and thought..."\(^2\) (my italics). And since, according to Nagel, continuity of the same functioning brain is what in fact underlies continuity of a person’s mental life, continuity of the same functioning brain is what is essential for personal identity. According to philosophers like Nagel, then, whilst psychological continuity is necessary for personal identity, it is not sufficient; physical continuity is also necessary to underpin psychological continuity. Other psychological continuity theorists, such as Parfit, however, are not so chauvinistic about what counts as the right kind of cause. They maintain that the ‘right kind of cause’ of the furthering of psychological connections into the future could be any kind of cause. As Parfit writes,

The abnormality of the cause seems to me trivial. Reconsider the artificial eyes which would restore sight to those who have gone blind. Suppose that these eyes would give to these people visual sensations just like those involved in normal sight, and that these sensations would provide true beliefs about what can be seen. This would surely be as good as normal sight. It would not be plausible to reject these eyes because they were not the normal cause of human sight.\(^3\)

Just as it would not be plausible to reject artificial eyes as providing something just as good as sight so, Parfit thinks, it would not be plausible to reject cases where psychological continuity is secured without its normal cause as cases of a person’s survival.\(^4\)

Of course, just as physical continuity theorists do not require that a person has to retain exactly the same cell for cell physical make-up over time in order to qualify as being the same person over time, so psychological continuity theorists do not maintain that a person who loses a few memories, or changes a few beliefs or intentions, is no longer the same person. Psychological continuity theorists draw a distinction between direct psychological connections among person-stages and psychological continuity. Direct psychological connections are the connections which hold between an intention and the intended-act,

---

\(^4\)Parfit canvasses a third possible option for psychological continuity theorists—a view, so far as I know, that no psychological continuity theorist has actually held—namely, that the ‘right kind of cause’ is a reliable cause. A reliable kind of cause need not be a normal cause (although the normal cause, continuity of the same functioning brain, is invariably reliable), any kind of mechanism that functions reliably (i.e. in over fifty per cent of cases) to underpin transmission of a person’s psychological characteristics is enough for the preservation of personal identity.
between a memory and the remembered experience, and between different expressions of some persisting character traits, beliefs, and desires. So a person X today will be directly connected to a person Y twenty years ago if and only if, for example, X can now remember having some of the experiences that Y had twenty years ago. But, on a psychological continuity account, even if there were no such direct memory connections between X now and Y twenty years ago, this does not necessarily mean that X is not the same person as Y. For there will be continuity of memory between X and Y if there has been an overlapping chain of direct memories between X and Y. If, in each day within the last twenty years, Y remembered some and enough of his or her experiences on the previous day. So psychological continuity involves the holding of overlapping chains of strong or enough connectedness.

But though there are these substantive points of dispute between non-reductionists and reductionists, and more recently amongst reductionists themselves, over exactly what relations our concept of personal identity picks out, there nevertheless remains a unifying and overarching point of agreement. What theorists of all these persuasions to date have invariably assumed is that *what makes it the case* that the concept of personal identity or survival consists in these relations is entirely independent of any concepts of subjects' person-directed practices viz., our natural and habitual practices of believing and judging that certain people are the same people over time; of identifying and reidentifying them; of trusting them; loving them; punishing them; being angry with them; caring about them; attributing them certain ongoing rights and obligations; ascribing to ourselves and others actions, beliefs, intentions, sensations, thoughts, feelings, memories, perceptions and physical traits; expecting to remember directly our own experiences; having a special and intense interest in what befalls ourselves; self-ascribing our own past actions and experiences; feeling remorse, embarrassment, guilt, regret, pride, shame...and so on. For nearly all these theorists to date the extension of the concept of personal identity is fully determined, in as much as it is determined, by some or other single and non-relative set of

---

5Parfit suggests that we should reserve the terms ‘earlier self’ and ‘later self’ to refer to the degree of psychological connectedness within a person’s life i.e. within the life of a single psychologically continuous person. So, although X and Y here are the same person because psychologically continuous, we could say that because the psychological connections between X now and Y twenty years ago are so weak, X is an ‘earlier self’ and Y a ‘later self’.
facts served up by nature or God alone. So that we can say, independently of any mention of subjects’ person-directed practices and concerns, what makes it true that two person-stages X and Y are stages of the same person. What Cartesians claim is that what makes two person-stages X and Y stages of the same person is entirely to do with whether Y has X’s soul, irrespective of whether or not subjects’ organize their person-directed practices and concerns around the soul. And what reductionists have claimed is that whether two person-stages X and Y are stages of the same person depends entirely upon whether Y is physically and/or psychologically continuous with X (and, on some accounts, whether Y is the later person who is most physically and/or psychologically continuous with X), quite independently of the way in which subjects’ organize their person-directed practices and concerns.

And, what both reductionists and non-reductionists have claimed is that it is when, and only when, the appointed facts obtain that people like us are justified or correct in applying the concept of personal identity. People of our biological kind who organized their person-directed practices and concerns around the assumption that a person at a later time was the same person as a person at an earlier time when the person at the later time was not psychologically continuous with the person at the earlier time (for psychological continuity practice-independent theorists), or not physically continuous with the person at the earlier time (for physical continuity practice-independent theorists), or did not continue to possess their soul (for Cartesian practice-independent theorists), would either be mistaken or would mean something significantly different by ‘personal identity’. Hence, though they differ as to exactly what the facts of personal identity are, what nearly all theorists of all these persuasions to date have assumed is that there is a single, non-relative answer to the question of what makes these set of properties (rather than some other(s)) count as the facts of personal identity entirely in terms of some or other set of facts served up by nature (or God). For both reductionists and non-reductionists to date, our person-directed practices are only contingently or accidentally connected to the concept of personal identity. They are no essential or necessary feature of the concept itself.

It may help here, in order to get clearer about the precise sense in which most accounts of personal identity to date have been, what I am
calling, practice-independent, to set out the general issue of practice­
independence in a little more detail.

1.2 Practice-independence versus practice-dependence.

Any adequate piece of conceptual analysis must put us in a position to
answer (at least!) the following three questions. The first of these is a
question about which things fall under the extension of a given concept:
for example, which person-stages are stages of the same continuing
person; which objects are red; which acts are good? The second question
is a question about what property (or relation) determines which things
fall under the extension of the concept in question: what property is it in
virtue of which these objects are red; what property determines which acts
are good; what relation between person-stages determines that two
person-stages are stages of the same continuing person? The third
question is a question about what justifies our answer to the second
question: in virtue of what does this relation (whatever it is) determine
that two person-stages are stages of the same person; in virtue of what
does this property (whatever it is) determine that all and only acts with
this property count as the good acts; in virtue of what are these properties
(rather than some other(s)) the red properties? The first question is about
which things fall under the extension of a given concept; the second, a
question about what property unites these things under a given concept;
and the third, a question about what justifies us in maintaining that things
with that properties (rather than some other(s)) count as the red things, or
the good things, or stages of the same person.

Here’s a commonplace example of the three questions in action.
Suppose we want an analysis of middle-eastern cooking: that is, we want
an answer to the question ‘which culinary dishes are middle-eastern?’
This question—which items of food are middle-eastern—is the first of our
aforementioned three questions; a question about which things fall under
the extension of the concept of middle-eastern cooking. However, in
order to answer this first question, we need first to answer the second of
our aforementioned three questions: what is the property of dishes in
virtue of which they count as middle-eastern dishes (as opposed to Asian
dishes, or western dishes, or some other sort of dish)? What unites certain
different token dishes under the extension of the type, middle-eastern
cooking? We cannot know which dishes are middle-eastern until we have
in place some story about what makes certain dishes count as middle­
eastern in the first place. Suppose we come up with the following story as an answer to this second question: suppose we say 'middle-eastern cooking is whatever sort of dishes people who live in the middle-east cook'. Now, and only now, are we in a position to say which culinary dishes are middle-eastern: *a posteriori*, they will be all and only those culinary dishes that people in the middle-east in fact cook. So, for example, falafels count as middle eastern by this definition, but yorkshire pudding and Asian noodles do not, since falafels are in fact cooked in the middle-east, whilst yorkshire pudding and asian noodles are not. We can only get a story about which things our concepts pick out (in this case, falafels and the like) by way of first answering the second question of what makes it the case that our concepts pick out those things (in this case, what property of dishes makes them count as middles-eastern dishes). Or, to put matters a slightly different way, we can only get a story about which things fall under the extension of a given concept (in this case, falafels and the like) after first answering the second question of what determines the extension of the concept of middle-eastern cooking (what determines the extension of middle-eastern cooking such that falafels but not Asian noodles are the things which fall under its extension?).

However, of this answer to the second question, someone might ask 'but *why is that* property of culinary dishes—the property of being cooked in the middle-east—the property that determines which items of food are middle-eastern?' In virtue of what does that property get to determine the extension of middle-eastern cooking? For surely, they might go on to argue, the dishes cooked by Arab emigres living in London or elsewhere overseas should count as middle-eastern just as well. This would be a question of the third kind: a question about what justifies us in defining middle-eastern cooking in terms of dishes having the property of being cooked in the middle-east, rather than in terms of some other property, say, the property of being cooked in the middle-eastern culinary tradition. Disputes at this third-level contrast with disputes at the second and first levels in that they are typically *a priori*, not *a posteriori*, disputes: disputes not about matters of fact (for example, whether or not Asian noodles are in fact cooked in the middle-east), but about the conceptual adequacy of the criteria for determining which dishes are middle-eastern.
Of course, sometimes there is no room to ask this third question. Consider, for example, what it is for a thing to be a triangle. Which things are the triangles? Answer: those things, and only those things, with three angles and straight sides. In the case of triangles, to go on to ask 'but why does the property of having three angles and straight sides determine which things are triangles and which not—why, for example, shouldn't things having the property of four angles and straight sides count as triangles?' is to ask for too much. For, in the case of triangles, a fair answer to this question seems to be 'but that's just what we mean by 'triangle'. There is nothing further to be said about triangles; our conceptual spade hits bedrock at this second level. Not all of our concepts, however, (in fact, very few of them), are like that of triangle. As we shall see, for many of our concepts, in particular, our richer ones, there is something very important and interesting to be said in answer to this third question.

From hereon in, call this third-question—the question of in virtue of what certain properties get to determine the extension of a concept—the meta-question. Call the second-question—the question of what properties of things determine that things with all and only those properties fall under the extension of a concept—the object-level question. (Since, for reasons I will discuss at the beginning of Chapter 2, I will not be directly interested in the first question—the question of what the things that fall under the extension of a concept are—I will not bother to give this first question a label.) It is important, however, not to let these labels confuse us. I have called the second question the 'object-level' question because it is a question about properties of objects or relations between object-stages. It is important to bear in mind, however, that it is not the first question of what the objects are. Similarly, I have called the third question the 'meta-question' because it is a higher-order, conceptual question about the properties or relations that determine that certain objects or object-stages fall under the extension of a concept, even though there is a sense in which the second, object-level question is a meta-question, for it is higher-order question about the first. Perhaps other terminology would be better, but none springs readily to mind. And

---

6I owe this point to David Braddon-Mitchell. However, some may think that rule-following considerations show that we can ask this third question of any of our concepts, including 'triangle'.

these labels should do the job, so long as we bear in mind exactly which question they are labels for.

We are now in a position to formulate more precisely the difference between practice-independent and practice-dependent accounts, as I am using these terms. What marks off practice-independent accounts from practice-dependent accounts is the kind of story told with respect to the *meta-question*; with respect, that is, to the question of what makes it true that *this* property (or relation), rather than some other, makes it the case that personal identity (or middle-eastern cooking, or...) picks out certain things. The dispute between practice-dependent and practice-independent accounts is a dispute, first and foremost, about what makes it true that a certain property (or relation) determines the extension of a given concept, *not* about which things fall under that extension (the first question), or what the property in virtue of which they fall under that extension is (the object-level question). Practice-independent accounts tell a story about what makes it the case that personal identity consists in a certain relation among person-stages which makes no essential reference to subjects' responses, or practices, or concerns; a story in terms of independent facts served up by nature or God alone. In contrast, practice-dependent theorists tell a story about what makes it the case that personal identity consists in a certain relation or set of relations among person-stages essentially in terms of subjects' responses or practices of some sort. Practice-dependent theorists regard our practices as determining what determines the extension of a given concept; practice-independent theorists maintain that what determines what determines the extension of a given concept are facts which obtain or fail to obtain essentially independently of our practices.

Now, of course, there is a trivial way of reading these claims: how we use our terms or concepts—which words or concepts we use to refer to which things, and properties of things, in the world—is up to us. We could, after all, have used the word/concept 'schmold' rather than 'gold' to refer to things with atomic number 79. And the story about why we use the word 'gold' rather than 'schmold' to pick out things with atomic number 79 is clearly going to make essential reference to our practices, in particular our linguistic ones. On this trivial reading, then, all conceptual

---

7 Although, of course, disputes at the meta-level may flow through to, or be reflected in, disputes at the object- and first levels as well.
analyses will turn out to be practice-dependent. However, conceptual analysis in general is not concerned with the question of why we happen to use the words we do; and the debate between practice-dependent theorists and practice-independent theorists is not a debate about the philosophically (though not anthropologically) uninteresting question of why we use the words we do in the way we do. What is at issue is much more substantial than that. And it is something that might perhaps be better captured like this: For any given concept, is there something in the world—some natural grouping—which demands that we group those things together in the extension of that concept, if we are not to miss something very intrinsically important about the way the world is, or, is the grouping more or less up to us, our interests, concerns and practices? I say ‘more or less’ because, of course, the world places some constraints on the ways in which we can group things. What practice-dependent theorists claim is that world constraints alone leave the extension of the concept in question crucially underdetermined.

Take the case of elements, for example. Everyone (that is, everyone who believes that there are such things as natural kinds) believes that someone who had a conceptual scheme which failed to group elements in accordance with their atomic number would miss something very important about the way the world is. So that people who went around grouping elements together solely in accordance with their atomic weight, for example, would be failing to describe something intrinsically important and interesting about the way the world is. They would be missing a natural joint. However, just as nearly everyone agrees that, as regards the case of elements, the world demands that we group things together in a certain way, so nearly everyone agrees that, as regards concepts such as embarrassing, it is not the intrinsic properties of things in the world, but rather our relationship to things in the world—our behavioural dispositions—wherein the commonality of interest lies. Which things are embarrassing—which properties of things it is in virtue of which they count as embarrassing—are whichever properties of things it is to which we are normally disposed to respond with embarrassment. There is no natural ‘embarrassing’ joint in the world which independently demands that a certain group of things count as embarrassing. Or, at any rate, to the extent that there is a natural ‘embarrassing’ joint, it is a behavioural joint: a joint defined by our behavioural dispositions, not by the intrinsic properties of things. What it is for a thing to be embarrassing
Chapter 1 Survival and Practice-Independence

is, first and foremost, for it to be the sort of thing to which normal people are disposed to respond with embarrassment.

Debates between practice-dependent and practice-independent accounts of concepts are rife in philosophy, though they go under a variety of different names. Consider the debates surrounding the following question in meta-ethics: What makes it the case that value concepts, such as good, necessarily consist in a certain property of acts, say, the property of being courageous? What makes it true that acts with certain properties, such as the property of being courageous, fall under the extension of the concept good, whilst acts with other properties, say, maximising pain, do not? According to Aristotle and some contemporary virtue ethicists, for example, good necessarily consists in a certain property—in this case, the property of being an act which is courageous— not because subjects approve of courageous acts, or are motivated to pursue acting courageously, or care about acting courageously, but because courage is an ultimate human virtue⁸; and would be an ultimate human virtue even if it were the case that subjects were not disposed to respond to it in any of the aforementioned ways. According to Aristotelian virtue ethicists, an act X is good if, and only if, X has, inter alia, the property of being a courageous act. Acts which have the property of being cowardly do not fall under the extension of the concept good because cowardice is not such an ultimate virtue. For Aristotelian virtue ethicists, then, what makes it the case that acts which have the property of being courageous fall under the extension of the concept good has nothing essentially to do with subjects’ responses. A complete analysis of what it is for an act to be good need make no essential reference to subjects’ practices. Aristotelian virtue ethics, then, is a practice-independent account of good.

Contrast this Aristotelian story with the following kind of story: What makes acts with a certain property or set of properties count as good is just that those are the properties which subjects in certain specified conditions are disposed to approve of and motivated to pursue: X is good, we might say, if, and only if, subjects in ideal conditions are disposed to

---

⁸Very crudely, Aristotle’s gloss on this was that substances, such as persons, have natural ends. These natural ends are the values; and the virtues—courage, generosity, and so on—are the means to promote or secure these natural ends. The natural ends, and the virtues which promote them, for Aristotle, count as values and virtues, whatever people should happen to care about, or approve of, or be motivated to pursue.
approve of X. Now, it may turn out, as a matter of contingent fact, that all and only those properties of acts that such subjects are disposed to approve of, are those of being courageous. In which case, Aristotelians and proponents of this latter story will agree about which things are actually good, at least in this world. (They will disagree of what to say of other worlds in which subjects happen not to approve of courageous acts). They will agree, that is, about the answer to the first of our questions.

Will they also agree about the answer to the object-level question—the question of what properties of acts in is in virtue of which they count as good? The short answer is that they may or they may not, depending on the more precise view the practice-dependent theorist takes. This marks an important difference between two sorts of practice-dependent accounts, and we will come back to it in greater detail in Chapter 5. The practice-dependent theorist has two options at this second level: either they may (contingently) identify good with the particular property that subjects in ideal conditions happen to approve of (in this case, courage), or they can identify it with a second-order property viz. the property of being a property that subjects in ideal conditions approve of. If the practice-dependent theorist takes the first line, they will agree with the Aristotelian about what property it is in virtue of which acts count as good (although, of course, they will disagree about the modal status of the property: Aristotelians will identify good with courage necessarily, whereas practice-dependent theorists will identify good with courage only contingently). If, however, the practice-dependent theorist takes the second option, and views good as the second-order property, they will disagree completely with the Aristotelian about the answer to this second of our three questions.

Nonetheless, whether or not the practice-dependent theorist of good agrees with the Aristotelian about the properties in virtue of which acts count as good, the two sorts of accounts fundamentally disagree about the answer to the meta-question of what makes it the case that good picks out the property of courage. Aristotelian virtue ethics is a practice-independent view—it claims that good picks out the property of courage because courage is an absolute, fundamental virtue—whereas the latter view is a practice-dependent view—it has it that good picks out acts with the property of courage (where it does, in fact, pick out acts with that property) because that is the property that ideal subjects in ideal
conditions happen to approve of. For practice-dependent theorists of good, courage has no claim to constitute the property of goodness independent of subjects’ dispositions to respond to that property with approval.

Or, for another example altogether, consider the debate over the nature of colour concepts. A practice-independent account of our colour concepts will tell a story about what makes it the case that, say, red consists in certain properties that makes no essential mention of observers, their colour experiences, or their colour-directed practices; a story entirely in terms of certain reflectance properties of objects in the world. As David Armstrong, for example, maintains, the concept of colour “does not yield any necessary connection between red objects and any sort of perceptual experience, such as looking red to normal observers in normal conditions.” According to Armstrong, what it is for a thing to be red is just for that thing to have a certain reflectance property: X is red, theorists of this ilk will claim if, and only if, X has (for example) reflectance property r. The story these philosophers tell about what makes it true that red consists in those reflectance properties makes no essential mention of observers, their colour experiences, or their colour judgements or dispositions to make certain colour judgements. The colour experiences or judgements of observers may function instrumentally as a reference-fixing device in this world, but they are no essential or necessary feature of colour concepts. For go to a world in which that property exists but does not cause objects to look red to normal observers in normal conditions (but, say, blue) and that property counts as red nonetheless. For Armstrong and other practice-independent theorists of colour concepts, a complete analysis of what it is for a thing to be red need make no essential reference to observers’ colour experiences or colour judgements.

In contrast, for practice-dependent theorists of colour concepts what determines that the extension of colour concepts such as red are determined by a certain property of objects is essentially to do with our colour experiences, perceptions or judgements. What it is for a thing to be red is, first and foremost, for it to be the sort of thing that produces red

9 Often this debate, as indeed the debate in meta-ethics, is characterized as a debate between subjectivists and objectivists. I am reluctant to use these terms since I think they carry with them some misleading connotations for our purposes.

colour experiences or prompts red colour judgements in normal observers in the appropriate viewing conditions. As Jackson and Pargetter put it, "the fundamental ground for ascribing a certain colour to something is the colour it looks to have". For these theorists, an object X is red if, and only if, X looks red to normal observers in normal conditions, or normal observers in normal conditions judge X to be red, or there is some property which causes X to look red to normal observers in normal conditions. On all these accounts, the colour experiences or judgements of normal observers in normal conditions are a necessary feature of colour concepts. Go to a world in which the same property of objects that causes objects in this world to look red to normal observers in normal conditions causes them to look blue to normal observers in normal conditions, and that property counts as blue in that world. The debate between practice-dependent and practice-independent theories of colour, then, is a debate about whether or not observers, their colour experiences or their colour judgements, need figure as an essential part of the answer to the question of what it is for a thing to be red.

1.3 Response-dependence

The sort of accounts of value and colour that I have called 'practice-dependent' have elsewhere been called 'response-dependent' accounts. In this complex and rapidly expanding literature, the claim that a concept is response-dependent is understood roughly as follows. A concept is response-dependent just in case there is a non-trivial, a priori conceptual connection between that concept and some relevant human response. More precisely, as Mark Johnston puts it,

If C, the concept associated with the predicate 'is C', is a concept interdependent with or dependent upon concepts of certain subjects' responses under certain conditions [i.e. a response-dependent concept] then something of the following form will hold a priori

\[ x \text{ is } C \text{ iff in } K, S's \text{ are disposed to produce } x\text{-directed response } R \]

(or \( x \) is such as to produce \( R \) in \( S's \) under conditions \( K \).)\(^{13}\)

---

\(^{13}\)Johnston (1989b), p. 145
So, the claim is, if something like the following biconditional holds substantively, *a priori* true, then the concept in question (in this case, *red*) is a response-dependent concept:

\[
x \text{ is red iff } x \text{ is disposed to look such and so (ostended) way to standard perceivers as they actually are under standard conditions as they actually are.}
\]

In contrast, as Richard Holton points out, a similar biconditional equation for the concept of square:

\[
x \text{ is square iff in } K, S's \text{ are disposed to, say, judge } x \text{ to be square,}
\]

would not hold *a priori* but, at most, *a posteriori* true.

Of course, as Holton points out, one could construct an equation for the concept of square which did hold *a priori* true in virtue of a trivialising reading of K, S, or R:

\[
x \text{ is square iff } x \text{ would be judged to be square by observers who are accurate at identifying square things in circumstances which are propitious for doing so,}
\]

is just such a trivialising reading. So, in order for a biconditional equation to be true to the response-dependent nature of a concept, K, S, and R must not be given a trivial, 'whatever-it-takes' reading which either overtly or covertly specifies the conditions, subjects, or responses, in whatever way is required to make the equation come out true.

On this way of thinking about these issues, then, the claim that a concept is *not* response-dependent amounts to the claim that such an equation only holds a priori true in virtue of a trivial or 'whatever-it-takes' formulation of the C-conditions or that such a biconditional, if true at all, holds at most *a posteriori* true.

'Practice-dependence' and 'response-dependence' are synonyms—they stand for exactly the same main idea. It is just that I find it easier and clearer to frame the idea in terms of the three questions I mentioned at the outset. The term 'practice-dependence' signals this difference of explanation (and my unwillingness to get involved in the various internal debates about response-dependence, which we can set aside). But, for
those who are familiar with the literature on 'response-dependence', the central idea is the same.

1.4 Practice-dependence again

A good way to understand the structure of the debate between practice-dependent and practice-independent meta-theories in general is to recall Plato's Euthyphro. Recall, in Euthyphro, Socrates and Euthyphro meet at the entrance of the law courts where Euthyphro is prosecuting his own father on what even he admits is a very dubious charge of murder. When Plato's Socrates discovers this he is astonished:

SOCRATES: But you, by heaven! Euthyphro, you think that you have such an accurate knowledge of things divine, and what is holy and unholy, that, in circumstances such as you describe, you can accuse your father? You are not afraid that you yourself are doing an unholy deed? Euthyphro replies that, with his special insight into right and wrong, he knows that, in prosecuting his own father, he is acting in the spirit of true piety. Socrates then challenges Euthyphro to define this true 'piety' which he claims justifies him in prosecuting his own father on such a dubious charge of murder:

SOCRATES: Well, then, show me what, precisely, this ideal is, so that, with my eye on it, and using it as a standard, I can say that any action done by you or anybody else is holy if it resembles this ideal, or, if it does not, can deny that it is holy.

EUTHYPHRO: Well, then, what is pleasing to the gods is holy, and what is not pleasing to them is unholy.

What Euthyphro appears to be claiming is that the gods' loves themselves determine the properties of things in virtue of which they are holy or pious and the properties of things in virtue of which they are not. Socrates doubts this; and he goes on to pose the following question:

---

16ibid., p. 174.
SOCRATES: We shall soon know better about that, my friend. Now think of this. *Is what is holy holy because the gods approve it, or do they approve it because it is holy?* 17 (emphasis added)

This question is left unresolved in Euthyphro, but its interest for us lies in the fact that the structure of the question Socrates poses is precisely the structure of the debate between practice-dependent and practice-independent theorists, here with regard to the concept of piety. The dialogue between Plato’s Socrates and Euthyphro was a debate over precisely the meta-issue of whether the facts about piety were determined independently of subjects’ practices, in this case the practices of the gods, or whether the facts about piety were determined by the gods’ practices: Is it because certain acts are pious that they are loved by the gods, or, is it because they are loved by the gods that certain acts count as pious? Plato’s Socrates and Euthyphro agreed that an act was pious if, and only if, the gods loved it. What they disagreed about was, if you like, the direction of determination. Socrates was advancing, or at least suggesting, a practice-independent account of piety: it is because certain acts are pious that they are loved by the gods. On this, practice-independent view, the gods’ loves did not determine that a certain group of acts count as pious (for that was determined quite independently of what acts the gods happen to love), rather, it was that the gods were very good trackers of the independently determined facts of piety. And the fact that the act was independently pious could figure as part of a (causal) explanation as to why the gods loved it. For the gods love pious acts. Euthyphro, however, maintained exactly the reverse: it is because they are loved by the gods that certain acts are pious. What makes certain acts count as pious is not determined by facts independent of gods (the gods do not love pious things because they are, independently, pious), rather it was just because the gods loved certain acts that they counted as pious. On this, practice-dependent view, the gods’ loves themselves make it the case that certain acts are pious, namely, exactly all and only those acts that the gods love. Practice-dependent views, then, claim that what makes it the case that a thing counts as ‘good’, or ‘pious’, or ‘red’ or the like, is determined by certain of the practices—judgements, perceptions, loves or the like—of certain subjects in suitably privileged conditions: something is red, good or pious because subjects, in suitable conditions, are disposed to see it as red, judge it to be good, or love it, for example. Practice-independent

17 ibid., 178.
theorists maintain exactly the reverse. For these theorists, insofar as subjects in suitably privileged conditions take things to be good, or red or pious, it is because those things are good, or red or pious. Just so, we might pose the question at issue for the case of personal identity: is what is personal identity personal identity because subject’s organize their person-directed practices around it or do/should subjects organize their person-directed practices around it because it is personal identity?

What, then, is there to choose between these two different kinds of meta-stories, practice-dependent and practice-independent? Why should we prefer one kind of meta-story to the other in any given case? The beginning (though by no means the end) of a very plausible answer goes by way of pointing to an intuitive difference amongst certain of our concepts. As Philip Pettit puts it, the difference between those concepts “that have a tenure in nature” and those concepts “whose tenure is tied to our interests and sensibilities”.18 The point is that some of our concepts seem to implicate subjects and subjects’ practices in a peculiarly intimate, seemingly inextricable, way; a way in which other of our concepts do not. (We saw this earlier with the contrast between the concept of embarrassing and concepts of the elements on the periodic table). Like the concept of embarrassing, concepts such as annoying, irritating, boring, exciting, humorous, entertaining, smooth, spicy, bland, beautiful, ugly, comely and so forth are not somehow served up by nature—there is no natural joint in the world which renders certain things intrinsically disgusting, annoying, embarrassing and the like—they constitutively involve our practices.19 What it is for a thing to be annoying, or embarrassing, or irritating or boring or the like, is first and foremost for it to be the sort of thing to which subjects’ in appropriate conditions are disposed to respond to with annoyance, or embarrassment, or irritation, or boredom. Other concepts, however, do not seem to implicate subjects and their practices in this same intimate way. Concepts such as carbon,

---

19To say that what things are annoying, disgusting, irritating embarrassing and so on, depends essentially on subject’s dispositions to respond to those things with annoyance, disgust, irritation, embarrassment and so on, is not necessarily to deny that there is a thoroughly objective, reductive story to be told about the properties of things to which people are disposed to respond with annoyance, disgust, irritation, embarrassment and so on. That is to say, there will be a thoroughly objective story to be told about what are the Fs where the Fs are, for example, the annoying things. We can look to the world, observe what properties of things cause people in the appropriate conditions to respond with annoyance, and things with those properties are the annoying things.
squareness, hardness and mass, for example, seem to implicate subjects and their practices at best only contingently or accidentally. What it is for a thing to be square, for example, seems to be determined quite independently of subjects' experiences of squareness, or their beliefs and judgements about squareness.

Another way we might get at the distinction between those concepts which essentially involve subjects' practices and those that do not, is to think of the former as concepts for which Moore's famous open-question argument has some intuitive force, and the later as concepts for which the argument intuitively does not. Recall, the open-question argument is that given any (reductive) analysis, Y, of a (normative) term, X, it still remains a substantial or 'open' question as to whether X is Y. So, for example, given the following analysis—an act X is good iff X maximizes happiness—it is still possible to ask in a substantial, non-trivial way, whether maximising happiness is good. Moore's point is not that such reductive analyses of normative terms will always admit of counterexamples (i.e. cases where something maximises happiness but we are reluctant to call it 'good', or vice versa), rather it is that reductive analyses of normative concepts in purely descriptive terms seem always to leave something absolutely crucial out. On one gloss, what such reductive analyses leave out is the normative element which gives us a sense of why we should care one way or another about those properties: why should we be motivated to pursue those properties; why should we be disposed to praise those who pursue those properties and condemn those who do not, for example. Why should we care one way or another about maximising happiness, or be motivated to maximise happiness, or condemn those who act in ways which systematically fail to maximise happiness? Why should maximising happiness, rather than some other property, say, exercising valour, figure in the extension of the concept good? What's so special about maximising happiness? It is not enough simply to reply that we should care about those properties because those

20Indeed, as Mark Johnston (1989b) suggests, the difference between those concepts that seem constitutively to involve subjects' practices and those which seem only contingently to involve them is one way of thinking about the intuition which lies at the heart of the distinction between primary and secondary quality concepts.

21Moore (1929), esp. pp.36-58. However, this may be only a rough test. Some, for example, think that Moore's open-question argument holds for consciousness but that this does not show that consciousness is essentially normative.
are the properties in which good consists. As John Perry writes as regards the concept of personal identity,

That I will be run over by a truck means...that the person who is run over by a truck will remember thinking and doing what I am thinking and doing now. But why would I care especially about that? Why should a person who is having such memories be of any more concern to me than anyone else? One is inclined to respond, “because to have such memories is just to be you”; but now the explanation goes the wrong way round; isn’t it fair to demand that the analysans shed light on why the analysandum has the implications for us that it does?22

Now, Moore himself thought that what the open-question argument showed, inter alia, was that good is unanalysable: good is a simple, irreducible, non-natural property. But this is a non-sequitur. What the open-question argument really shows is not that good is a simple, unanalyzable, non-natural property but that no non-normative or straightforwardly descriptive analysis will suffice for concepts such as good. For no such analysis will be able to capture the ineluctably normative nature of the concept. What Moore’s argument shows, then, is that if our analysis of normative concepts such as good is to capture the nature of the concept, we must analyze these concepts in a way which accords normative considerations an appropriately central role in the analysis. Exactly how conceptual analysis can reconcile descriptive reduction with this normative element will, I hope, become apparent in Chapter 7 if not before.

How does all this connect up with the issue of practice-dependence and practice-independence? Well, it’s a truism that conceptual analysis should aim to capture and explicate the nature of the concept in question. That’s what conceptual analysis, at base, is: an account of the meaning or nature of the concept in question. For those concepts which intuitively implicate subjects’ practices non-accidentally or constitutively, then—those, as we might call them ‘normative’ concepts—we should tell a practice-dependent meta-story. This, I take it, is the proper upshot of the ‘open-question’ argument. For those which do not—for those concepts which ‘have their tenure in nature’ rather than in ‘our interests and sensibilities’—we should tell a practice-independent meta-story.

22Perry (1976), p.68.
Most philosophers are happy to grant that concepts such as nauseating, disgusting, irritating, tasty, embarrassing and the like are normative or practice-dependent concepts which should be given a correspondingly practice-dependent analysis. These are concepts which, as Mark Johnston puts it, "wear their response-dependent nature on their face". Many, too, are happy to think that an analysis of this sort should be extended to secondary quality concepts such as colour and value as well. For these concepts are, in sufficiently salient respects, in the same conceptual boat as the first. But that personal identity too might be such a practice-dependent or normative concept might not be so initially intuitively obvious. And so it may require a little more philosophical work to make it apparent.

1.5 Locke: An unexpected ally.

This claim—that the concept of personal identity, unlike concepts of squareness and the like, is a normative concept which should be given a correspondingly practice-dependent analysis—though rare in the contemporary debate about personal identity, is certainly not entirely unprecedented. Famously, John Locke claimed that the concept of personal identity was above all a ‘forensic’ concept, appropriating actions and their merits, being, by definition, that which grounds moral responsibility: “In this personal identity is founded all the right justice of reward and punishment; happiness and misery being that for which everyone is concerned for himself...” For Locke, this forensic role was precisely what marked off the concept of personal identity from concepts of the identity over time of ‘man’ and ‘substance’. The identity of the same man (human being), like the identity of brutes (animals!) over time, consists “in nothing but a participation of the same continued life, by

---

24And some have even claimed that causation and modal notions such as possibility and necessity should be given this kind of analysis.
25Locke (1961), Book II, Chapter XXVII, 26: 291: “Person...is a forensic term, appropriating actions and their merit, and so belongs only to intelligent agents, capable of a law, and happiness and misery.”
26Locke (1961), Book II, Chapter XXVII, 18: 287.
27Locke (1961), Book II, Chapter XXVII, 7:278: “it is not...unity of substance that comprehends all sorts of identity...but...we must consider what idea the word it is applied to stands for: it being one thing to be the same substance, another the same man, and a third the same person, if person, man, and substance are three names standing for three different ideas; for such as is the idea belonging to that name, such must be the identity, which if it had been a little more carefully attended to would possibly have prevented a great deal of...confusion...especially concerning personal identity...”
constantly fleeting particles of matter, in succession vitally united to the same organized body." 28 In contrast, the self is "that conscious thinking thing...which is sensible or conscious of pleasure and pain, capable of happiness or misery, and so is concerned for itself, as far as that consciousness extends." 29 Person, he writes,

is the name for this self...This personality extends itself beyond present existence...only by consciousness: whereby it becomes concerned and accountable, owns and imputes to itself past actions, just for the same ground and for the same reason that it does the present. 30

Locke was, as Alston and Bennett point out, a 'conceptual pragmatist' who believed that the extension of our concepts are as they are, not because they have some privileged or independent claim to carve up nature at its joints, but because it best serves our interests and activities to give them the extension we do. Our concepts carve up the world in this way, rather than some other out of the multitude of possible carvings, because carving up the world this way rather than another best suits our concerns and practices. Given Locke's general conceptual pragmatism, we can understand his claim that personal identity is a forensic concept as the claim that the extension of our concept of personal identity picks out the things it does rather than some other set of things not because those things have any privileged, independent, natural claim to be the facts of personal identity, but because incorporating these things rather than those in the extension of the concept of personal identity best suits the interests of law and morality. 31

Central to the interests of law and morality is that personal identity turns out to consist in a relation that matters to us. Hence, as Harold Noonan stresses, one of Locke's chief concerns was to provide an account of personal identity which could explain why personal identity matters to us: which explained why we can be indifferent as to the punishment or reward received by others, but cannot be so indifferent as to our own fate. In particular, Locke wanted an account of personal identity which could ground and explain why we cannot help but be specially concerned about the prospect of punishment or reward for acts which we acknowledge and appropriate as our own. Something which, Locke was the first to point

31Noonan (1989) also suggests this way of understanding Locke's claim.
out, a view of personal identity as consisting in continuity of same (immaterial or material) substance rendered problematic, indeed, he thought, incomprehensible. He writes

Let anyone reflect upon himself and conclude that he has in himself an immaterial spirit...; let him also suppose it to be the same soul that was in Nestor...at the siege of Troy; but he now having no consciousness of any of the actions...of Nestor...does or can he conceive himself the same person with [him]? Can he be concerned in [his] actions, attribute them to himself, or think them his own, more than the actions of any other man that ever existed? [N]o more...than if some of the particles of matter that were once a part of Nestor were now a part of this man: the same immaterial substance, without the same consciousness, no more making the same person...than the same particle of matter, without consciousness, makes the same person.

This Locke took to constitute a reductio of the supposition that personal identity could consist in sameness of either material substance or immaterial substance. For neither, in the absence of sameness of consciousness, could account for the role that personal identity plays in our forensic practices, including the special and intense kind of interest that we have in our own identity over time. The mere fact that I become convinced that I am immaterially or partially materially identical with Nestor, Locke thinks, would not make me any more interested in Nestor than I would be interested in any other person from the past; nor could it cause me to 'own and impute' to myself his past actions: regarding the things that he has done with the kind of pride, or shame, or remorse, or embarrassment, or the like, that typify my retrospective attitudes to my own remembered past. Likewise, if I came to believe that some future person was going to be immaterially or materially identical with me now, I would not, indeed could not, as a result become specially concerned for that future person in the same way that I am ordinarily specially concerned for my own future self and what befalls me. I would not, and could not, feel towards her future actions and experiences the kind of

---

32 As the following quote makes apparent, Locke was here assuming that the soul's continued existence might be independent of sameness of material substance and sameness of consciousness. Indeed Locke (Book II, Chapter XXVII, 14: 284) says as much for the case of material substance "...for souls being, as far as we know anything of them, in their nature indifferent to any parcel of matter, the supposition has no apparent absurdity in it...". Locke's thought was that any relation, such as sameness of immaterial spirit, which could come apart from sameness of consciousness was not an acceptable candidate relation for personal identity, for such a relation could not play the forensic role definitive of personal identity. I will discuss this issue more fully in Chapter 8.

33 Locke (1961), Book II, Chapter XXVII, 14: 284.
future-oriented attitudes of trepidation, anticipation and the like, which typically characterize my attitudes to my own future actions and experiences. According to Locke, sameness of consciousness is required to underpin and make sense of these concerns:

For as to this point of being the same self, it matters not whether this present self be made up of the same or other [material or immaterial] substances, I being as much concerned and justly accountable for any action that was done a thousand years since, appropriated to me now by this self-consciousness, as I am for what I did the last moment.³⁴ (Later emphasis mine)

So, Locke thought that a necessary condition for sameness of person over time was sameness of consciousness. (And this is the Locke with which commentators are most familiar.) But, and here's the crucial question, why did Locke think that sameness of consciousness was necessary for personal identity? Clearly not because, as many (psychological continuity) practice-independent theorists have it, sameness of consciousness (memory) has any privileged, independent, natural claim to be the relation in which personal identity consists. For, in insisting that personal identity was above all a forensic concept, Locke plainly conceived of our practices of punishment and related person-directed concerns as an essential defining feature of the concept of personal identity, and not something merely accidentally or contingently associate with it. For Locke, sameness of consciousness, rather than sameness of substance or..., determines the extension of our concept of personal identity, not because sameness of consciousness has any independent claim to determine the extension of the concept, but because sameness of consciousness, rather than sameness of substance or any other of the superabundant array of possible candidates, best suits the interests of morality and law. Sameness of consciousness, then, for Locke, was certainly necessary for personal identity, but it was clearly not sufficient. It was necessary for personal identity, but only because, as Locke thought, sameness of consciousness was necessary in order for personal identity to play the forensic role definitive of the concept. For Locke, then, our moral and forensic practices, not nature or God, determined what determined the extension of the concept of personal identity; and what determined that extension was necessarily sameness of consciousness.

³⁴Locke (1961), Book II, Chapter XXVII, 16:286.
Hence, though many psychological continuity practice-independent theorists nowadays are happy to call themselves 'Lockeans' or 'neo-Lockeans', they have overlooked a crucial aspect of Locke's thought. For Locke, though certainly a psychological continuity theorist of sorts, was not a practice-independent theorist of personal identity. Locke was, quite clearly, offering a sort of practice-dependent account of personal identity.  

So, to sum up so far. The debate between practice-independent and practice-dependent theorists has usually been couched, where it has been explicitly couched at all, as a debate over whether certain of subjects' practices (in certain specified conditions) are constitutive of a given concept or, at best, only contingently associated with it. And so I couched the debate at the outset as well for the case of personal identity. We are now in position to see that there is, however, a crucial ambiguity in this way of putting things. For there are three levels at which subjects' practices might be said to be constitutive of a given concept: the things which fall under the extension of a concept may be practices of ours, the property in virtue of which things fall under the extension of a concept may be practices of ours, or it may be that subjects' practices are what determine the property (or relation) that determines what things fall under the extension of the concept. As I hope is now clear, the sense in which practice-dependent accounts make subjects' practices constitutive of a concept is that they claim that subjects practices (in suitably specified conditions) are what make it the case that a certain property determines the extension of the concept, irrespective of whether that property, or the things picked out by that property, are or are not practices of ours. The debate between practice-independent theorists and practice-dependent theorists is agnostic as to the question of exactly what properties our

35 Though a practice-dependent theorist, Locke was not a pluralist about personal identity. For Locke thought that continuity of consciousness was necessary for personal identity.

36 As Harold Noonan (1989, pp.57-63) points out not in so many words, so was Leibniz. Like Locke, Leibniz thought that personal identity was a forensic concept. Our moral and forensic practices determined the extension of the concept of personal identity. Leibniz differed from Locke, however, in maintaining that sameness of substance, as well as sameness of consciousness, was necessary for personal identity. In other words, although Leibniz agreed that our forensic practices determined the extension of the concept of personal identity, he differed as to the object-level question of exactly what things our forensic practices demanded be incorporated in the extension.
Chapter 1  Survival and Practice-Independence

concepts pick out. As we shall see, in the case of personal identity, as in the case of value, practice-dependent theories and practice-independent theories may agree about what properties or relations personal identity consists in, but they will nonetheless disagree deeply about what makes it the case that personal identity consists in those properties. Practice-dependent theorists will claim that personal identity consists in a certain relation, say the relation of psychological continuity, because that is the relation around which subjects organize their person-directed practices and concerns; practice-independent theorists will claim that personal identity consists in a certain relation because that is the grouping which nature (or God, or some other such independent fact) demands that we include in the extension of the concept.

Distinguishing between these two questions in conceptual analysis should help, I hope, to clear out of the way a potential confusion in the case of personal identity. Psychological continuity accounts, it might be thought, cannot be practice-independent because they claim that certain of subjects' person-directed practices—viz. beliefs, desires, values, intentions, projects... —are constitutive of personal identity. (In just the same way as it might be thought that Aristotelian virtue ethics is not a practice-independent account because it claims that good consists in our psychological states of courage or dispositions to act courageously). With the above distinction between the three levels of conceptual analysis in place, however, we can see why most psychological continuity accounts to date (like Aristotelian accounts) have nonetheless been practice-independent. True, they claim that personal identity or survival consists in properties which are themselves person-directed practices. But the story they invariably tell about what makes it the case that personal identity consists in those properties makes no essential or necessary reference to those practices. Personal identity consists in those properties, viz. psychological continuity and connectedness, not because those are the properties around which subjects are disposed to organize their person-directed practices, but because those are the properties in which personal identity independently consists; and would consist even if it were the case that subjects were not disposed to organize their person-directed practices around those facts.

Except, as earlier noted, insofar as different meta-views may naturally lend themselves to different object-level views.
2

VARIETIES OF PRACTICE-INDEPENDENCE

To date, the question of personal identity has been almost exclusively conceived as an object-level question. As we saw from our brief survey of the personal identity debate, the question of personal identity has typically been posed as the question of what are the necessary and sufficient conditions for a person identified at one time being the same person (or surviving) as a person identified at another; that is to say, as a question about what properties or relations determine the extension of the concept of personal identity. For theorists of all persuasions to date, the "fundamental question" (in Parfit's words\(^1\)) has been the object-level question of what facts determine the extension of personal identity.

Of course, as we saw in our example of middle-eastern cooking, the answer we give to this object-level question will also, and fairly straightforwardly, put us in a position to answer the question of which person-stages are stages of the same person. They will be all and only those person-stages which share the properties or relations which our answer to the object-level question tells us determine which person-stages are stages of the same person. Just as our answer to the question of what determines the extension of middle-eastern cooking (say, being cooked in the middle-east) tells us that felafels count as middle-eastern cooking and asian noodles do not (on the empirical assumption that felafels are cooked in the middle-east and asian noodles are not), so maintaining that, say, psychological continuity determines the extension of personal identity, makes it true that person-stages related by psychological continuity are stages of the same person, but person-stages related only by physical continuity are not (on the empirical assumption that two person-stages are in fact related by psychological continuity). Because our answer to the object-level question (together with some \textit{a posteriori} information) gives us an answer to the first, different answers to the object-level question will give us different answers to the first. This is how the first and second questions are related.

The fact that answering the object-level question (in tandem with some \textit{a posteriori}, empirical investigation) answers the first for us, explains why philosophers working on the question of personal identity have been so concerned to address the object-level question and, on the whole, not particularly concerned directly to address the first. (Indeed, as our earlier example of middle-eastern cooking illustrated, it is hard to see how we \textit{could} answer the first question without first answering the object-level question.) Similarly, in this thesis, I shall not be interested in the first question (at least, not directly). Rather, what I shall be interested in is the relationship between the object-level question and the meta-question. In particular, I shall be concerned to ask (and answer), what has to date been left almost entirely explicitly unasked (and explicitly unanswered): 'In virtue of what does a certain property or relation get to determine the extension of the concept of personal identity?' And, as we shall see, the sort of answer we give to this meta-question may quite dramatically affect the answer we give to the object-level question, which has so preoccupied personal identity theorists to date.

In Chapter 1, I distinguished two broad sorts of answers to the meta-question: the practice-independent answer, according to which certain properties or relations get to count as the extension-determining facts of personal identity (or good, or red, or...) because those are the facts that nature or God demand we use to determine the extension of the concept; and the practice-dependent answer, according to which these properties or relations earn their right to count as the extension-determining facts only in virtue of the fact that those are the facts around which we organize some or other of our practices. This characterization of practice-independence (and practice-dependence, for that matter) is very general. However, practice-independence in the case of personal identity, as elsewhere, may, and indeed I suspect has, taken a variety of more specific forms. In this chapter I want to try to get a little clearer about these more specific forms. For if we can better understand exactly what meta-level theoretical commitments underlie practice-independence in the case of personal identity, we can better understand, not only potentially problematic background assumptions at work in the object-level debate, but also exactly what assumptions motivate philosophers to practice-independence in the first place. (It will also give us a clearer target for the arguments of Chapter 4!)
Chapter 2 Varieties of practice-independence

Giving more precise characterizations of the various forms that practice-independence has taken in the particular case of personal identity, however, is neither as easy nor as straightforward as it really ought to be. As just mentioned, the problem is that most personal identity theorists to date have said an awful lot about what relations might determine the extension of personal identity, but very little explicitly about how and why those relations get to count as the facts of personal identity in the first place. As should be apparent, even from our brief survey, the personal identity debate to date has been almost entirely a debate about exactly what relation determines which person-stages fall under the extension of the concept of personal identity (physical continuity, psychological continuity, sameness of soul...?), with little, if anything, explicitly said about in virtue of what these relations get to determine the extension of personal identity in the first place. Cartesians, for example, are quite explicit in asserting that personal identity over time consists in sameness of an immaterial entity or soul. But they say very little about what makes it the case that personal identity consists in sameness of soul. Why does personal identity consist in sameness of soul and not some other relation among person-stages, say, physical continuity? Reductionists too, both physical and psychological continuity theorists, have insisted explicitly and at great length that personal identity consists in relations of physical continuity or psychological continuity respectively. But few have said much at all about what makes it true that personal identity necessarily consists in those relations among person-stages respectively.

This is a very striking and peculiar oversight, not only because this meta-question has loomed so large in debates in nearly all other areas of conceptual analysis, but also because, in the face of such substantial, sometimes seemingly intractable, object-level disagreement about exactly what property or relation determines the extension of personal identity, one obvious strategy would be to investigate the meta-level to see exactly what is going on there. Indeed, this is exactly what has occurred elsewhere in philosophy (in meta-ethics, for example) in response to just such object-level disagreement. For, as our earlier discussion of the debates surrounding value and colour concepts illustrated, answers to the object-level question are often, for a large part, a product of answers to the meta-question: view the facts about colour as determined independently of our responses, for example, and the natural object-level upshot is to
type-identify colours with the particular property of objects that causes them to look that colour to us; view the facts about colour as determined by our responses, and what turns out to be essential to colour concepts is not the particular property of objects that causes them to look a that colour to us, but, at best, the fact that there is some property or other of objects which causes them to look that colour to us. And some practice-dependent theorists of colour even deny that colour is to be identified with a property of objects at all, rather it is a (dispositional) property of observers. So exactly what meta-theory, or meta-theories, have been influencing object-level analyses in the case of personal identity?

2.1 Old-style Cartesianism: The one-phase *a priori* analysis.

Let’s begin by reconstructing the Cartesian meta-theory. As we have seen, Cartesians claim that personal identity consists in a further fact, namely, sameness of an immaterial entity or soul. This is the relation which, according to Cartesians, determines the extension of the concept of personal identity. But how and why, for Cartesians, does sameness of soul necessarily determine the extension of personal identity? In particular, what, if any, kind of role do our person-directed practices play in determining that sameness of soul determines the extension of personal identity?

Old-style Cartesians, such as Reid and Butler, typically believed (some, such as Chisholm, still apparently do) that first-person experience, in particular introspection, gives us direct and privileged access not merely to the meaning or semantics of personal identity, but to the very nature or underlying metaphysics of personal identity itself. There are two more particular sorts of views that old-style Cartesians take with respect to the sort of introspective access that we have to the nature of personal identity. One of these is to view our introspective access to the soul as quasi-causal or perceptual. In introspection, we perceive—we are literally in causal contact with—the soul, just as in perceiving a tree we are literally in causal contact with the tree. The second view regards introspection as giving us non-causal or logical access to the soul—as if, in introspection, we are peering into Plato’s heaven. On this second view, claims about personal identity are on par with claims about other platonic entities, such as mathematical entities. In the abstract this methodology may seem obscure, so a geometric analogy may help. Imagine holding a picture of a cube before our mind’s eye. We can imagine that if we hold it
there for long enough, bit by bit, side by side, angle by angle, the whole nature of the cube will be revealed to us. When we first begin to hold the cube before our mind’s eye, for example, we may know only that it is a three-dimensional square. But if we hold it in our mind’s eye for long enough, examining it from different angles and different perspective’s, eventually it will become transparent that the three-dimensional square has six sides. And, we might think, if we keep examining the cube like that for long enough, everything there is to know about the nature of the cube will be revealed to us. So, for old-style Cartesians who take the second view, it is with personal identity.

But introspection, on either view, although certainly a person-directed practice, does not determine what determines the nature of personal identity, any more than examining a cube before our mind’s eye determines the nature of a cube. Rather, it is our best means of obtaining access to the (independently determined) nature of personal identity. It is through introspection that we are best, most reliably and most vividly aware of ourselves as a special sort of non-reducible entity (a Cartesian ego or soul) continuing through time.

Later Cartesians, such as Richard Swinburne, have invoked additional arguments of a different sort in favour of the Cartesian view. According to Swinburne, more recent personal identity puzzle cases, such as those of reduplication and fission, and cases which seem to allow (on reductionist views) that personal identity may be indeterminate, pose problems for reductionist or, as Swinburne calls them, ‘empiricist’ views of personal identity. In response to these cases, Swinburne claims, reductionists have to reject as misguided some very strong intuitions we seem to have about personal identity. The fact that reductionist accounts face these problems and are forced to deny some of our deeply held beliefs about personal identity, Swinburne claims, gives us good reason for thinking that these reductionist accounts are wrong. And the fact that the non-reductionist view can deal with these problems in a way which saves our intuitions, is an argument for thinking that the non-reductionist view of personal identity is right.
We will discuss the details of some of these cases later\(^2\), but for now the point to note is the more schematic one about the way in which personal identity theorists have gone about answering our second question. And, in this context, Swinburne's 'argument by default' in favour of the Cartesian view on the basis of intuitions gleaned from puzzle cases, not simply introspection alone, seems to represent a departure from the one-phase \textit{a priori} method of theorizing about personal identity (i.e. of theorizing about the answer to the second question) which typified old-style Cartesianism.\(^3\)

2.2 The role of the method of cases.

Few philosophers nowadays, including non-reductionists like Swinburne, seem to believe that introspection alone is sufficient to reveal the whole nature of personal identity. Contemporary theorizing about personal identity, as elsewhere, has been dominated by what we might call the 'method of cases'.\(^4\) Just as old-style Cartesians believed that introspection was our best means of accessing the nature of personal identity, so contemporary (and some, like Locke, not so contemporary) personal identity theorists believe that plumbing our intuitive responses to actual and hypothetical puzzle cases involving questions of personal identity is the best means of accessing or elucidating the meaning and nature of personal identity. To this end, theorists formulate and present us with a vast range of actual and hypothetical personal identity puzzle cases: people are destroyed, recreated, duplicated and transported; bodies are dismembered, remembered and recombined; brains are thawed, unthawed, disintegrated and reintegrated; psychologies are removed, switched and halved, and so on. The right account of personal identity (i.e. the right answer to the second, object-level question of what relation determines the extension of personal identity), it is assumed, will be the

\(^2\)For an excellent summary of these puzzle cases and responses to them, see Noonan (1989), Chapter 1, esp. pp.14-23.

\(^3\)Of course, it may not represent such a radical departure. For one possible, though I think pretty implausible, way of understanding what is going on in the method of cases (shortly to be discussed) is that the method of cases is itself a method of introspecting, or at any rate, of activating our introspective access to the nature of personal identity. On this view, the method of cases, by prompting our introspective faculties, would give us similarly privileged access to the nature of personal identity. Whether Swinburne himself takes this view is hard to tell.

\(^4\)I borrow this phrase from Johnston (1987).
one which best predicts and explains our intuitive reactions (our beliefs about who is who) to these puzzle cases.

But exactly what function are our intuitions about personal identity supposed to be playing in our analysis of personal identity? In particular, what relation do they bear to the metaphysics of personal identity? Are they tracking or discovering the (independently determined) nature of personal identity in the same way in which, for example, scientists use an electron microscope to discover the (independently determined) facts about atoms? Are our intuitions actually determining which relation is the relation of personal identity? Are they serving to fix on a natural kind? Are they merely helping to make transparent our ordinary beliefs about personal identity by revealing when we are disposed to apply the concept and when we are not? Are they doing more than one of these things? Or are they doing something else entirely?

Well, as I have stressed, strangely, given the incredible dominance of this method of theorizing about personal identity, very few personal identity theorists have asked or answered these questions. And thought experiments could be playing any one and sometimes more of these functions. So it’s hard to be sure of exactly what is going on in theorizing about personal identity. Rather than second-guess which theorists subscribe to exactly which view of the function of the method of cases on the basis of scanty evidence, then, I shall simply canvass a few of the more plausible options and, where possible, attempt to put a concrete name to an option. Let’s begin with what would seem to be the most likely option.

2.2.1 Option 1: Elucidating ordinary beliefs about personal identity

I suspect that for most contemporary theorists of personal identity (as, indeed, for most theorists who employ the method of cases in general) the method of cases is serving as a means to make transparent and consistent—to reach reflective equilibrium regarding—our ordinary, implicit grasp of the concept of personal identity. Puzzle cases serve primarily to elicit information about under exactly what circumstances we are disposed to believe that a later person is the same person as an earlier one.⁵ We ask respondents, of any given case, whether or not they think

---

⁵Puzzle cases also often test for another belief closely related to our beliefs about survival, namely, our beliefs about how it is rational for the people involved in the puzzle cases to behave, given the beliefs they are described as holding. These beliefs are,
that a later person is the same person as an earlier one, and then we ask: exactly what relations among person-stages explain these beliefs about personal identity? On this view of the role of the method of cases, then, the project of analysing personal identity is, for the most part, the project of finding out exactly what relation(s) among person-stages underlie and explain our beliefs about who is who in puzzle cases. And the right account of the personal identity (that is, the right account of the relations which determine the extension of personal identity) is the one which does the best job of explaining and predicting these intuitions about who is who in the many and varied puzzle cases.

If this is the role that the method of cases is playing in theorizing about personal identity, then it goes some way to explaining why personal identity theorists have, on the whole, not thought to ask or address the meta-question. For the project of theorizing about personal identity just amounts to the project of figuring out and rendering explicit exactly what implicit object-level commitments underlie, guide and explain our intuitions. The question of personal identity, on this view, is just the question of what is the relation among person-stages in virtue of which we are disposed to believe that someone is or is not the same person. The further question, 'in virtue of what do these relations underpin our beliefs about personal identity?' is no part of this project, and so simply not a question that naturally arises. Briefly considering the sorts of the puzzle cases typically used to elucidate the meaning of personal identity (where that, on this view, means making our pre-theoretic object-level commitments explicit) may help to illustrate.

As already mentioned, the puzzle cases which have dominated contemporary theorizing about personal identity have typically involved pulling apart the various relations among person-stages which usually go together and asking whether or not a person survives such transformations, or, in cases where more than one person is involved, who is who. For we want to find out exactly which relation(s) among person-stages our beliefs about personal identity are tracking. So, for example, we are asked to imagine that the brain of a person, Mr Brown, is successfully transplanted in the skull of Mr Robinson. The resulting

in Parfit's famous phrase, beliefs about 'what matters in personal identity', and Chapter 3 discusses these sorts of beliefs and the role they have played in theorizing about personal identity in greater detail.
person, call him Brownson, wakes up after the operation with Mr Robinson’s body, but is, in all psychological respects, indistinguishable from Brown. Who is Brownson? Is Brownson the same person as Brown? Or is Brownson Robinson? Suppose we answered that Brownson is Robinson. Then it would appear that we believe that having the same body is required in order to be the same person. Suppose, on the other hand, that we answer that Brownson is the same person as Brown. Then it would appear that we do not require that a person have the same body in order to be the same person. At most, as regards physical continuity, we would seem to require that a person have the same brain. Or, for another typical example, we are asked to imagine that there is a machine which is capable of copying all the psychological information (including dispositional information) from the brain of one person, call her Susan, and transferring it into the brain of another, call her Carol. So that Carol wakes up from the procedure with all of what was Susan’s dispositional psychology. Who is Carol? Is Carol Susan? If our intuition is that Carol is Susan, we would seem to think that psychological continuity is more important for personal identity than physical continuity. A person can be the same person when they have the same psychology, irrespective of whether or not they have the same body. Having the same body is not necessary for personal identity. Suppose that we think that Carol is not Susan. Then we would seem to be committed to the view that psychological continuity, in the absence of physical continuity, does not secure personal identity. Physical continuity would then seem to be necessary for personal identity. Or, to give another standard sort of example, you are asked to suppose that a machine will record all the information about your bodily states and psychology, after which your brain and body will be destroyed. Simultaneously, the machine will use the atomic blueprint to make a qualitatively identical body and brain out of different matter. Will you survive this? Is the person after the procedure the same person as the person before? If we answer ‘yes’, it appears that we do not regard physical continuity as necessary for personal identity. Psychological continuity, even in the absence of physical continuity, is sufficient for personal identity. If we answer ‘no’, it appears that we regard physical continuity as necessary for personal identity. Or, for another typical sort of variation on the above puzzle case, suppose that the machine in the prior example did not make a qualitatively identical person from your blueprint straightaway, but
waited five years to construct the qualitatively identical person? Would the qualitatively identical person still (if it ever was) be you?

Our intuitive responses to these puzzle cases and a myriad of others like them, it is assumed, individually and collectively reveal exactly what relation between person-stages underpins and explains our beliefs about who is who in puzzle cases or, as I have couched matters, our answer to the second, object-level question of what relation determines which person-stages are stages of the same person. But they do not tell us anything about the sort of answer we would give to third, meta-question of in virtue of what does the relation among person-stages, whatever it is, (revealed by our intuitive responses to these puzzle cases) get to count as the relation of personal identity. The method of possible cases, as typically deployed, gives us information about what we use the words 'survival' or 'personal identity' for, but it does not tell us what governs what we use the words for. In order to plumb our intuitions about the meta-issue, we need to present people with puzzle cases of a different sort: puzzle cases which ask people about their own or other people's intuitive reactions about who is who in puzzle cases—puzzle cases, for example, which take conflicting opinions whether or not someone is the same person in a given puzzle case and ask whether or not these disagreements are disagreements about personal identity. If respondents answer 'yes' to this question, the obvious further question is 'in virtue of what are these people talking about personal identity?' If they answer 'no', the question is 'why are these disagreements not disagreements about personal identity?' In Chapter 4 I will put some flesh on these as yet schematic bones. But the point to note for now is that this view of the role of the method of cases in theorizing about personal identity—as elucidating our pre-theoretic, object-level commitments—is, in itself, neither obviously practice-dependent nor obviously practice-independent, for it does not explicitly address the meta-issue one way or another.

However, it might be, not simply and incidentally that the meta-question has not arisen for theorists employing the method of cases in theorizing about personal identity, but that these theorists actively assume that there is no such further question to be asked. The method of cases serves as a means to find out what we mean by personal identity (i.e. what relation among person-stages underpins our personal identity talk), and there is no further fact of the matter (except of the anthropological or
sociological kind of finding out how we came to mean what we mean by the words ‘personal identity’ or ‘is the same person’). This view of the further, third or meta-question as simply the anthropological-cum-sociological question of why we happen to use the words we do to refer to things, is the trivial reading of practice-dependence that we noted in Chapter 1. But, if theorists of personal identity believe that this trivial question is the only further question about personal identity there is to be asked, this amounts to a substantive thesis. For, according to these theorists, ‘personal identity’ would be like ‘triangle’. A triangle, we might say, is a geometric object with three angles and straight sides. But, as we earlier noted, it seems a fair response to say to someone who asks the further question ‘but in virtue of what does that property—the property of having three angles and straight sides—determine which objects are triangles?’; the answer is that’s just what we mean by triangle. Plausibly, there is no further question to be asked or answered about the meaning of ‘triangle’, except the anthropological-cum-sociological question of how we came to use the word ‘triangle’ to refer to geometric objects with the property of having three angles and straight sides.

This may well be what some personal identity theorists think about personal identity (and if they do, the arguments of Chapter 4, in particular, should, I hope, give them reason to think again). But I suspect that, in fact, many theorists employing the method of cases do not take this ‘no-further-question’ view. The structure of the personal identity dialectic, and the function the method of cases has served therein, is the subject of Chapter 9. But let me briefly pre-empt some of this later discussion, for it is relevant here. People’s intuitions gleaned from actual and hypothetical puzzle cases have proved very problematic for personal identity theorists. The problem has been that people’s intuitive responses to puzzle cases do not seem to converge on any one of the competing views about our object-level commitments in personal identity mentioned at the outset; and, what is particularly problematic, they seem not to converge, even when all is known about person-stages and their inter-connections and respondents have well-informed, consistent and considered opinions about who is who in the various puzzle cases. Of course, if we were all Cartesians, we might well have grounds to complain that the full information is not given when all is known about person-stages and their inter-connections. For a crucial detail, what happens to the soul, is not settled when we are told everything about the relations
which hold between person-stages in puzzle cases. But unless philosophers, contrary to their professed views, are closet Cartesians, the lack of convergence cannot be explained away as the product of ignorance, error or lack of full information. Philosophers such as Shoemaker, Parfit, Lewis, Noonan and a host of others on the one hand, and Williams, Wiggins, Nagel, Unger and a host of others, on the other, all have exceptionally informed and considered opinions about who is the same person as who in puzzle cases. But their informed and considered opinions do not converge. The former insist that physical continuity is no necessary condition for personal identity (psychological continuity is necessary and sufficient); whilst the later insist that physical continuity is necessary for the preservation of personal identity.

Now, given this situation, if theorists really held the 'no-further-question' view of personal identity, what they would say (or what they ought to say, at any rate) about this sort of entrenched object-level disagreement, is that what it shows is that the meaning of 'personal identity' is just vague or indeterminate, and that's all there is to say about personal identity. But this is not at all what most theorists of personal identity employing the method of cases say in response to such interpersonal object-level divergence. They go on to formulate and reformulate more and more elaborate puzzle cases, the intuitive reactions to which, they claim, support their preferred account; and engage in more and increasingly desperate attempts to explain away as mistaken or irrational intuitive reactions which would seem to support some other competing account. The feeling seems to be that if we could just get the presentation of puzzle cases quite right, then people's intuitive responses will converge, and we will have finally discovered the answer to the question of personal identity.

However, if the no-further-fact view does not explain the desperation with which philosophers have sought intuitive convergence on the question of personal identity, what does? One explanation that springs readily to mind is that, although the method of cases has replaced introspection as the chief vehicle for theorizing about personal identity, and although few theorists still believe in Cartesian egos or souls, one element of Cartesianism—what we might call the 'real-essence' view—has remained. The idea that there is a real essence or nature of personal identity determined quite independently of whatever we may happen to
think about things has carried over from old-style Cartesian thinking; and
the method of cases has simply replaced introspection as the means best
to access or discover those facts. On this view, intuitions gleaned from
puzzle cases play much the same role as an electron microscope plays for
physicists: they serve as our best means of discovering or gaining access to
the independently determined facts of the matter. Just as atoms exist out
in the world waiting to be discovered and revealed by the electron
microscope, so the facts of personal identity exist out there in the world
waiting to uncovered via introspection and/or intuition. And just as
scientists persisted in refining microscope technology until they
discovered atoms, so personal identity theorists persist in refining thought
experiments in the hope that, once they are sufficiently fine-tuned, they
too will reveal the nature and structure of personal identity.6

Admittedly, that theorists assume that there is a privileged
independent fact of personal identity waiting to be discovered like a
fourth fundamental force is not the only explanation for the desperation
with which philosophers have sought intuitive convergence on the
question of personal identity. They might, for example, believe that
convergence is important, not because convergence will show that we
have finally discovered the real nature or essence of personal identity, but
because such convergence around a single relation is somehow
constitutive of a relation counting as the relation of personal identity. Or,
and perhaps relatedly, they might believe that there is a deep, but
presently obscured fact of personal identity implicit in our practice upon
which we would converge were we in better conditions. This sort of
meta-view is not practice-independent, it is rather a sort of subjective
universalism, but it would nonetheless explain the desperation with
which theorists of personal identity have sought intuitive convergence.
For, on this sort of view, in the absence of convergence, we have no choice
but to endorse an error theory about personal identity; and that an error
theory should be true for the case of personal identity, Hume aside, may
seem intuitively incredible. Or the explanation might have nothing to do
with meta-level views or ontological commitments as such. It might just
be that theorists think that personal identity is semantically determinate
and/or that it is intuitively incredible to suppose that personal identity

---

6If this is the right reconstruction of what is going on at the meta-level for at least some
theorists of personal identity—if personal identity is such an independently determined
fact—then there is a real question as to how personal identity influences our intuitions.
might turn out, like baldness, to be semantically indeterminate. Or the desire for convergence might just be a brute psychological disposition of ours. Although the personal identity dialectic does indicate that few involved in it hold the no-further-fact view, then, it does not decisively show that theorists are committed to the ‘real-essence’ view of personal identity as consisting in a wholly independently-constituted entity or relation best accessed by the method of cases, in just the same way as old-style Cartesians thought personal identity to consist in an independent entity best accessed via introspection. Nonetheless, this ‘real-essence’ view is one sort of practice-independent view and a perfectly possible meta-view for a theorist employing the method of cases to have. And, especially given the Cartesian origins of much thought about personal identity, it is not wildly implausible to suppose that some theorists actually hold it.

However, at least one personal identity theorist who makes similarly heavy use of the method of cases takes a different, and undeniably practice-independent, view of personal identity. And given that, in a symposium on Peter Unger’s book, philosophers such as Parfit, Swinburne and Shoemaker never passed comment on his explicitly practice-independent approach (choosing rather to debate his object-level claims), it would not be particularly surprising to discover that this sort of practice-independent meta-assumption has underlain much of the contemporary theorising about personal identity. On this view, and in contrast to the ‘real-essence’ view, intuitions gleaned from puzzle cases are not taken to discover or reveal the independently determined metaphysical facts about personal identity, rather they actually determine, or fix, the reference of ‘personal identity’. This view may initially sound like a practice-dependent view but, as we shall now see, in at least one popular incarnation (and the one favoured by Peter Unger), it is thoroughly practice-independent.

2.2.2 Option 2: Practice-independence as rigidification

It may help to get clear about this third sort of approach to theorizing about personal identity by first considering some better known analogues from elsewhere in philosophy. Consider, for example, analytic functionalist accounts of the nature of mental states.
Analytic functionalism is the doctrine that the type-identity conditions for mental states are given in terms of their actual causal interactions, and dispositions to enter into causal relations, with sensory inputs, behavioural outputs, and other mental states. Mental state-types are defined by their characteristic causal or functional role in relation to inputs, outputs, and other mental states. So, very crudely put, pain is, *inter alia*, the state-type such as typically to be disposed to utter 'ouch', and to believe that you are in pain, when pricked by a needle. Less crudely (though still over simplistically), to give the Lewis-Ramsey sentence for pain:

\[ X \text{ is in pain} = \text{def} \text{there are two states (properties) the first of which is caused by skin damage and causes, or is disposed to cause, both the emission of "ouch" and the second state, and the second state causes, or is disposed to cause, brow wrinkling, and } X \text{ is in the first state.} \]

Analytic functionalists further draw a distinction between the functional role or software definitive of some mental state-type, and the occupant or hardware which realizes that functional role, usually (but not necessarily) some physical property. A state or property will *realize* any given functional state when it plays the functional or causal role definitive of that functional state. Of this much, all analytic functionalists are agreed. But here is where agreement ends.

Some functionalists, call them (following, amongst others, David Braddon-Mitchell and Frank Jackson\(^7\)) 'scientific' or 'empirical' functionalists, maintain that each mental state-type name is definable in terms of a specific causal role, and that name necessarily denotes whatever property happens to occupy that role in the actual world. Mental states are type-identified with the neurophysiological state which actually plays the functional role definitive of that state-type for us. Suppose that, *a posteriori*, (say) C-fibres firing is the brain-state that plays the functional role definitive of pain for human beings in the actual world, then pain is necessarily identical to C-fibres firing in all worlds; and so on, for all other mental states. 'Pain' rigidly designates whatever actually realizes the pain role in humans; and, given that C-fibres firing is what actually realizes the pain role in humans, pain is necessarily C-fibres-
firing. On this view, no C-fibres firing, no pain. It follows from scientific or empirical functionalism, then, that beings who are differently physically constituted from us—dogs, dolphins, silicon-based Martians and computers—have no mental states. For this reason, scientific or empirical functionalist accounts have often been accused of chauvinism—of denying mental states to other differently physically constituted beings, no matter how apparently ‘intelligent’ their behaviour.

Scientific or empirical functionalists identify pain with the first-order property that actually plays the pain role in us. In contrast, for ‘second-order’ functionalists mental state-types are to be identified with functional role or software, not actual realizer or hardware. ‘Pain’, for example, does not rigidly designate the brain-state that actually realizes the pain role in humans, namely C-fibres firing; ‘pain’ refers rather to a (second-order) property that the brain-state C-fibres firing has when it realizes the pain role: namely, being a typical realizer of the pain role. Pain, on this second-order view is not the brain-state C-fibres firing, but rather being in some brain-state which typically plays the pain role. Since some property other than C-fibres firing might equally well have the second-order property—it might equally well be a state which typically plays the pain role—different properties can realize the pain role. To use some standard functionalist jargon, on the second-order view, mental states such as pain are multiply realizable. For different brain-states might have the property of being typical realizers of the pain role in different subjects; and those states would count as being in pain just as well.

Sometimes it is thought that these two views are the only two options open to functionalist accounts of mental states: either functionalists are first-order and chauvinistic or they are second-order and mental states are multiply realizable. If they are first-order then they meet the pretty powerful objection of chauvinism; if they are second-order, they are forced to deny a common and central intuition that there is something very important to pain about the hardware that realizes the pain role. For it is the brain state-type, not the second-order property of being in a state-type which typically plays the pain role, which causes pain behaviour. However, there is a third-option; a sort of half-way house between the first- and second-order views. This option is to be a first-order functionalist—that is, to identify mental state-types with first-order properties or hardware, not role—but to maintain, nonetheless, that
functional roles are multiply realisable. How can this be? The claim is that, although pain is identical with the hardware or brain state-type, whatever it happens to be, that realizes the pain role in subjects, it is only *contingently* identical with that brain state-type. For us, *homo sapiens*, say, C-fibres firing is the hardware that realizes the pain role, but go to a world of Martians in which some other property, say, M-fibres firing, realizes the pain role and that property counts as pain round there just as well. Pain is, at best, only *contingently* identical to C-fibres firing. Pain, on this view, although a first-order property, may be different first-order properties in different worlds (much as the colour of the sky is different colours in different worlds). For different first-order properties may play the pain role in different subjects and those properties count as pain just as well. As Braddon-Mitchell and Jackson\textsuperscript{8} state the view,

Premiss 1. Pain = occupant of causal role R (Accepted conceptual fact)
Premiss 2. Occupant of causal role R = brain state B (Empirical discovery)

Conclusion. Pain = brain state B (Transitivity of identity)

I mention this third option briefly here, for, I think, something very similar to this in essential structure is the right analysis of personal identity.

For another example altogether, consider the famous (or perhaps better, infamous!) debate over the nature of our concept of water. In an influential article ‘The meaning of ‘meaning’\textsuperscript{9}, Hilary Putnam presents the following science-fiction case. Suppose that somewhere in the Universe there is a planet pretty much exactly like Earth, call it Twin Earth. One of the very few differences between Earth and Twin Earth is that the liquid people on Twin Earth call ‘water’ is not \textit{H}_2\textit{O}, like it is on Earth, but a different liquid, whose chemical configuration we might abbreviate as XYZ. XYZ is indistinguishable from water at normal temperatures and pressures; it tastes like water and quenches thirst; it fills the oceans, lakes and rivers, falls from the sky as rain, and so on. In short, XYZ does all the same things on Twin Earth, as \textit{H}_2\textit{O} does on Earth. Now,

\textsuperscript{8}\textit{ibid.}, Ch. 6. For other statements of the view see David Armstrong (1968) and David Lewis (1983).
\textsuperscript{9}Putnam (1975), pp. 215-71.
do the people on Earth and Twin Earth mean the same thing when they utter 'water'? In particular, what settles the reference of the term 'water'?

What we might call the 'rigidifying view' (the analogue of the scientific or empirical functionalist view for the case of water\textsuperscript{10}) has it that water is the stuff which \textit{actually} fills the oceans, falls from the sky, runs from taps, gets called 'water' by experts, is essential for life on earth, is odourless and colourless, etc.; or it is the stuff which \textit{actually} does most of the foregoing; or it is the stuff which \textit{actually} comes close to doing most of the foregoing (perhaps it isn't completely colourless, for instance); or it is the stuff which \textit{actually} comes close to doing most of the important parts of the foregoing, and in any case comes closer than anything else around; or... On this view, the word 'water' functions like an indexical to rigidly designate the stuff it picks out in the actual world. Water, in all worlds, is what realizes the water role in the actual world. \textit{A posteriori}, what realizes the water role in the actual world is H\textsubscript{2}O. And so the rigidifiers' conceptual claim (that 'water' is a rigid designator), combined with some empirical information (that what realizes the water role in the actual world is H\textsubscript{2}O), delivers the claim that 'water=H\textsubscript{2}O' expresses a necessary truth.\textsuperscript{11} Since 'water' rigidly designates whatever actually realizes the water role, and since (\textit{a posteriori}) what actually realizes the water role is H\textsubscript{2}O, water (by transitivity of identity) is necessarily H\textsubscript{2}O. In all worlds, if there is any water there is H\textsubscript{2}O, and if there is H\textsubscript{2}O there is water. And this remains true, of course, even in a world such as Twin Earth where XYZ, not H\textsubscript{2}O, is the chemical configuration that plays the water role. Just as, for scientific functionalists, no C-fibres firing, no pain, so, on the rigidifying view of 'water' (combined with information about what actually realizes the water role), no H\textsubscript{2}O, no water. So that a person on Twin Earth who uttered the word 'water' to refer to XYZ would mean something different from what we, here on Earth, mean by it. They

\textsuperscript{10}One could be a scientific functionalist without being a rigidifier i.e. one could hold that pain is the brain-state that actually plays the pain role without maintaining that 'pain' is a rigid designator. Scientific functionalists could think that it is \textit{a priori} true that pain is the brain-state that actually plays the pain role, for example. As a matter of fact, however, scientific functionalists are typically also rigidifiers.

\textsuperscript{11}This is why those who regard 'water' as a rigid designator typically also believe that 'Water=H\textsubscript{2}O' expresses a necessary, rather than a contingent, truth. But the two questions—whether 'water' is a rigid designator and whether 'Water=H\textsubscript{2}O expresses a necessary truth—are different. You could believe, for example, that 'Water=H\textsubscript{2}O' expresses a necessary truth without believing that 'water' is a rigid designator: it just so happens that the stuff which plays the water role in every possible world is H\textsubscript{2}O.
would be talking about water*. According to this view, then, 'water' rigidly designates, not the role that water plays, viz. 'being the stuff (whatever it is) that has the property of falling from the sky, running from taps...etc...', but the stuff which actually realizes that role here on Earth. Since what actually realizes the water role here on earth is H2O, water is necessarily H2O; just as, on scientific functionalist accounts, since 'pain' rigidly designates the brain-state that actually plays the pain role in us, and (a posteriori) the brain-state which actually plays the pain role in us is (say) C-fibres-firing, pain is necessarily C-fibres firing.

The rigidifying view, which is the analogue of the view about personal identity shortly to be discussed, contrasts with a view which often goes under the name of the 'definite description' view. Whereas the rigidifying view rigidly identifies 'water' with its actual content (namely, H2O), this view identifies the meaning of the word 'water' with its character or role viz. being the stuff that falls from the sky, runs from taps, fills the oceans...etc...12 According to the definite description view, then, the reference of 'water' is fixed by the role water plays, conceived as the definite description: 'the stuff which fills the oceans, falls from the sky, runs from taps, gets called 'water' by experts, is essential for life on earth, is odourless and colourless, etc.; or it is the stuff which does most of the foregoing; or it is the stuff which comes close to doing most of the foregoing (perhaps it isn’t completely colourless, for instance); or it is the stuff which comes close to doing most of the important parts of the foregoing, and in any case comes closer than anything else around; or...

Both the rigidifying and the definite description views agree that being H2O is essential to actual water.13 However, they disagree about whether 'water' is a rigid designator (i.e. refers in all worlds to the stuff which actually plays the water role on Earth) and whether 'water=H2O' is necessary or contingent. The rigidifier maintains that 'water' is a rigid

---

12On another way of putting the point, whereas what I have called the 'rigidifying' view has it that 'water' rigidly designates H2O ('water' picks out H2O in all worlds), the definite description view has it that 'water' rigidly designates the water role given by the definite description. So, in this sense, the definite description view is also a rigidifying view. The difference between the two views lies in what they rigidify on—actual realizer for what I have called 'rigidifying' theories; role or description for what I have called the 'definite description' view. So in calling the first view a 'rigidifying' view I do not mean to imply that the second is not. The crucial difference between the two views lies in the different things that they rigidify on: actual realizer or second-order property (being a chemical configuration that has the property of satisfying the definite description).

13That is, they agree that it is essential de re that water is H2O, they disagree about whether this is also a de dicto necessity.
designator and, typically, that 'water=H2O' is necessary. The definite description view disagrees on both counts. 'Water' does not rigidly designate the actual realizer of the water role, it designates the role itself: namely, being the stuff which falls from the sky, runs from taps... Any stuff (H2O, XYZ...) which falls from the sky, runs from taps etc is properly called 'water'. On the assumption that there are possible worlds (such as Twin Earth) in which some chemical configuration other than H2O (such as XYZ) satisfies the description, then, in such a world, that other chemical configuration counts as water just as well. Earth people and Twin Earth people, then, mean the same thing by 'water'. For water is defined by its role, and XYZ plays the same role on Twin Earth as H2O plays on Earth: both H2O and XYZ have the (second-order) property of playing the water role in their respective worlds. This definite description view of water is analogous to the second-order functionalist account of mental states.

The third option—contingently identifying water with the first-order realizer of the water role—is also available here (whether it is right for the case of water I leave an open question). On this view, water is whatever first-order property it is that plays the water role—whatever stuff it is that falls from the sky, runs from taps, fills the oceans...etc... Earth people and Twin Earth people mean the same thing by 'water'—they mean 'whatever stuff plays the water role'—it's just that, as a contingent matter of fact, different chemical configurations happen to realize the water role on Earth and Twin earth. Here on Earth, the water role happens to be realized by H2O, but go to another world—say, Twin Earth—where people use the word 'water' to refer to another property, XYZ, and that property counts as water just as well. For that stuff plays the water role on Twin Earth. Water, for us on earth, although identical with H2O, is only contingently identical with H2O. For go to Twin Earth, and XYZ counts as water round there. On this view, water on Twin Earth is (contingently) XYZ; water on Earth is (contingently) H2O.

What unites the two latter views, as against the rigidifying view, is their common denial of the related claims that 'water' rigidly designates the first-order property that actually realizes the water role on Earth, and that 'water=H2O' expresses a necessary truth. According to both the 'second-order' view and the 'contingent-identity' view14 (the view that

14I will later call this view 'pluralism'.
water in any world is contingently identical with the first-order property that realizes the water role in that world), in order for some stuff to count as water in any world it must play the water role in that world. Any stuff which does not play that role (fall from the sky, run from taps etc) is not water. What we use the word 'water' to refer to, these two will both have it, is not the particular chemical configuration that happens actually to satisfy the description, but the functional role (or the first-order property that realizes that role) that that chemical configuration plays. (In much the same way as 'doorstops' refers not to the messy bunch of actual realizers—books, irons, drunks, chairs and a whole host of other reasonably heavy, solid, extended objects—but rather, to the common role propping doors open that this messy disjunctive group of objects play.) Since we use the word 'water' to refer to role (or first-order property that happens to realize that role), not actual realizer, we should take the reference of 'water' to be fixed by that role, and leave it open that various other chemical configurations which played that role might count as water just as well.\footnote{I suspect that, in the case of 'water' (thought not in the case of 'survival') we probably do not use the words in this way, at least not unambiguously.}

However, despite their common denial of both of the rigidifier's claims, the 'second-order' and 'contingent identity' view differ amongst themselves on the question of whether water should be identified with that role or whether, rather, it should be (contingently) identified with the first-order property \((\text{H}_2\text{O} \text{ on Earth}, \text{XYZ} \text{ on Twin Earth})\) that happens to realize that role. As with the case of mental states, the question hinges on whether it is a sufficiently central intuition that water is causally efficacious. If it is—if, for example, we think things of the sort 'It was because I left my bike in the rain that it rusted'—the first-order view will be right. For it is the first-order property \((\text{H}_2\text{O}, \text{in this case})\) which caused my bike to rust, not the second-order property of being the stuff which falls from the sky, fills the oceans etc. For the second-order property is causally inert. If, on the other hand, we do not say things of the above sort—if we do not seem to require that water be causally efficacious—the second-order view may be right. The 'contingent identity' view, then, agrees with the second-order view that what unites different chemical configurations under the concept of water is the common role that these chemical configurations play, and that, as a consequence, the water role is
multiply realizable. But it denies that water should be identified with that role. Rather, water should be (contingently) identified with the first-order property which happens to realize the water role for people in a given world.

If it sounds as if the 'contingent identity' view has to tread a subtle line between the rigidifying view (in similarly identifying water with a first-order property) and the second-order view (in similarly maintaining that it is the role or character of water in virtue of which different first-order properties earn their right to count as water), that is absolutely right. That it is not only possible, but plausible, to tread such a subtle line in the case of personal identity is the subject matter of chapters 5 and 6.

2.3 'Survival' as a rigid designator

Peter Unger is an example of a contemporary personal identity theorist who seems to hold the rigidifying view for the case of personal identity or survival. Unger believes that the reference of 'personal identity' or 'survival' is fixed by a certain (limited) class of our person-directed practices: namely our 'deepest beliefs' about our survival conditions. For Unger, the character of personal identity (that the conditions of our survival are whatever the conditions which jibe with our deepest beliefs about our survival conditions are) is given a priori, but the content or referent of personal identity is revealed a posteriori by noting people's intuitive reactions to puzzle cases, properly presented. For Unger, then, the method of cases is a posteriori not a priori theorizing. He writes,

By disclosing our deep beliefs about our survival, I indirectly articulate what are, as we most deeply believe, some quite general conditions of

---

16 In fact, I suspect that theorizing about the meaning of terms by way of the method of cases is invariably a posteriori, rather than a priori, philosophizing. All contemporary philosophers of personal identity seem to believe that it's an a priori truth that the right account of the relations in which personal identity necessarily consists is the one which best accords with and explains people's intuitive response to puzzle cases. The difference is that, unlike Unger, they appear to think that what the relations that jibe with people's intuitive reactions to puzzle cases are, is also revealed a priori. I think that this rests on a mistaken conception of what it is to do a priori as opposed to a posteriori philosophizing. But it's easy to see why the mistake might arise. Consider this: Are anthropologists, who investigate what other communities mean by certain of their words, doing a priori or a posteriori work? Well, what people mean by words is certainly ana priori matter. But the work that the anthropologists are doing is in observing what those communities mean by certain words is surely a posteriori, empirical research. What the anthropologists are doing is a posteriori investigation into an a priori matter. But, isn't that exactly analogous to what philosophers do when they note and record people's intuitive reactions to puzzle cases?
Chapter 2

Varieties of practice-independence

our survival. How do I detect these deep beliefs? Doing informal psychology, I uncover them by noting people’s responses to examples.17

For Unger, then, theorizing about personal identity proceeds in two stages. First, we have an *a priori* conceptual claim, roughly: our survival conditions are those conditions that we most deeply believe our survival to require. This fixes the reference of ‘personal identity’ or ‘survival’. Second, there is the *a posteriori* task of finding out what the conditions that we most deeply believe our survival requires actually are. The first claim is Unger’s meta-story: his story about in virtue of what some relation gets to count as the relation of personal identity. The second stage is his object-level claim: for Unger, an empirical claim about what relation actually underpins our ‘deepest beliefs’. For Unger, this second stage goes by way of the method of cases: by way of actual and hypothetical puzzle cases, properly presented18, we elicit our deepest beliefs about what are our survival conditions. And then, by transitivity of identity, those are our survival conditions.

Unsurprisingly, given this, Unger’s chief argument against both Cartesian and psychological continuity theories is that, as our intuitive reactions to puzzle cases illustrate, these theories do not cohere with our deepest beliefs about our survival conditions. The object-level story Unger takes this meta-story to deliver up is a version of a physical continuity account of survival:

Briefly and roughly, as part of his total (dispositional) psychology, each of us has a core, or a basic, psychology that, as a matter of fact, he has in common with all normal, and most subnormal, human beings. Then, also briefly and roughly, one of us will survive from an earlier to a later time when whatever physically realizes that person’s core psychology continuously (enough) realizes that core of basic psychological capacities from the earlier time to the later time.19

Unger’s object-level account differs from most psychological continuity theories in maintaining that continuity of a person’s core, not

---

18Unger places a list of stringent constraints on both the manner of presentation of puzzle cases and the sorts of subjects whose intuitions we may take seriously. Unger believes that these constraints are required in order for people’s intuitive reactions to puzzle cases to be truly revelatory of their ‘deepest beliefs.’ In Chapter 9 we will find some reason to doubt that these constraints are principled and, as a result, Unger’s object-level claim itself.
distinctive, psychology is what is important for personal identity. And it differs from standard physical continuity accounts in insisting that the sort of physical continuity required for our survival is not confined to continuity of the same body or brain but to any physical realizer (silicon-based circuitry, for example) which realizes continuity of a person’s core psychologically sufficiently continuously (i.e. with no time lags or gaps). We need not pursue the detail of Unger’s object-level account any further here, for what we are interested in is his meta-story. And, as should now be apparent his meta-story is this: what makes it true that survival requires physical continuity is that physical continuity is the relation that we most deeply believe survival to consist in.

As Stephen White is at pains to point out, prima facie this meta-story appears intriguingly different from the meta-stories we have canvassed above. In fact, at first glance, it looks as if it might be a practice-dependent account of survival (indeed, White goes on to argue, in not so many words, that it is): our survival conditions are determined by our deepest beliefs about our survival conditions, not by some independent set of facts served up by nature or God. However, the crucial question to ask Unger (or indeed any personal identity theorist) in this regard is the following counter-factual question: were our deepest beliefs about our survival conditions different—were they not to deliver up physical continuity but, say, psychological continuity—would personal identity then consist in psychological continuity? Unfortunately, this is a question that Unger himself never really directly addresses. He appears to think that our deepest beliefs about our survival conditions are not only uniform, but also pretty much immutable. And so talk of what we might have most deeply believed is not only moot, but is just to change the subject.

20 A person’s core psychology is that part of her dispositional psychology which she shares with all normal and many subnormal human beings, notably a capacity for conscious experience, a capacity to reason in at least a rudimentary way, and a capacity to form some simple intentions. A person’s distinctive psychology is that part of their psychology that is (more or less) particular to them, for example, my memory of having eaten chocolate-chip ice cream at Moomba or of having written the first part of this thesis.

21 Of course, initial appearances might be misleading. Unger might hold the background assumption that our deepest beliefs reliably track the independently determined facts about personal identity, in just the same way that Plato’s Socrates thought that the Gods were very good trackers of the independently determined facts of piety. If this were so, Unger would hold the ‘real essence’ view of personal identity, regarding our intuitions from possible cases not as determining the facts about personal identity, but as the best means to discover or uncover them. In fact, however, as should become clearer, Unger does not appear to hold such a background assumption.
When pressed by White\textsuperscript{22} on a similar point he writes

In trying to gain \textit{wide acceptance} for certain of my arguments, I presented things in a certain \textit{ecumenical way}. Just so, in the book I said things like this: Suppose that there are people some of whose key concepts, beliefs and values are quite different from ours, but whose whole psychology is just as coherent as ours. For example, coherent with the rest of their mentality, they prefer cheap, quick purely informational “teleportation” over expensive, slow, ordinary travel. Seeking \textit{wide agreement on some} matters of rational choice and value, I then \textit{granted} that, for \textit{those} people, such a preference \textit{might be} rational. But, as I thus ecumenically argued, \textit{even if} that \textit{might} be so—and I never said that it actually \textit{was} so—nonetheless, for \textit{us}, such a preference is \textit{not} rational. For, \textit{at the least}, there is clearly \textit{this difference} between us and them: By contrast with the noted preference that their totality of attitudes might indicate for \textit{them}, our \textit{own different} total psychology will indicate for \textit{us}, a quite \textit{opposite} preference as most rational, namely, the preference for the expensive and slow, but genuine travel.\textsuperscript{23}

This sort of ‘ecumenism’ stands in contrast to many practice-independent theorists who maintain that people who believe that personal identity consists in some relation other than the particular relation in which the given theorist claims it does are simply mistaken. For, they claim, personal identity necessarily consists in the property in which they claim it does, and anyone who thinks it doesn’t is simply wrong. Moreover, since such a community would be wrong about the conditions their survival necessarily requires, it would be irrational of them to organize their person-directed practices around those facts. Unger, it appears, does not think such a community of people would straightforwardly be wrong or, loosely speaking\textsuperscript{24}, irrational, rather he thinks that they would mean (though, by their lights, quite acceptably so) something different to what we mean by personal identity.\textsuperscript{25} They would

\textsuperscript{22}White (1992), pp. 153-8.
\textsuperscript{24}Unger does, however, go on to maintain that they may be, \textit{strictly} speaking, irrational. The issue of the relationship between the facts about survival and the organization of person-directed practices will be discussed more fully in Chapter 3.
\textsuperscript{25}Unger (1990), pp.89-92. As Unger (p. 91.) writes, Shoemaker says that the people in radiation-land have a word ‘person’ that means the same thing as does our word ‘person’. Given his story, the claim that they have a word ‘person’ cannot be questioned. But, even if expressed by the same sort of marks and sounds, it is at least somewhat questionable whether our ‘person’ means the same as a word employed by people whose practices of individuating themselves are so markedly different from our own... Shoemaker appears to assume that because both societies will agree about which entities are persons, and which are not, there will be agreement on all the substantive questions that might ground a
mean something different because the predicate ‘is the same person’ would have a different extension in their mouths to the extension it has in ours. Whereas they would regard the statement ‘A person survives informational teleporting’ as true, we, because (according to Unger) we most deeply believe that a person’s survival requires physical continuity, regard it as false.

Although more ecumenical than some of his practice-independent counterparts, then, Unger is no less practice-independent for that. For in insisting that personal identity necessarily consists in what we now most deeply believe it to consist in, Unger rules out that another community who most deeply believed that their survival required some other relation might have a concept of personal identity like ours. A community of people who most deeply believed that personal identity consisted in psychological continuity (as opposed to physical continuity) and who organized their person-directed practices and values around psychological continuity rather than physical continuity would, nonetheless, mean something different by personal identity. What turns out to be necessary for personal identity, for Unger, then, contrary to initial appearances, is not peoples’ person-directed practices or their deepest beliefs or their values, but the particular property which happens to organize a limited class of those practices (viz. deepest beliefs about survival conditions) for us now. A community of people just like us (perhaps even us a hundred years on) who most deeply believed that personal identity consisted in psychological continuity (without its normal cause, physical continuity) would have changed the subject. They would have ceased to talk about personal identity, for personal identity is necessarily what we now most deeply believe it to be. They would be talking about something else, personal identity*. According to Unger, then, the reference of 'personal identity' is given not by the definite description 'the relation (whatever it is) that a community most deeply believe their survival conditions to involve' but rather, by the particular property that we, as we actually are now, most deeply believe it to involve. Like the rigidifying theories of water and mental states just considered, then, Unger rigidly identifies 'survival' with the property that we as we actually are now most deeply believe it to involve viz. physical continuity.

difference in meaning for their ‘person’ and our ‘person’. But we will disagree with them about when a certain situation contains the same person as does a later, or an earlier, situation. These substantive differences are not to be ignored.
broadly construed. Although, because Unger takes our intuitive reactions
gleaned from puzzle cases to determine the extension of personal identity,
he sounds more practice-dependent than philosophers who hold the ‘real
essence’ view, in fact he is thoroughly practice-independent. For go to a
world where physical continuity does not play the survival role, and
physical continuity counts as survival in that world nonetheless. For the
rigidifier, that a relation plays the survival role in the person-directed
practices and concerns of a community is not necessary in order for that
relation to count as survival. What is necessary for survival for any
people in any world is that there obtains between person-stages the
relation that plays the survival role for us.

2.4 Survival as a natural kind

On a popular account of natural kind terms, they are terms that refer to
kinds that figure in the best scientific explanation of some phenomenon,
that is, the scientific explanation with the best explanatory power and
completeness. So, for example, ‘water’ is often thought to be a natural
kind term, because there is a presupposition that we are referring to
whatever property or nature of the substance best explains why water
boils at one-hundred degrees Celsius, causes certain metal objects to rust,
freezes at zero degrees and exhibits various other behaviours which
chemical science aims to explain. As it happens, this property is the
property of being H$_2$O. And so, *a posteriori*, ‘water’ refers to H$_2$O. In the
case of survival the phenomenon to be explained are our person-directed
practices. So the claim that ‘survival’ is a natural kind term will amount to
the claim that ‘survival’ refers to the kind which best explains why people
identify and reidentify people over time in the way in which they do; why
people normally anticipate the experiences of a future person when and
only when that person is them; why people normally hold a person
responsible for a crime when and only they are the same person as the
person who committed the crime; and so on.

So stated (and as I shall argue in more detail in Chapter 4) the view
that something (‘water’, ‘survival’...) is a natural kind term (rather than a
term which names a social or functional kind) is not necessarily at odds
with practice-dependence; and, indeed, may actually support it. For go to
a world in which some other chemical configuration best explains why
water behaves in certain ways, and ‘water’ will name that kind round
there. A view of ‘survival’ as a natural kind term does not, then, in itself,
rule out multiple realizability: for different groupings of things or
properties of things may be explanatorily significant in different worlds.
The view of 'survival' as a natural kind term will, however, place a
constraint on multiple realizability: namely that the realizer must be a
natural kind in order to count as survival. To put the thrust of this
general point another way: we can be natural kinds' theorists without
being rigidifiers. We can say that 'survival' names the natural kind that
plays the survival role in any world, providing that what plays the role in
any world is a natural kind'. On this view, 'survival' refers to a natural
kind (obviously), but it may refer to different kinds in different worlds
(H2O on Earth; XYZ on Twin Earth, for example). The constraint on
multiple realizability is that where what plays the role in a world is not a
natural kind we have an error theory—there is no survival in that world.
For it is a priori true, on this view, that 'survival' refers only to a natural
kind. (As I hope shall become apparent as we go along, an unfortunate
consequence of this view is that people do not survive in this world. For
the relations that play the survival role in this world seem too disjunctive
and gerrymandered to form a natural kind).

Insofar as the practice-dependent theorist about survival has a
quibble with the view of natural kind terms, so (non-rigidly) understood,
it will be a reasonably marginal quibble—a quibble about whether or not
it is a priori true that 'survival' non-rigidly refers to a natural kind, rather
than just to whatever plays the survival role in a world, natural kind or
not. And this turns on the empirical question of whether we use the word
'survival' to refer only to a natural kind or whether we are happy to have
it refer to a reasonably diverse disjunction of relations which play the
survival role (in the same way in which we are happy to allow 'doorstops'
to refer to a diverse disjunction of things which prop doors open). In
short, it turns on the question of whether or not we think that in a world
(such as our own!) where 'survival' refers to a disjunction of relations
people do not survive, or whether we are happy, in that case, not to
embrace an error theory, but to retreat to the view that 'survival' names a
functional kind. My bet is on the latter.

However, although it's perfectly possible (and, I think, often
plausible) to be a non-rigidifying natural kinds' theorist in this way, as
matter of fact, those who are natural kind theorists are typically also
rigidifiers. And those who think that 'survival' is a rigid designator are
practice-independent. For they think that survival rigidly designates whatever it designates in the actual world, even in worlds where what it designates does not play the survival role. What underpins this natural kind/rigidifying tandem act is, I think, the semantic intuition that if survival is a natural kind, then we had better use the word 'survival' rigidly to designate that kind. That 'survival' names a natural kind is what justifies the chauvinism. In this case, the view that 'survival' names a natural kind will be a motivation for rigidifying on actual realizer, rather than role. But just as natural kinds' theorists need not be rigidifiers, rigidifiers need not be natural kinds' theorists. They might believe that 'survival' is a rigid designator, not because it names a natural kind, but because that is the way we use the word: we use the word 'survival' to refer to the relation that actually plays the survival role, not to the role.

We might think of the difference between what I have called 'non-rigidifying' natural kind view, and the rigidifying natural kind view in the following way. It is perhaps a better way, since calling the non-rigidifying natural kind view, 'non-rigidifying', is potentially misleading. For there is a sense in which the non-rigidifying view is not non-rigid: it claims that 'survival' names a natural kind in all worlds (just not necessarily the same one). So we might perhaps better put the issue like this. There are two views about natural kind terms. One (what I have called the 'non-rigidifying' view) has it that a natural kind term is a term which refers only to a natural kind. The second (what I have called the 'rigidifying' natural kind view) has it that a given natural kind term refers only to a particular natural kind (most likely, though not necessarily, the kind to which it refers in the actual world).

The dialectical point to bear in mind for what follows is that the practice-dependent theorists' main opponent here is the second, not the first, view of natural kind terms. As I have pointed out, there are two possible motivations for believing the second view: the first is that, as a matter of fact, we use the word 'survival' to name a particular natural kind (not any natural kind, or a social or functional kind). The second is that 'survival' is a natural kind term, and so (perhaps even if we did not in fact use it in this way) ought to rigidly designate a particular kind, typically the kind that it names in the actual world. In one way or another, the practice-dependence theorist must deny both these claims.
David Wiggins is a philosopher who explicitly states that person is, or is "akin to", a natural kind concept. He asks: "Is there some scientifically palpable real essence, some nature that is presupposed to a kind's being nomologically founded, which not only underlies in fact but also must underlie the kind person (must if members of it are to be subjects whom we can interpret, and react to in the way in which we react to what we recognize as people)?"26 Wiggins' answer to this question is, of course, 'yes'. According to Wiggins, the sortal person is a functional restriction of the very general natural kind, animal. Persons, for Wiggins, are all and only those animals or animal-kinds which satisfy a certain functional requirement. Wiggins calls his 'natural kinds' view of persons, the 'animal attribute view':

This sees person as a concept whose defining marks are to be given in terms of a natural kind determinable, say animal, plus what may be called a functional or (as I shall prefer to say) systemic component. Perhaps x is a person if and only if x is an animal falling under the extension of a kind whose typical members perceive, feel, remember, imagine, desire, make projects, move themselves at will, speak, carry out projects, acquire a character as they age, are happy or miserable, are susceptible to concern for members of their own or like species...[note carefully these and subsequent dots], conceive of themselves as perceiving, feeling, remembering, imagining, desiring, making projects, speaking...have and conceive of themselves as having, a past accessible in experience-memory and a future accessible in intention..., etc.27

Wiggins's analysis proceeds in the familiar two stages. The first stage involves the a priori, conceptual claim just enumerated. This, for Wiggins, captures the character of 'person'. Just as, in the case of water, the a priori, conceptual or reference-fixing component (character) is given by 'the stuff that [actually] falls from the sky, runs from taps, fills the oceans...', so for Wiggins, the reference-fixing component of the definition is given by 'animal falling under the extension of a kind whose typical

27Wiggins (1980), p.171. Note that, according to this definition, it is a supposed to be an a priori, not an a posteriori, truth that all persons must be animals. Wiggins seems to oscillate between the strong thesis that there is only one animal kind, namely *homo sapiens*, of which all persons are members, and the weaker thesis (expressed in the above definition) that, for any actual or possible person, there is some animal kind of which that person is a member. This latter weak thesis does not rule out the possibility that dolphins and chimpanzees, were they to meet the functional/systemic requirements, might also count as persons. Even the weak thesis does, however, rule out a priori the possibility that sophisticated silicon-based creatures or a brain in a vat (assuming that a brain, though part of an animal, is not itself an animal) might count as persons.
members perceive, feel, remember...’ These list, if you like, the respective phenomena to be explained. The second stage involves \textit{a posteriori} investigation to discover whether or not a given animal or animal-kind (say, a chimpanzee) does in fact satisfy the functional requirement: being a member of a kind whose typical members perceive, feel, remember... And if it does, and only if it does, that animal then counts as a person. As Wiggins writes,

According to this view, a person is any animal that is such by its kind as to have the biological capacity to enjoy fully the psychological attributes enumerated; and whether or not a given animal qualifies is left to be a strictly empirical matter.\textsuperscript{28}

Wiggins, then, views 'person' as a natural kind term which names any animal that satisfies the functional requirement. Wiggins' account is interesting because 'person' does not straightforwardly name a natural kind, rather it names a functional kind which necessarily involves a natural kind. Persons for Wiggins are a functional kind which, necessarily and \textit{a priori}, only animals can satisfy, rather like doorstops are a functional kind which only solid, extended objects can satisfy. And just as different solid, extended objects (books, tables, baskets, dogs...) can play the role of doorstop, so Wiggins leaves it open that different animal-kinds might count as persons. In leaving the person role thus nominally multiply realizable, then, Wiggins is like the second-order functionalist, definite description' or 'contingent identity' theorists. But in maintaining that it is an \textit{a priori} constraint on persons that they be members of some animal-kind, Wiggins denies that 'person' names anything which is not an animal. Elaborating on exactly why Wiggins thinks this would involve a substantial reconstruction of his views about identity in general, and I shall not attempt such a reconstruction here. It seems enough in this context to note that, \textit{contra} Wiggins, there seems nothing particularly intuitively incredible about the thought that, for example, a spiritual entity which satisfied the systemic requirement might count as a person just as well. Indeed, exactly this view of persons was the orthodoxy for a long time (and, I suspect, for many people, still is). Historically, it was the thought that persons might be \textit{material} entities which seemed incredible.

Wiggins' account is, of course, an account of what it is to be a person, not of what it is to be the same person over time. However,

\textsuperscript{28}Wiggins (1980), p. 172.
Wiggins' account of what it is to be a person yields an account of what it is to be the same person over time: what it is to be the same person over time is to continue to exist by whatever conditions are required for the continued existence of the animal-kind of which it is a member. The identity conditions for persons over time are the identity conditions for the particular animal-kind in question, which satisfies the functional requirement. So, at the level of what relation determines the extension of personal identity, Wiggins can be understood as advocating a sort of physical continuity account of personal identity: in order for a person to continue exist they must be the same animal. Unlike practice-dependent accounts, Wiggins takes being the same animal to be necessary for survival: there is no world in which a person survives but is not the same animal. But, unlike standard (practice-independent) physical continuity accounts, Wiggins will presumably hold that role is also necessary for a person's survival: where being the same animal does not meet something like the functional or systemic requirement for persons—where it does not play the survival role—there is no survival. On Wiggins's view, as I understand it, both role and actual realizer are necessary (and jointly sufficient) for survival. If either is absent, there is no survival.

There is one final, possible practice-independent approach to analysing personal identity worth mentioning (although I cannot think of a particular personal identity theorist who obviously takes it), for it is a reasonably popular theory of reference more generally. Whereas almost all the theories discussed so far maintain that we need some sort of description to fix the reference of terms, according to proponents of the causal theory of reference, reference-fixing need not go by way of a description. This approach need not, and typically will not, invoke the method of cases at all. For according to the causal theorist, what is important is not our theory about what properties our words refer to (revealed by the method of cases), but rather what property in the world in fact causes our utterances, and this is something our theory may get wrong. What property causally grounds our utterances is a matter, not for thought experiments, but for hard-core scientific investigation. Since the causal theory of reference need not go by way of the method of cases, the causal theory merits a new category of its own.
2.5 The causal/historical theory

Return to the case of water. The causal theory of reference has it that the word 'water' refers to the stuff water, not (or not necessarily) in virtue of the fact that the stuff satisfies the description 'falls from the sky, runs from taps, gets called 'water' by experts...etc...', but rather because there is some property of the stuff water that caused our initial dubbing, designation or calling of that stuff 'water' (and a subsequent causal chain linking the property of the stuff that caused that initial dubbing to all later utterances of the word 'water'). The substantive (of course, a posteriori) task is that of discovering what the property of the stuff, water, causally responsible for our utterances of the word 'water', is: is it $H_2O$, XYZ,...etc...?

Applied to the case of personal identity, this view will have it that 'personal identity' refers to a certain relation among person-stages in virtue of the fact that that relation caused the initial dubbing of that relation 'personal identity' (and a causal chain linking that relation which caused the initial dubbing to all later utterances of 'personal identity' or 'is the same person as'). This is the causal theorist's meta-theory about why 'personal identity' picks out a certain relation among person-stages. The remaining task for the case of personal identity, as for the case of water, is the task of ascertaining what the relation between person-stages causally responsible for utterances of 'personal identity' or 'is the same person' is: is it physical continuity, psychological continuity, the property of being a relation that plays a certain role in our person-directed practices,...etc...? And to find this out we need to investigate the world, in just the same way as scientists had to investigate the world to discover that the property of the stuff water which was causally responsible for utterances of the word 'water' was $H_2O$ not XYZ. 

---

29 According to proponents of this view, such as Devitt and Sterelny, the causal theory of reference is entirely a posteriori. It is supposed to be a thoroughly a posteriori truth that reference is a causal relation between things and words.

30 As Braddon-Mitchell and Jackson op. cit. point out, there is a problem here for the causal theorist (like Kim Sterelny and Michael Devitt) who maintains that the causal theory need not invoke a description at all in order to fix the reference of terms. The problem has two dimensions. First there is what they call the 'width problem'. Identify the stuff water at a given (late) stage in its development and there will be numerous candidate properties for the nature of water: its mass, weight, chemical configuration, liquid, not a fictional object, and so on. Which of these properties constitute the nature of water? Second, there is the depth problem: which part of the causal chain between our utterances 'water' and properties of the stuff water do we count as the nature of water: which of these properties constitute the nature of water?
considered, the causal theorist takes her theory, including the reference-fixing component or meta-theory, to be thoroughly *a posteriori*. According to the causal theorist, not only does science tell us what properties of things are causally responsible for our utterances, but moreover, science tells us (or will tell us) what reference is.

The causal theory of reference applied to personal identity is yet another sort of practice-independent view. ‘Personal identity’ picks out whatever relation between person-stages it does, not in virtue of the fact that that relation plays a certain role in our person-directed practices, but because that is the relation which grounds the causal chain. According to the causal theorist of personal identity, what we mean by ‘personal identity’—what relation ‘personal identity’ picks out—is, necessarily, whatever the privileged relation is which grounds the causal chain. So that someone who used the word ‘personal identity’ to refer to some relation between person-stages other than the privileged one, would either be mistaken (if the causal theorist thinks that the causal theory actually latches onto the pre-existing, privileged fact of the matter served up by nature or God) or (if the causal theorist merely thinks that the dubbing rigidly fixes the reference of ‘personal identity’) would mean something different by ‘personal identity’ from what we mean by it. Either way, the causal theorist will be practice-independent.

This concludes my comparatively broad-brush reconstruction of the various ways in which at least some of the most prominent philosophers of personal identity to date have been practice-independent. Perhaps there is some meta-view that I have overlooked, but I hope to have canvassed the most obvious and plausible of the practice-

---

According to Braddon-Mitchell and Jackson, to answer these objections the causal theorist must invoke some sort of description: because scientists think that this is the most interesting or appropriate property to associate with water and we defer to the experts in the scientific community, or some such thing. Then the question is: is water to be rigidly identified with H\textsubscript{2}O (assuming the that is the property taken to be the nature of water, rather than some other)? According to Jackson and Braddon-Mitchell, then, the causal theory just ends up as a version of the description theories just discussed.

31 Of course, there is a sense in which the causal theory of reference has the metaphysics of personal identity dependent on our practices: the initial dubbing or naming of a thing is up to us. But this is the trivial sense in which all accounts of the meaning of terms will be practice-dependent that I distinguished in Chapter 1. The relevant question is rather the question of what makes it true that a certain relation counts as the relation of personal identity, and to this the causal theorist answers that it is in virtue of the thoroughly independent fact that that relation (rather than some other) is the relation which grounds the causal chain.
2.6 Retaxonimizing the personal identity debate

What we have now is a new way, or perhaps better, a new dimension, along which we can taxonomize the personal identity debate. Answering the metaphysical-cum-semantic question of personal identity—the question of what are the necessary and sufficient conditions for a person identified at one time being the same person as a person identified at another—requires answering not one but (at least) two questions. It involves answering the meta-question (of what determines what determines the extension of personal identity) as well as the object-level question (of what are the properties (or relations) in virtue of which certain person-stages fall under the extension of personal identity).

Since, to date, the personal identity debate has been focussed almost exclusively on the object-level question, standard taxonomies of the personal identity debate have carved up the debate simply in terms of the different object-level answers offered. So we have the terrain typically carved up as a debate between Non-Reductionists (who claim that personal identity picks out a further fact such as an immaterial entity or soul) and Reductionists (who claim that personal identity consists in the holding of certain more particular facts) and then as a debate amongst Reductionist themselves about exactly what more particular facts personal identity picks out (physical continuity, psychological continuity, the soul, the second-order property of being a property which plays a certain role...?).

But we can, and indeed, should, carve up the debate at the meta-level as well, in terms of the different sorts of answers offered to the meta-question. Here, I suggest, we might broadly carve it up as debate between practice-dependent accounts of personal identity, according to which some or other of subjects' person-directed practices make it the case that personal identity consists certain relations among person-stages, and practice-independent accounts, according to which some independent set of facts served up by nature or God make it the case that personal identity consists person-stages united by a certain relation; and I have attempted to show at least some of the various more plausible forms that practice-independence might take. Along the way, I shall canvass some of the
more specific forms that a practice-dependent account of personal identity might take.
PRACTICE-INDEPENDENCE AND THE CONTINGENCY OF OUR PERSON-DIRECTED PRACTICES

Because practice-independent accounts of survival, whatever their precise form, regard our person-directed practices and concerns as only contingently associated with the concept of survival, for practice-independent theorists, the question of personal identity has been not one, but really two, questions. The first, the constitutive question about what personal identity or survival necessarily involves or consists in, and the second, the evaluative and motivational question about what we care about in personal identity or survival. Or, as it's sometimes differently, and I think wrongly, put: what it is rational for us to care about in personal identity or survival. Put this way, this second question is, in Parfit's famous phrase, the question of 'what matters in personal identity'.

And these two questions have to be two distinct questions given a practice-independent analysis of survival. For, on a practice-independent analysis of the concept of survival, it is always conceptually possible, for all that the analyses themselves tell us, that survival might turn out to consist in something that we don't care specially or directly about. So, for practice-independent theorists, as well as the challenge of providing an account of the logically necessary and sufficient conditions for survival, there is the additional project of accounting for the significance of the role that survival plays in our person-directed practices and concerns; of explaining how and why, given any particular account of survival, survival matters to us in the way that it does. Why is it that, for example, we should be specially and intensely interested in what befalls some future self when and only when that future self is physically continuous with us; why is it that the fact that a person at a later time is physically continuous with a person at an earlier time who committed a crime makes it just to punish the person at the later time; why is it that we should continue to love a person only when they're physically continuous; and so on. As John Perry, a practice-independent theorist himself, puts the challenge

Most of us have a special and intense interest in what will happen to us. You learn that someone will be run over by a truck tomorrow; you are
Chapter 3 The contingency of our person-directed practices

saddened, feel pity, and think reflectively about the frailty of life; one bit of information is added, that the someone is you, and a whole new set of emotions rise in your breast. An analysis of this additional bit of information, that the person to be run over is you, is offered by theories of personal identity, for to say it is you that will be hit is just to say that you and the person who will be hit are one and the same. And so it seems that those theories should shed some light on the difference this bit of information makes to us. If it gives us more reason to take steps to assure that the person is not run over, our theory should help explain why that is so. And if this bit of information gives us reasons of a different kind than we could have if it were not us to be hit, our theory should help explain this too.¹

In a similar vein, Sydney Shoemaker, another practice-independent theorist, writes

A...source of perplexity about personal identity has to do with the special concern persons have for their own continued existence and their own future welfare. Imagine that a wizard demonstrates to you his ability to reduce any object to a pile of dust by a wave of his hand and then, with another wave, to create an exact duplicate of that thing out of another pile of dust. If one really believes that he can do this, one probably would not be too averse to letting him do it to one's kitchen stove. But only a monster would offer his wife or child as a subject for the wizard's trick, and only a madman (or a suicide) would offer himself. Or so it initially strikes us. Our concern for personal identity, the kind of importance it has for us, seems totally different in kind from the concern we have for the identity of other sorts of things. And this is linked to the special concern each person has for his or her own future welfare. It is this that gives point to many of our moral, social and legal practices, and explains the significance they attach to considerations of personal identity. If a person does an action, it is that same person who can later be held responsible for the action, and whom it is appropriate to punish or reward for doing it. If someone buys something, it is that person who is subsequently entitled to the use of the item purchased. These principles, which are constitutive of the institutions of punishment and property and the concept of moral responsibility, are intelligible only against a background of a conception of human motivation in which a central role is played by the special concern each person has for his own future well-being... An account of personal identity...ought to make intelligible the special sort of importance personal identity has for us.²

And what Bishop Butler took to be the most serious objection to Locke's account of personal identity was that it "rendered the inquiry concerning

a future life of no consequence."³ From Locke’s account of personal identity, Butler claimed, it followed “that it is a fallacy upon ourselves to charge our present selves with anything we did, or to imagine our present selves interested in anything which befell us yesterday, or that our present self will be interested in what will befall us tomorrow.”⁴ Why, Butler asks, should the fact that some future person will remember thinking and doing what I am thinking and doing now make me care one bit about what will befall that future self?⁵ According to Butler, Locke’s account of personal identity must be mistaken, for there is nothing in his account which can explain the special sort of concern that we have for our future self.

Of course, this objection misunderstands the sort of account of survival that Locke was offering. For as we saw in chapter 1, Locke took it to be constitutive of some relation being the relation of survival that it could underpin and explain our forensic practices, including the special sorts of attitudes, like non-derivative concern, that we have towards our past and future self. So, on Locke’s account, it was conceptually impossible that personal identity could consist in a relation that rendered us indifferent as to what befalls our future self or what befell our past self. However, despite the fact that his objection was misplaced, Butler, nonetheless, like all personal identity theorists, agrees that it is a requirement on any adequate account of personal identity that it make intelligible and explain our person-directed practices and concerns.

Unsurprisingly, however, the project of accounting for the importance of the role that survival plays in our person-directed practices and concerns has proved a daunting and recalcitrant problem for practice-independent analyses according to which the important role that survival plays in these and other of our person-directed concerns and practices is something that is only contingently associated with the concept. For practice-independent accounts of survival leave a gaping conceptual gulf.⁶

³Butler (1906), p.257.
⁴Ibid., pp. 260 ff.
⁵In criticizing accounts of personal identity in analytic philosophy, Ross Poole (1992, p.15) similarly objects that "it is hard to see how a psychological fact such as memory could have the enormous social and moral consequences that are supposed to follow from personal identity". Why, he asks, "should my being able to recollect having done something explain why I should be punished for it?"
⁶The same is true of practice-independent accounts in the case of value. Since practice-independent accounts of value make value concepts conceptually independent of concepts of subjects’ motivational dispositions, there is the additional project of giving an
3.1 Practice-independence and subject’s person-directed practices: Attempts to close the conceptual gulf.

In the case of survival, practice-independent theorists have sometimes gone to seemingly incredible lengths to attempt to close this gulf. Perry himself attempts to account for this special kind of interest that we have for our future selves by resorting to the claim that the importance of personal identity is really only derivative. We do not, or should not, care that we ourselves continue to exist, for what we really do or should care about, are our impersonal projects. And these might be realized just as well, if not better, by someone who was very like us, though not us. And Parfit, for example, claims that on a psychological continuity account it follows that we should love types not tokens. In continuing to love someone we should not care that the person we love continues to exist. For what we should care about in loving someone are the attractive or desirable qualitative properties they happen to exhibit, and these would obtain if they were killed and replaced by a qualitatively identical replica. So, Parfit argues, it should not matter to us if someone we love is killed and replaced by someone exactly like them, we should continue to love that person’s replica just as we would have loved the person had they not been killed and replaced. What’s more, according to Parfit, it should not matter to us if we ourselves are killed and replaced by a qualitatively identical replica. For what we really do (or should) care about in our own survival, as well as in that of others, is that there be someone psychologically connected to, or continuous with, us, no matter if we ourselves had to be killed so that they could live. Shoemaker moots a similar view.7

However, despite the lengths to which practice-independent theorists have been prepared to go to explain how their account of personal identity explains what matters in personal identity, very few of these theorists have made it entirely clear exactly what the question of ‘what matters’ in survival is.

Chapter 3 The contingency of our person-directed practices

3.2 Two senses of ‘what matters’ in survival

We have many different kinds of concern. Some of these kinds of concern do not directly involve questions of our own survival at all. Amongst our non-survival-involving concerns, for example, is our altruistic concern: concern that the lives of others and the world in general go better, quite independently of whether that promotes our own interests and goals. So, for example, motivated by such unself-interested concern, some people give more than half their income to charity, even though such gifts mean that their own lives are lived in less comfort. Since such altruistic concern, by definition, has nothing to do with questions of what we care specially and non-derivatively about in our own survival over time, concerns of this kind are clearly not at all relevant to the question of what matters in survival. The kind of concern that is relevant to questions of what matters in survival is a particular species of the kind of concern with which altruistic concern is often contrasted: egoistic or self-interested concern.

However, there are (at least) two different kinds of egoistic concern, and it is vital to an adequate understanding of the question of what matters in survival to distinguish between them. On the one hand, in our concern for our own future, we are concerned that the things which we care about in life—our projects, aspirations, values, intentions, relationships and the like—continue on. That is, in our concern for our own future, we are concerned for continuity of the things which give our lives meaning or make our lives worth living. However, we also have a distinctively different kind of egoistic concern for our future self: a kind of concern that we have for our future self and what befalls us quite independently of whether or not our relationship to that future self preserves continuity of the things which give our lives meaning or make our lives worth going on with.

An example of a situation in which the two kinds of egoistic concern come apart might help here. If I were to discover that tomorrow, as a result of some terrible accident, I was to become a paraplegic amnesiac, my life would not seem to me to be worth going on with. For

---

8This is just a report of my own psychological attitudes given my present values. Since being able to play sport, go for bush-walks, be physically independent and remember my life history mean an awful lot to me, being deprived of those things would be sufficient to rob my life of most of the value and meaning that is has for me. Others may not value these things so highly. They may need to construct another example where their
very few of the things that give my life meaning would be preserved in my relationship to this poor future paraplegic amnesiac. In this ‘what makes life worth living’ sense of ‘what matters’, my relationship to my future self would contain little, if anything at all, of what matters in survival. Nonetheless, on any standard account of personal identity, that poor future paraplegic amnesiac will be me; and, precisely because she would be me, I would care very strongly and deeply about what befell that future paraplegic amnesiac. Though my continued existence may have negative value for me—I may prefer to die now than continue to live under those circumstances—I would be far from indifferent as to what befell my future paraplegic amnesiac self. The mere fact that I would be prepared to kill myself, rather than continue on to live a life like that, is itself proof that, although my relationship to my future self does not contain much of what makes life worth living, I nonetheless have some sort of deep and strong concern for that future self. In this second sense of ‘what matters’, then, what matters is preserved in my relation to my future paraplegic, amnesiac self, even though my relationship to that future self contains little, if anything at all, of what makes life worth living.

Steven White is one of the very few who have drawn attention to this critical ambiguity in uses of ‘what matters in survival’. He writes

The preservation of what matters might mean the preservation of what makes life worth living. In this sense, what matters might not be preserved in the existence of some future self in whom one would not hesitate to say one survived. One might, for example, be tortured at some time in the future to such an extent that one would prefer, both before the torture and during it, not to go on living. But the preservation of what matters might have a stronger sense. It might mean that there are future person stages for which one would have (or be justified in having) the sort of special concern that one ordinarily has for one’s future self. In this sense, what matters is preserved in the torture case. One would normally make very extreme sacrifices where one’s present desires were concerned to alleviate some of the pain of the self to be tortured. It is this stronger sense in which it is plausible to identify survival with the preservation of what matters.9

Peter Unger is another of the very few who have explicitly noted this distinction between the two senses of ‘what matters’ in survival.

Chapter 3  The contingency of our person-directed practices

Unger labels the first, weak sense the desirability use: "what it is that one gets out of survival that makes continued survival a desirable thing for one, a better thing, at least, than is utter cessation." Unger distinguishes this desirability use from what he terms the prudential use of 'what matters' in survival (White's strong sense of what matters). Roughly, From the perspective of a person's concern for herself...what future being is there or, possibly, which future beings there are, for whom the person ... should be "intrinsically" concerned. Saying that this ... concern is "intrinsic" means, roughly, that, even apart from questions of whether or not he might advance the person's present projects, there is this ... concern for the welfare of the future being...

And, he goes on

It should...be clear that this prudential use is the motivational use that is most closely related to questions of personal identity. It is the prudential use, not the desirability use, that is the motivationally relevant counterpart of the constitutive use [viz. what it is about a case that counts toward the case being one that involves a person who does survive].

Unger offers a test to employ in our thinking about whether or not some hypothetical process preserves what prudentially matters (as distinct from what desirably matters) in survival; a test which is implicit in the characterization of the strong sense of what matters offered above by White. Unger calls this test, the avoidance of future great pain test; and we are instructed to apply it as follows. In any given puzzle case involving a question of survival, we should imagine ourself to be in the position of the person who is to undergo the hypothetical process which the case involves. If we are not already, we should suppose ourselves to be self-concerned and not to be masochists. We should also suppose that we have no option but to undergo the process. Then we are presented with the following choice:

I may elect to experience, before the process, very considerable pain and, as I know full well, thus ensure that the being who emerges will, after awakening to consciousness, feel no pain. Alternatively, I may elect to endure no pain before the process begins and, as I also know full well, thus ensure that the being who emerges from the process will, after awakening, undergo really excruciating tortures for quite some time.

10Unger (1990), p.93.
11Unger, ibid., p.94.
12Unger, ibid., p.97.
13Unger, ibid., p.29.
If we believe that some hypothetical process preserves what (prudentially) matters in personal identity we will take the first option: we will elect to endure the early pain so that our future self will not have to endure any. We will make sacrifices now to alleviate some of the greater suffering of our later self. If, on the contrary, we believe that the process does not secure what (prudentially) matters in survival, we will choose the second option: because we do not (prudentially) care about what befalls our future self, we will choose to endure no pain now before the process and thus ensure that our future self undergoes the more prolonged excruciating torture.

Unger and White are onto a crucial and often overlooked distinction between two importantly different senses of ‘what matters’ in survival. And, I think, they are right to stress that the relevant sense of ‘what matters’ for questions of what matters in survival is not the desirability or weak sense of what matters, but something like their strong or prudential sense. For although survival is invariably a precondition for the weak or desirability sense of what matters—in order for our lives to be worth living or to continue to have meaning we need, at the very least, to survive—this kind of concern, as we have seen, can and often does come apart from clear cases of a person’s survival, which is the question that, in discussions of personal identity over time, we are interested in. In the personal identity debate, we are interested in the question of what matters in personal identity or survival, not of what gives our lives meaning or makes life, once there, worth going on with. However, the characterization both Unger and White offer of the nature of the strong or prudential kind of egoistic concern relevant to questions of what matters in survival seems to me not quite right.

Unger and White, rightly, want a sense of ‘what matters’ in personal identity which captures the distinctively self-involving or egoistic nature of the special sort of concern which we have for our future self, and for our future self alone. But the way they have characterized this ‘distinctively egoistic’ concern leaves it far from distinctively egoistic. For their characterization in terms of preparedness to make great sacrifices leaves it open that we might have this kind of concern for others. Indeed, as they have characterized it, we often do. We are often prepared to make extreme sacrifices now for others who we care about deeply—our friends, family, culture, even nation, and the like—to avoid their suffering
more greatly later. Making extreme sacrifices (even to the extreme of killing ourselves), is not something we are prepared to do only for ourselves. What is distinctive about the kind of concern that we have only for our future self, and for no one else, then, cannot be just that we are prepared to make extreme sacrifices to avoid or alleviate greater later suffering.

The distinctive (phenomenological) feature of such egoistic concern is, I think, rather this. It is that in our distinctive, non-derivative egoistic concern for our future self, we directly anticipate what will befall our own future self, in a way in which we do not directly anticipate what will befall others, no matter how deeply we may care about their future and what will befall them. We anticipate our own sufferings and joys in a way in which we do not anticipate the sufferings and joys of others, no matter how great a sacrifice we are prepared to make so that they have the greater future joy or do not have to undergo the later suffering. If I know that tomorrow my child is to have a root-canal operation without anaesthetic I may be more concerned and make even greater sacrifices than I would if I knew that I was to have the painful operation. But I do not fearfully anticipate my child’s pain. I do not brace myself, do special mouth exercises and mentally rehearse stoic platitudes, as I do when I know that I will have to undergo the horribly painful operation.

In talking about what people care about in personal identity or survival, and in thinking about what matters to ourselves in personal identity or survival, then, it is important always to bear in mind that the relevant sense of ‘what matters’ for questions of personal identity (though not, of course, for questions of the meaning of life) is the second, what we might call, anticipatory sense. (For those worried about vicious circularity here, chapter 7 is the place to look for reassurance).

One important reason for bearing this in mind is, of course, that this is correct sense of ‘what matters’ for questions of what matters in survival. But there is another important and related reason as well. One surprising thing about the contemporary personal identity debate is that psychological continuity accounts which have it that physical continuity is no necessary condition on a person’s survival have been so dominant. This is somewhat surprising because physical continuity is so incredibly important for survival in our actual lives. Given our present levels of medical technology, the only way a person can survive in actual life is for
them to be physically continuous. So it is somewhat surprising to find that people apparently place so little importance on a condition which is so important for survival in people’s actual lives. What, then, might explain the surprising dominance of the psychological continuity view?

Unger, who, against this trend, argues for a physical continuity account of survival, is, unsurprisingly, particularly concerned to answer this question. And he suggests that psychological continuity accounts may derive a significant amount of their dominant intuitive appeal by illicitly implicitly trading on the desirability sense of what matters in survival. For, for many of us, psychological continuity—continuity of our particular projects, aspirations, values, interests, intentions, and the like—is a substantial part of what makes life worth living.

It is certainly true that few, if any, psychological continuity theorists have drawn attention to the ambiguity in the question of ‘what matters’ in survival. And, indeed, some prominent psychological continuity theorists have quite explicitly couched the question of what matters in survival in the desirability sense, as a question about what makes life worth living. Parfit, for example, glosses the question of what matters in survival as follows

What is it rational to care about, in our concern about our own future?...What is the relation that would justify egoistic concern about this resulting person? If the rest of this person’s life will be well worth living, in what way should I want to be related to this person?14

And Parfit regularly employs this desirability/‘what makes life worth living’ sense of what matters in arguing that psychological connectedness and continuity, rather than some other competing candidate relation, preserves what matters in survival15. More ambiguously, but still

14Parfit (1983), pp. 282-3. Although, in a later piece for popular consumption (Parfit, 1992), he does distinguish a variety of possible senses of ‘what matters’, including the anticipatory sense.
15Interestingly, however, not always. Parfit himself recognizes that there is a tension, a possible trade-off, between the ‘what life worth living sense of what matters’ and survival. For our values may well be better realized in some future person who is a less good survivor, that is to say, less, rather than more, psychological connected to us. Thus, for example, when discussing whether or not fusion preserves what matters (glossed in the ‘what makes life worth living’ way) he writes,

Few of us think of ourselves as perfect. Most of us would welcome several changes in our physical and mental features. If the changes were improvements, we would welcome the partial reduction of both kinds of connectedness. I should avoid fusion if it would predictably involve subtracting features that I value, and adding features that I find repugnant. Suppose that there are only two things that give my
potentially misleadingly, David Lewis glosses the question of what matters in survival as follows:

What is it that matters in survival?...What do I really care about? If it can happen that some features of ordinary everyday survival are present but others are missing, then what would it take to make the difference between something practically as good as commonplace survival and something practically as bad as commonplace death?\(^{16}\)

Unlike Unger, however, who maintains that once this ambiguity is cleared up—once people realize that the relevant sense of what matters is the ‘prudential’ sense—they will come unequivocally and univocally to believe that physical continuity, not psychological continuity, is what they really care about in survival, I am less sanguine about the likelihood of such unequivocal convergence. Nonetheless, Unger is quite right to point out that the relevant sense of what matters is not the ‘desirability’ sense, but the prudential, or rather, anticipatory sense. And this may be especially important to bear in mind in light of the way in which the desirability use of ‘what matters’ may illicitly prejudice us in favour of a psychological continuity account of survival.

Now, this non-derivative, distinctively self-involving anticipatory concern is necessarily future-oriented. For, of course, we cannot anticipate things that have already happened in the past (at any rate, on the plausible assumption that the past is fixed). What, then, is the retrospective corollary of anticipatory concern?

It is interesting, I think, that it is very hard to pinpoint a retrospective counterpart of anticipatory concern. It is difficult to find a retrospective reactive attitude that applies only to cases of personal identity i.e. a reactive attitude that we have only towards the past actions and experiences of past person-stages that we regard as stages of us. We ‘self-ascribe’ (or, as Locke put it, ‘own’ to ourselves) past actions and experiences—taking them to have been done or experienced by person-

---

\(^{16}\)Lewis (1976), p.17.
stages which we acknowledge to be our own, in a way in which we do not self-aspcribe what we take to be the past actions and experiences of others. And this 'self-ascription' or 'owning' of past actions or experiences is, by definition, distinctively self-involving. But the reactive attitudes which typically go along with such retrospective self-ascription—pride, shame, remorse, guilt, embarrassment and the like—are attitudes which are not distinctively self-involving. For we also have these attitudes towards past actions and experiences that we do not self-aspcribe, that we do not regard as having, literally, been done or had by us. Conversely, some people have absolutely no reactive attitudes whatsoever towards past actions or experiences which they unhesitatingly self-aspcribe. Present-day Germans who 'impute' to themselves past Nazi atrocities of which they disapprove, feeling shame, remorse, guilt and the like, are examples of the former; people who feel indifference (no reactive attitudes whatsoever, except that of indifference) towards past experiences which they unhesitatingly self-aspire are examples of the latter. Interestingly, self-ascription seems neither necessary nor sufficient for the having of any of these sorts of retrospective reactive attitudes.

These attitudes are intimately bound up with issues of identity—familial identity, national identity, cultural identity, religious identity, gendered identity, and so on—but do not restrict themselves to personal identity. The present day German who 'owns and imputes' to themselves past Nazi atrocities does not literally think of themself as identical with the person or persons who committed the atrocities—they do not expect to remember directly from the inside the experience of being a prison guard marching Jews to their death, for example—nonetheless, they regard themselves as members of a continuing national group, Germans, who did these things. It is identity of this broader sort, in this case national identity, which renders their attitude of identification and resulting reactive attitudes—of feeling responsibility, guilt, shame and remorse—understandable and justifiable (if it is), in a way in which they may not be were these people not to regard themselves as somehow or other confederated with, or related to, the people who did these things. Identification and corresponding reactive attitudes, although not distinctively and non-derivatively self-involving, are, nonetheless, invariably derivatively self-involving.
If, however, for some reason or another, we wanted a distinctively self-involving attitude which had retrospective application we could give anticipatory future-oriented concern a retrospective application. For we can recast all concern as future-directed concern. For our past self was to our current self exactly as our current self is to our future self. All person-stages are then related to other person-stages by, *inter alia*, anticipatory concern.

3.3 Whether identity matters distinguished from practice-dependence versus practice-independence.

As we saw at the beginning of this chapter, since practice-independent theorists believe that subjects' person-directed practices and concerns are only contingently associated with the concept of personal identity, answering the question of personal identity has involved answering, not one, but two, conceptually distinct questions: the first, the question of what personal identity necessarily involves or consists in; and the second, the question about what we do or should care about (in the aforementioned, anticipatory sense) in questions of personal identity.

A reasonable reaction, however, might be to think that these two questions aren't really two distinct questions, but rather two aspects of one. For it seems reasonable to suppose that, especially given the distinctively self-involving nature of anticipatory concern, we would anticipate the experiences of some future person when, and only when, that future person will be identical with us. In other words, it's an intuitive assumption that I will be specially and directly concerned for some future person when and only when that future person is me.\(^\text{17}\) I think, in an important sense, this intuitive reaction is right. An analysis which adequately captures the nature of our concept of survival should not allow that our survival conditions might come apart from what matters to us in survival.

Famously, however, Parfit has challenged the widely-held tenet of common sense: that identity matters. Paradoxically, according to Parfit, identity is not what matters in personal identity. At first sight, it might appear that a practice-dependent account of personal identity—according to which subjects' person-directed practices, including their anticipatory

\(^{17}\)This is not to say that I am indifferent as to what befalls others, but just that I have a special, direct and intense interest in what will befall me.
concerns, are what make it true that personal identity consists in a certain relation—simply rules out Parfit's argument *a priori*. For, it might be thought, such an account makes it constitutive of personal identity that it be a relation which matters to us, and so conceptually impossible that identity be not what matters. However, although a practice-dependent account certainly makes it constitutive of some relation counting as the relation of *survival* (we will appreciate the force of this term in a moment) that that relation be one that matters to us, it does not rule out Parfit's argument *a priori*. For, the debate about whether practice-dependence or practice-independence is the right meta-story for the case of personal identity, is a different debate from the debate about whether or not identity matters in personal identity.

According to Parfit, personal identity (PI) consists in R—relations of psychological connectedness and/or continuity, with the right kind of cause—where, and only where, R holds uniquely (U) i.e. where, and only where, R holds between one present person and *only one* future person. As Parfit writes,

The view that I accept can be stated with this formula: PI = R + U.18

This is what we might call a *constitutive* claim about the nature or facts of personal identity: about what object-level relation personal identity consists in, and (although Parfit never explicitly mentions it) about what makes it true that personal identity consists in that relation.

Now, for Parfit, since personal identity (R + U) is determined independently of our person-directed practices and concerns, there is the second, what we might call, *motivational* question of what we do, or, for Parfit, should, care about in personal identity—that is, a question about the motivational significance of the facts of personal identity. (The fact that these two questions are distinct for Parfit is itself proof that Parfit is practice-independent about personal identity. For the practice-dependent theorist regards these two questions as inextricably intertwined: personal identity is analysed, *inter alia*, in terms of our person-directed practices and concerns). Parfit's answer to this second, motivational question is that what it is rational to care about in personal identity is not uniqueness

---

(U), but R. And, since identity requires uniqueness, identity cannot be what matters in personal identity.

Parfit's main argument for the claim that uniqueness, and so identity, doesn't matter in personal identity goes by way of the case of fission, and since the conclusion may seem surprising, it may be worth a brief, schematic rehearsal. Briefly, in fission, one person stands in a relation to two later people which, but for the existence of the second later person, we would be happy to call personal identity. Because of the existence of two excellent and equally good physical and psychological continuers, rather than one, we cannot say that the original person continues to exist: given the logic of identity, two numerically different people cannot be one and the same person. Nonetheless, Parfit claims, all that it is rational for us to care about, in just the way in which we normally care about our future selves, is preserved in the original person's relation to the two resulting people: it is just preserved twice over. So, Parfit concludes, identity cannot be what matters in personal identity. That is, uniqueness or our continued existence cannot be what it is rational for us to care specially and directly about in our future-directed concern. What matters in personal identity, or what is properly the focus of our future-oriented concern in survival, Parfit positively claims (although not from the case of fission alone), are the more particular relations of psychological continuity and connectedness, (R), which happen to coincide with personal identity (i.e. hold uniquely) in ordinary life only to the extent that cases of fission-like branching do not in fact occur. And since, where they come apart, the more particular relations and not uniqueness are what matter, we ought care about the more particular relations and not uniqueness, even in ordinary, everyday life where R and U happen to coincide.

Having distinguished between the constitutive question of what are the facts of personal identity, and the question of the motivational significance of the facts of personal identity, we are in a position to see why the question as between practice-dependence and practice-independence is a different question from the question of whether or not identity matters. The question of whether identity matters is a question about the motivational significance of facts about personal identity, not about the constitutive question of what the facts of personal identity are, or what determines what determines the facts of personal identity.
Whereas, the debate between practice-dependent and practice-independent theorists is a debate about the constitutive question—about what the facts of personal identity are—in particular, about what determines that these relations rather than those count as the relations of personal identity.

Unfortunately, however, despite the fact that these are different issues, we need to introduce some terminology to separate them, and keep them separate. Because the practice-dependent theorist, unlike the practice-independent theorist, regards the constitutive and motivational questions as inextricably intertwined—the latter determining what determines the former—if the practice-dependent theorist couches their analysis as an analysis of personal identity, which builds in the requirement of uniqueness, the practice-dependent theorist will unwittingly succeed in ruling out Parfit's argument a priori: something I, at any rate, do not want to do. So, in order to leave the question of whether uniqueness matters an open question, we need some terminology for talking about something which is just like personal identity—viz. being the relation which organizes our person-directed practices—but which does not build in the requirement of uniqueness.

Let us say that personal identity over time consists in the holding among person-stages of certain properties or relations (for Parfit, R; for physical continuity theorists, physical continuity; for Cartesian dualists, the soul\textsuperscript{19}) when, and only when, these relations hold uniquely or in a non-branching, one-one form (i.e. when cases of fission-like branching or fusion do not occur). Personal identity, then, equals a certain property or relation among person-stages plus uniqueness. (This is the standard usage). Let us say, less standardly (and here's the technical terminology that the practice-dependent theorist needs), that a person's survival over time consists in whatever property or relation among person-stages it is in which personal identity is taken to consist (be it physical continuity, psychological continuity, the soul...), but without or minus the requirement of uniqueness. So talk about what makes for a person's survival over time is just the same as talk about what makes for personal identity over time, it's just that talk about survival leaves it open as to

---

\textsuperscript{19}In point of fact, of course, since for Cartesian Dualists the soul is necessarily indivisible, it is impossible for it to hold among person-stages in anything but a one-one or non-branching form.
whether the relations in which personal identity is taken to consist hold uniquely, in a one-one or non-branching form. Since the question of what survival consists in is just the question of what personal identity consists in, without the requirement of uniqueness, theorists who are practice-independent about personal identity will therefore, ipso facto, be practice-independent about survival.

With this distinction in place, we should simply recast the issue between practice-dependence and practice-independence as an issue about what determines what determines the extension of survival, in the aforementioned technical sense. The practice-independent theorist, then, will claim that survival necessarily consists in a certain relation (psychological connectedness and/or continuity, or physical continuity, the soul, etc.) in virtue of those being the facts that nature or God demand that we incorporate in the extension of survival; the practice-dependent theorist will claim that survival consists in a certain relation because that is the relation around which subjects' organize their person-directed practices and concerns. But this way, the practice-dependent theorist does not beg-the-question against Parfit's claim that uniqueness, and so identity, does not matter; for survival, unlike personal identity, does not require uniqueness. Although the practice-dependent theorist claims that survival necessarily matters, then, it is an open and separate question as to whether uniqueness also matters in survival.

One very important point to note. My technical sense of 'survival' is not the sense in which the term 'survival' is normally used in the personal identity debate. This should be no surprise, since few have addressed the issues that I am addressing, and so there has been no need for such a technical sense. Perhaps I should have found another term, but I want a word that captures the important sense in which what we have called 'survival' is similar to personal identity, viz. their both being concepts of a relation around which we organize our person-directed practices. Most often in the personal identity debate, 'survival' is just used interchangeably with 'personal identity', to mean personal identity in my aforementioned technical sense. Potentially more confusing in this context, however, is Parfit's use of 'survival'. For, Parfit sometimes uses the term 'survival', not in my constitutive, technical sense outlined, but in the motivational sense of what we do (or should) care about in questions of personal identity. That is to say, Parfit sometimes uses the term
'survival' just to mean 'the relation that we do or should care about in personal identity', namely, for Parfit, R, but not uniqueness. I cannot stress strongly enough that this not the sense in which I am using the term 'survival'. Were Parfit's motivational use of 'survival' my use, Parfit would turn out to be practice-dependent about survival, for, on his use, survival just is the relation that we do and should care about. But Parfit is not practice-dependent about survival in my sense. For Parfit believes that the facts about survival in our sense, namely R (with no requirement of uniqueness), are determined quite independently of our practices and concerns. My constitutive use of the term 'survival' and Parfit's purely motivational use will be coextensive—on both uses Parfit's answer will be R. But this is not because Parfit is practice-dependent about survival in our constitutive sense. Rather it is because Parfit thinks that survival in our constitutive sense, and what we do (or should) care about in personal identity, happen to coincide. But for Parfit, unlike the practice-dependent theorist, there is no conceptual necessity about this coincidence. For, for Parfit, like all practice-independent theorists, our person-directed practices, including what we care about in survival, are only contingently associated with the concept of survival.

This is despite Peter K. McInerney's strange remarks to the contrary. McInerney appears to think that Parfit departs from the orthodoxy in the personal identity debate by being, in my terms, notably practice-dependent. He begins by noting, correctly, that

In memory and psychological connectedness theories, some form of the epistemic relation of first-person remembering plus the endurance of other psychological characteristics unites later person-stages and earlier person-stages into one person; whether one is proud, ashamed, or feels responsible for characteristics of earlier person-stages is thought to be dependent upon personal identity, but not in any way to constitute it. Similarly, John Perry thinks that whatever relations between a present person-stage and later person-stages unite these into one person should underlie and explain concern for one's own future; concern for (one's own) future person-stages does not in any way constitute those future person-stages as one's own.20

But, curiously, in a footnote to this passage, he writes

Parfit is the major exception; he maintains that valuational attitudes towards certain earlier person-stages play some role in making these part of one’s own past. See ‘On “The importance of self-identity”’...”\(^{21}\)

McInerney must have in mind the following passage in ‘On “The importance of self-identity”’, where Parfit writes

The man who feels remorse can reflect upon the lessening in psychological connectedness. If this weakens his remorse, he is to some extent deciding what is to be the part of his past with which he identifies. He is, to this extent, deciding what is to count, for him, as the history of his present self.\(^{22}\)

As should be apparent, however, even from this passage from where McInerney must have drawn support for his interpretation of Parfit (since it is the only paragraph where Parfit sounds even remotely practice-dependent), Parfit is in fact no exception to the practice-independent accounts of survival which have dominated the personal identity debate. Note the give-away wording of the second sentence: ‘If this weakens his remorse, he is to some extent deciding what is to be the part of his past with which he identifies.’ As should be apparent from his wording, for Parfit, what counts as the man’s past is determined independently of his attitudes of identification; the man is simply reflecting upon what part, if any, of that independently determined past he identifies with. That this is Parfit’s view, and not simply a loose use of words, is blatantly apparent from everything else he writes in that article and elsewhere. For example, a few pages earlier, he writes

The man that I described has a divided attitude towards his past. The latter part he regards with pride and shame, pleasure and regret; the earlier part he regards with indifference. The man’s divided attitude has the following cause. Between him now and his recent self there are strong psychological connections; between him now and his earlier self there are only weak connections. \textit{If we take the Complex [reductionist] View, it will be the strength of these connections that we think important. Where the strength differs, we may think it justified to have a different attitude.}\(^{23}\) (Emphasis mine)

As this passage should make apparent, Parfit, just like Perry, Swinburne, Williams and other practice-independent theorists, believes that what underlies, explains and justifies our retrospective (and future-oriented)

\(^{21}\)ibid., p.244
\(^{23}\)ibid., p. 686.
concerns and affective attitudes of pride and shame, pleasure and regret are the facts in which survival independently consists: in Parfit's case, R. To put the point the Euthyphro way: we are justified in organizing our person-directed practices around some relation because that is the relation of survival; R does not count as the relation of survival because we organize our person-directed practices around it. For Parfit, our reactive attitudes and practices no more constitute past or future person-stages as our own, than do electron microscopes make it true that these particles (rather than those) are the atoms.

Parfit's view that identity does not matter in personal identity may give rise to another, related confusion that needs to be cleared out of the way. Because Parfit believes that identity does not matter, he sometimes says things that may sound deceptively practice-dependent. For instance, he frequently writes things of the sort: "we could say that the resulting person will be me, or we could say that I shall die and he will be someone else". How we choose to apply the language of personal identity—whether we decide to call a future person me or not—is up to us.

However, we should not let Parfit's way of talking mislead us. Parfit is not claiming that our application of the language of personal identity itself makes it true that a later person is the same person as an earlier person. Quite the contrary. As we have seen, Parfit believes that the facts of personal identity are determined independently of us, including our use of language. Parfit can afford to be so ecumenical about the language of personal identity only because he believes that identity is not the motivationally significant feature of personal identity, and so that the language of personal identity is neither here nor there in latching onto the motivationally significant facts. Parfit's interest is in survival, not personal identity; for survival (for Parfit, R), not uniqueness, is what it is rational for us to care about in survival. And the facts of survival, for Parfit, like the facts of personal identity, are determined quite independently of us, our practices and concerns, and our use of language.

We should think of the debate between practice-dependent and practice-independent accounts of personal identity, then, as concerning survival, in the aforementioned precise technical sense. In particular, as concerning the question of whether or not subject's person-directed

---

practices, including their anticipatory future-oriented concerns, are an essential feature of the concept of survival or something only contingently associated with it. That way we can avoid some confusion and, for the time being, leave aside the separate issue of whether uniqueness matters or not. So nothing in practice-dependence as such, rules out Parfit's argument that uniqueness doesn't matter a priori. (On the practice-dependent view I shall advocate, whether uniqueness matters or not (i.e. whether a community of people anticipate the future experiences only of future selves who are unique continuers of them or not), all depends upon whether the community in question happens to care about uniqueness or not.) From hereon in, then, I will use the terms 'survival' and 'personal identity' in the constitutive, technical senses outlined, unless otherwise noted.

3.4 Exclusive and inclusive justifiers

All practice-independent theorists, be they Cartesian, causal/historical, rigidifying, real-essence or some other, believe that survival necessarily consists in the relation in which it consists for us, so that people, even people just like us, who took survival to consist in some other relation would either be mistaken (for Cartesians and real-essencers) or, for rigidifiers, would mean something different from what we mean by 'survival'. (It is in this way that practice-independent accounts, whatever their precise form, are chauvinistic, parochial or ethnocentric.) All practice-independent theorists, too, and relatedly, think that survival has nothing essentially to do with subjects' person-directed practices. Because they think this, practice independent theorists view the facts of personal identity as independent justifiers: insofar as we are justified in organizing our person-directed practices and concerns around some relation, it is because survival necessarily consists in this relation, not the other way round. However, practice-independent theorists appear to divide as to the normative and motivational significance of survival. They divide as to the question of whether or not the facts of survival are, as we might put it, 'exclusive justifiers'. We saw this earlier when, in chapter 2, we discussed

\(^{25}\)Some practice-independent theorists may take a different view, according to which there is a hierarchy of candidate relations for survival. So, for example, (and I think Parfit holds this view) survival consists in psychological continuity in a soulless world, otherwise it consists in continuity of the soul. The soul is, as it were, the best candidate for survival, but the next best thing is psychological continuity. This view is practice-independent because the hierarchy is determined by our concept.
Unger’s relative ‘ecumenism’ with respect to the community who organized their person-directed practices around psychological, not physical, continuity. I want now to bring out this difference more explicitly.

According to one group of practice-independent theorists, survival is the relation that all people, at all times, should care about. So that people (or, at any rate, people sufficiently like us) in this world or any other who organized their person-directed practices around some other relation would be irrational, mistaken and unjustified. These philosophers regard the facts of survival not merely as independent justifiers, but as exclusive justifiers: for they, and they alone, are the facts which justify a community’s person-directed practices, not just for some limited group of us, but for everyone everywhere. For these philosophers, there is only one thing that people like us can correctly organize their person-directed practices and concerns around, and that is the relation that we organize (or ought to organize) our person-directed practices around. There are no true or acceptable alternative concepts of survival for people of our biological kind. So that subjects like us who organized their person-directed practices and concerns around some other relation, say, survival*, would simply be mistaken.

However, a second group of practice-independent theorists, Unger amongst them, appear to take a different, more ecumenical, view. True enough, these philosophers will say, survival is the relation around which we organize our person-directed practices, and which independently justifies us in organizing our person-directed practices around that relation. However, as against the first group, this second group do not require that all people everywhere are similarly rationally obliged to organize their person-directed practices around the facts of survival just as we do. A community who organized their person-directed practices around some other relation, survival*, are not necessarily irrational, mistaken or unjustified. In contrast with the first group who maintain that the facts of survival are exclusive justifiers, these second group of philosophers have the facts of survival, as we might put it, as inclusive justifiers: facts which justify preferences where they cohere with the views and concerns of the people involved.

On the first, exclusive justifier view, then, survival is the only thing that people are justified in organizing their person-directed practices
around, on the second view, people might quite correctly and rationally (or at any rate, not clearly irrationally) organize their person-directed practices around some other relation, survival*, or survival**, etc. Both these views are practice-independent, for they both maintain that the facts of survival are essentially determined by facts which obtain or fail to obtain independently of people’s person-directed practices, they just differ about the motivational status of those facts.

Parfit is an example of a practice-independent theorist who takes the first view. In summing up his central claims about survival, Parfit, for example, writes

...I believe that I have now considered those views that, in this debate, need to be considered. I may be unaware of some other published view. And I have not considered views held in different ages, or civilizations. This fact suggests a disturbing possibility. I believe that my claims apply to all people, at all times. It would be disturbing to discover that they are merely part of one line of thought, in the culture of Modern Europe and America...Fortunately, this is not true. I claim that, when we ask what persons are, and how they continue to exist, the fundamental question is a choice between two views. On one view, we are separately existing entities distinct from our brains and bodies and experiences, and entities whose existence must be all or nothing. The other view is the Reductionist View. And I claim that, of these, the second view is true...Buddha would have agreed. The Reductionist View is not merely part of one cultural tradition. It may be, as I have claimed, the true view about all people at all times.26

At first sight, this may appear to be a claim only about the intrinsically privileged status of Parfit’s claims about survival, not about the evaluative or motivational significance of those facts. Parfit believes that his account of survival applies to all people at all times. However, implicit in this, is a claim about what matters in survival or, as Parfit prefers to put it, what it is rational to care about in survival. For, according to Parfit, the Reductionist view of survival is inextricably entwined with a particular view about what it rational to care about in survival. For, according to Parfit, “on a reductionist account, personal identity cannot be what matters”.27 Reductionism about survival goes hand and hand, for Parfit, with ‘reductionism about significance’: the view that “when one fact just consists in certain others, it can only be these other facts which matter.”28

---

28loc.cit.
Indeed, the reason Buddha would have agreed (conveniently italicized by Parfit), is that, according to Parfit, Buddha too believed that, in one's concern for one's future, it was irrational to care about identity. Not only is it irrational to care about uniqueness, according to Parfit, but people who care about physical continuity, or even that their psychology be furthered by way of reliable cause, are irrational. He writes, for example,

...can I rationally care a great deal about whether this [future] person's brain and body will be my present brain and body? I believe that, while it may not be irrational to care a little, to care a great deal would be irrational...Why would it not be irrational to care a little? This could be like one's wish to keep the same wedding ring, rather than a new ring that is exactly similar. We understand the sentimental wish to keep the very ring that was involved in the wedding ceremony. In the same way, it may not be irrational to have a mild preference that the [future] person...have my present brain and body.29

According to Parfit, what it is rational to care about in survival is psychological continuity, with any cause, and any person at any time who did not care about this relation would be irrational. Why? Good question! We will come to that shortly.

In contrast to Parfit, Sydney Shoemaker takes the second sort of view; the view of the facts of survival as inclusive justifiers. Shoemaker asks us to imagine the following case (and since the case will be useful in illustrating the difference between the views of the facts of survival as inclusive and exclusive justifiers, it may be worth presenting it in full).

Imagine a society living in an environment in which an increase in some sort of radiation has made it impossible for a human body to remain healthy for more than a few years. Being highly advanced technologically, the society has developed the following procedure for dealing with this. For each person there is a stock of duplicate bodies, cloned from cells taken from that person and grown by an accelerated process in a radiation-proof vault, where they are then stored. Periodically a person goes into the hospital for a 'body-change'. This consists in his total brain-state [all his present mental states or psychology] being transferred to the brain of one of his duplicate bodies. At the end of the procedure the original body is incinerated. We are to imagine that in this society going in for a body-change is as routine an occurrence as going to have one’s teeth cleaned is in ours. It is taken for granted by everyone that the procedure is person-preserving. One frequently hears remarks like ‘I can’t meet you for lunch on Tuesday, because that is the day for my body-change; let’s make it Wednesday

instead.' All of the social practices of the society presuppose that the procedure is person-preserving. The brain-state recipient is regarded as owning the property of the brain-state donor, as being married to the donor's spouse, and as holding whatever offices, responsibilities, rights, obligations, etc. the brain-state donor held. If it is found that the brain-state donor had committed a crime, everyone regards it as just that the brain-state recipient should be punished for it.30

Shoemaker further asks us to grant, for the sake of argument, that materialism is true in this world: there are no non-physical entities such as souls; and that the members of society know this, and know exactly what the brain-state transfer procedure involves, so that “[t]here is no clear sense in which they can be said to be mistaken about a matter of fact in regarding the procedure as person-preserving.”31 Now, Shoemaker says, confronted with such a society, there is strong reason to think that what they mean by ‘person’ is such that the BST-procedure is person-preserving (using ‘person’ in their sense). And, what goes with this, it would be very hard to maintain that they are being irrational when, being under no misconception concerning matters of fact, they willingly submit themselves to the BST-procedure.32

Swinburne, however, disagrees. Like Parfit, Swinburne holds the first view of the facts of personal identity as exclusive justifiers. He writes,

There may well be, as Shoemaker supposes, some reason for supposing that a society, in which BST is common, means by ‘person’ (pp.109f) ‘what we mean by it’. But in that case there is good reason for anyone who reflects on the issue to doubt whether ‘the BST procedure is person-preserving’, even if that society unthinkingly supposes that it is. Many societies have believed that present persons are their dead ancestors reincarnated. The fact that the society does hold that view about personal identity has no tendency to show that that view is true and the same goes for any society which holds the view that the BST-procedure is person-preserving.33

For Swinburne, whether or not a community is justified in organizing their person-directed practices around some relation depends on whether they are right about what survival necessarily involves, irrespective of the fact that, by their lights, survival* gives the conditions for survival, and

31loc.cit.
32loc.cit.
33ibid., p.133.
survival* is what organizes their person-directed practices and concerns. For Swinburne, survival is what all people everywhere ought to organize their person-directed practices and concerns around, and those who organize them around some other relation, survival*, are simply mistaken and unjustified in so doing.

What might underpin or explain this division amongst practice-independent theorists? I suspect that it has very little to do with views about survival as such, and much more to do with views about rationality and morality. In particular, it has to do with whether the facts of survival are regarded as intrinsically motivating—as being, in and of themselves, the sorts of facts that people ought care about—or whether such motivational status as the facts of survival have derives from the views and preferences of the people involved. Unger implicitly offers this sort of explanation when he writes,

Owing to a certain flexibility in the semantics of “rational”, we may employ the term to set contexts where the satisfaction of lenient conditions suffices for its correct application and, as well, we may employ it to set stricter contexts. Properly setting a lenient context, we can have it so the satisfaction of certain (quasi-)formal conditions by a person’s attitudes may suffice, near enough, for his desires, values, choices, etc. correctly to be called rational... Though not useful for my... ecumenical presentation, there’s also room to use “rational” more strictly, and that room is helpful for certain other things. For example, when used more strictly, often the terms is employed to get “the authority of reason” behind values, etc. that the speaker recognizes to be good ones—or, for those less objectively inclined, behind values, etc. that the speaker endorses. Two simple examples may clarify my point: (1) At least when I’m the one speaking, even on a very strict use, a person’s deep basic desire for loving relations with others is correctly called rational. (2) But, at least when I’m speaking, on such a strict use, someone’s basic desire to spend his life collecting string is not correctly called rational, no matter how well it may cohere with the rest of his psychology.34

Unger himself does not take a strong stand as between the lenient and the strict use of ‘rational’. He merely points out that, being lenient or ecumenical, a society such as the BSTs can be correctly called rational (or, at any rate, cannot properly be called, irrational); being strict and chauvinistic, they could not properly be called rational. But, in permitting the lenient use as at least a valid one, Unger allows that the facts of survival could be inclusive justifiers. (He is not, of course, so ecumenical

about the motivational significance of survival for us. For, according to Unger, our views and preferences have it that the facts of survival—for Unger, physical continuity—and not the facts of the BSTs’ survival*, are the motivationally significant facts for us.

The distinction Unger draws between the lenient and the strict sense of ‘rationality’ is, roughly, the difference between the Humean and the non-Humean view of motivational reason or rationality. Famously, David Hume maintained that reason is slave to the passions. For Hume, non-instrumental brute preferences or desires, unlike beliefs (which aim to represent the way the world is and so can represent it truly or falsely), cannot be assessed in terms of truth and falsity, and fall outside the sphere of rationality—at least, providing that they do not rest on any false beliefs. Instrumental desires are those desires which serve to maximise satisfaction of our brute desires. For example, if we had a brute, basic or non-instrumental desire to drink coffee, an instrumental desire which served this brute desire would be the desire to go to a coffee shop. Instrumental desires are the means to satisfy our brute desires. Instrumental desires can serve our brute desires better or worse—some means are better than others for getting to a desired end. As a result, for the Humean, instrumental desires, where they do not serve to maximise satisfaction of our brute preferences, can, and will, be irrational. But, for the Humean, the brute desires that our instrumental desires serve are themselves never properly called irrational. For the Humean, such non-instrumental preferences or brute desires may be immoral—they may stand in need of moral justification—but they are not properly called irrational. The fact that we have a certain non-instrumental desire (say, a basic sentimental desire to keep the very same wedding ring that was placed on our finger at the ceremony) is just a brute fact about us, and one that, where it is not the product of some other false belief (for example, the belief that the wedding ring has a magical power to keep demons away), cannot be subjected to rational criticism.

In contrast, the non-Humean believes that some, although not all, such brute desires can be intrinsically irrational. For the non-Humean, brute desires, like beliefs, can properly be subjected to rational criticism, and not just where or because they rest on false beliefs. In support of her case, the non-Humean cites our folk intuitions and practices of saying that ‘so-and-so has such-and-such a "stupid" desire’, where the desire in
Chapter 3  The contingency of our person-directed practices

question does not appear to rest on a false belief. The non-Humean takes our dispositions to say, for example, that 'the desire to spend one's life collecting string is "stupid"', to support a non-Humean, as against a Humean, account. To non-instrumentally desire to spend one's life collecting string, the non-Humean might say, is intrinsically irrational, so that anyone who had such a preference would properly be called irrational. (The Humean, in contrast, will say that this desire is not properly called irrational, although it may well be called immoral—there may be better things we could be doing with our life. From the Humean's point of view the folk are disposed to use the word 'irrational' too loosely.) However, since our intuitions about whether or not a desire is intrinsically irrational vary from case to case—we do not, for example, tend to think that the desire to spend one's life having loving relationships is irrational—the anti-Humean's claim that certain desires are intrinsically irrational must go on our intuitions on a case by case basis.

Whereas Shoemaker and Unger are, if somewhat tentatively, Humean; Swinburne and Parfit, and others who view the facts of survival as exclusive justifiers, are non-Humean. They claim that the facts of survival—in Swinburne's case, the soul; in Parfit's, R—are intrinsically motivating, so that someone who did not care about those facts would be irrational. According to Parfit, it is not irrational to have mild preference for one's original wedding ring, but, by implication, it would be irrational to care a great deal. It would be irrational to care so much that, for example, when one lost the ring, one refused the offer of a qualitatively identical replacement. Why would it be irrational? It would be irrational because having such a great desire for one's original wedding ring is an intrinsically irrational desire.

The Humean will happily admit that there are circumstances under which a person's great desire for their original wedding ring might be irrational. It might be irrational, if the person had false beliefs about the ring. It might be irrational, too, if what the person really cared brutely about in their caring about their wedding ring was the colour, shape or feel of the ring. For then their great desire for the original wedding ring would be an instrumental desire—they greatly desire the token ring only because it happens to instantiate the properties which they really, brutely desire—and, what's more, a potentially irrational one: if, for example, they were, as a result of this instrumental concern, to turn down the offer
of a qualitatively identical replacement ring. For all the things they
brutely care about in their caring about their original wedding ring would
be secured by a qualitatively identical replacement. Such an instrumental
desire would be irrational, because it would not serve, but thwart,
maximal satisfaction of the person's brute preferences.

Now, Parfit seems to think that our concern for uniqueness, and so
identity, in survival is just such an instrumental desire. According to
Parfit, we think we care non-instrumentally about identity or uniqueness,
but we are wrong. What we really care non-instrumentally about are the
more particular facts (R) which, mistakenly, we confuse with a brute
desire for personal identity, since in everyday life the two things go
together. Our desire for identity or our own continued existence is then,
for us, an instrumental desire: we care about identity only because, as
things actually stand, ensuring that we continue to exist is the only way of
ensuring that there will be some future person who is R-related to us.
That such instrumental desires can be irrational is common ground
between the non-Humean like Parfit and the Humean. The Humean can
readily agree that, if Parfit is right, and people actually do have a brute
desire for R rather than uniqueness in questions of survival, then caring
greatly about uniqueness is rational where, as is the case in our actual
lives, it is the only way to guarantee satisfaction of the brute desire for R;
but irrational where, in cases like fission, it does not best serve to
maximize satisfaction of the brute desire for R. But, the Humean will
baulk at Parfit's further claim, that subjects are rationally obliged so
brutely to care about R so that a community, neither in ignorance or error,
who had a brute, non-instrumental desire for physical continuity, for
example, would be irrational. In the case of survival, as we shall see, the
Humean's case has extra support. For, not only is it doubtful that such
brute, non-instrumental concern is obviously in the same intuitively
irrational boat as the desire to devote one's life to collecting bits of string
(rather than desiring to spend one's life in loving relationships) in the first
place, but, moreover, there is substantial disagreement about what facts, if
any, it is in virtue of which they are irrational.

Note that this is contra Parfit's claim that, since R and not uniqueness is what we care
about where the two come apart, R but not uniqueness is what it is rational to care about
even in ordinary, everyday life where the two go together.
I am Humean about motivational reason. But, and I should stress this point, whether one is Humean or non-Humean is a question about theories of rationality, not about personal identity: it is a view about the motivational status of facts about personal identity, not about what those facts are. The view of personal identity I advocate, like any other, is compatible with either Humeanism or non-Humeanism, and so one which theorists of either persuasion are free to endorse. Unsurprisingly, however, since I am Humean, the non-Humean may (and probably will!) disagree with some, although by no means all, of the normative and motivational upshots I draw from pluralism. But to argue for Humeanism is a thesis in itself.

One final point to note before we move on. Recall Parfit’s claim that reductionism about personal identity is inextricably bound up with a certain motivational view of those facts, namely, ‘reductionism about significance’. According to Parfit, reductionists about personal identity “cannot defensibly deny” that the more particular facts, not identity or uniqueness, are what it is rational to care about in survival. The foregoing should give us reason to doubt that views about personal identity and particular views about the motivational status of facts about personal identity, such as reductionism about significance, are necessarily as inextricably entwined as Parfit supposes. Parfit’s ‘reductionism about significance’ is a non-Humean doctrine; and the reductionist about the facts of personal identity who is also Humean about the motivational status of those facts can quite defensibly deny reductionism about significance. Here we may have the bones of a new sort of argument against the normative and motivational conclusions Parfit draws from reductionism about personal identity, including the claim that identity does not matter.
TOWARDS A PLURALIST ACCOUNT OF SURVIVAL

As we have seen, practice-independent theorists claim that what makes it the case that a certain relation (rather than some other) constitutes the facts of survival is wholly determined, in as much as it is determined, by some single and independent set of facts served up by nature (or God). What makes these relations rather than those count as the relation of survival is wholly independent of any concepts of subjects' person-directed practices and concerns. And it is when, and (for those who take the facts of survival to be exclusive justifiers), only when, the appointed facts served up by nature (or God) obtain that people like us are correct or rationally justified in organizing their person-directed practices around those facts.

In this chapter, I want to provide some reason for reconsidering these meta-assumptions. There is no independent set of facts served up by nature (or God) that demand that we use a certain relation or set of the relations to determine the extension of survival if we are not to miss something important about the way the world is. What criteria we use to determine what determines the extension of survival is mostly up to us; in particular, to the way in which we organize our person-directed practices and concerns.

Before we turn to the case of persons, however, it will be helpful to consider this case from the realm of artefacts.

4.1 The viking ship

Suppose that an ancient viking ship, a country's prized national emblem, sets off on an historic round the world voyage retracing the steps of the first voyage of discovery of the country. At each port of call on this voyage, an old, rotting ship's plank is replaced by a new one exactly like it so that the ship can continue its voyage safely. When the ship finally arrives back at its home port in the capital city of the country for the grand centenary celebrations, not a single one of its original planks remains, each having been gradually replaced, one by one, by a plank of exactly the same sort at each port of call on the journey.
When news of this leaks out, many of the citizens of the country are in uproar. Their prized national emblem, they claim, has been destroyed. The ship that now sits glistening proudly in their harbour is no longer the ship that was their prized national emblem, but a completely new and different ship which happens to look exactly like their old one. Distraught, these citizens claim that they can no longer display the ship at their centenary celebrations, for it is no longer the same ship on which the famed explorers sailed.

Other citizens disagree. It is, they claim, simply wrong to think that the ship in their harbour now is a new, different ship from the ship that set sail a year ago. For, they argue, it looks just the same as the old ship, and is spatio-temporally continuous with the ship that left port. And these things, they argue, and not whether or not the ship is composed of exactly the same token planks, are what are important in determining whether or not the ship is the same ship over time.

But still the debate rages on as the centenary celebrations grow nearer...

After a while, certain reflective peacemakers attempt to intervene in the argument. What, they ask the warring citizens, could make it true that one of them is right and the other wrong? What claim has spatio-temporal continuity or strict identity of the very same token planks, respectively, to constitute the facts of the ship's identity over time, independent of the particular interests and concerns of the people involved? Surely, these citizens argue, whether or not strict plank identity or spatio-temporal continuity get to determine whether or not this is the same ship, depends upon our particular interests in the matter, in particular, on the function the ship plays in our interests and concerns. Given that this particular ship is of great historical value and interest, they argue, it is understandable to think that it may matter a great deal that the ship be composed of the original planks. On the other hand, the ship's planks were replaced gradually (the last original plank being replaced just before the ship entered port) and, in keeping with the historical value of the ship, by planks of exactly the same sort as the original planks; and the explorers themselves replaced many of the ship's planks on their voyages in it. And, in general, our interest in the
identity of ships over time does not seem to require that they retain exactly
the same token planks over time. Neither group of citizens are wrong, these
peacemakers argue, for what privileged, independent facts could there be to
make them wrong? Rather, they just place different amounts of importance
on different aspects of our interests and concerns regarding the ship. They
just care about different things in the ship’s identity over time.

Some may think that the case of survival is very different from the case
of artefacts in this respect. Whereas what determines the identity of artefacts
may well be a conventional matter—a matter of our interests and concerns—
what determines the identity of people over time is not such a conventional
matter. Whether or not a person survives some process is not up to us in the
way in which whether a ship is the same ship, or a watch the same watch, or
something annoying, may be up to us. But, as we shall see, surprising though
it may be, survival is very much like this. Survival is an artefact of our
person-directed practices.

4.2 The Teletransporters and the Somataphiles: A challenge for practice­
independent theorists.¹

Imagine two communities of people just like us: communities made up of
people of our biological kind, Human Beings, living in a world in which there
are no non-material entities such as Cartesian Egos or souls. One community,
being highly technologically advanced, has developed a procedure they call
teletransportation. In teletransportation, a person steps into a booth at one
location where a machine scans the person’s brain and body and beams a
blueprint of it to a receiving booth where another machine makes a cell for
cell exact organic replica of that person, while the original person’s brain and
body is destroyed. The person who emerges from the teletransporter booth
at the receiving end thinks that he or she is the person who entered it, seems
to remember living the life of the person who entered the teletransporter, has
all their intentions, beliefs, desires, and other character traits, and looks the
same. Not only this, but when the teletransporter mechanism constructs the
blueprint, it rids the person of all disease and disability and regenerates their
molecular make-up, so that the person who emerges at the receiving booth is

¹This case is amended from an example of Johnston’s (1989c), esp. pp.454-58.
young, healthy, and able again. From the Teletransporters' point of view, teletransportation provides not only a means of cheap, quick travel, but also, barring unexpected accidents, a means of attaining eternal life.

Being teletransported is as routine an occurrence in the Teletransporter society as boarding a plane or going to the doctor is in ours. It is taken for granted by everyone that teletransportation is survival-preserving: everyone who enters the teletransportation booth has a special and intense interest in what will befall the person who leaves the teletransportation booth at the other end; when their spouse and children are teletransported they continue to love them, be annoyed with them, or whatever such reactive attitudes they had towards them before they were teletransported, and so on. And all the broader person-directed practices of the society presuppose that teletransportation is survival-preserving: the person who emerges from the teletransporter at the receiving end is regarded as owning the property of the person who entered it at the initial end, as being married to their spouse, and as holding whatever offices, rights, obligations etc. the person who entered the booth had. If it is found that the person who entered the booth had committed a crime, everyone regards it as just that the person who emerges from the teletransportation booth should be punished for it. Let us suppose, moreover, that the Teletransporters do not have any false beliefs about the operation of the teletransporter mechanism—they don't believe that it transfers souls or whatever—they simply take themselves to survive the teletransportation procedure.

But now imagine another community, the Somataphiles, just like the society of Teletransporters, except for the fact that the Somataphiles, unlike the Teletransporters, regard the teletransportation procedure as determinately always failing to secure a person's survival. The Somataphiles, having fallen under the sway of certain practice-independent theorists among them, regard the Teletransporters as deeply mistaken, indeed absolutely insane. From the Somataphiles' point of view, the Teletransporters are perfectly prepared to regularly commit suicide without a moment's thought, to periodically murder their children, loved ones, and other members of their society without reason, have a special and intense interest in what will befall people who are mere replicas of them, systematically blame and punish the wrong people for misdemeanours, are continuously deluded about exactly
who had the experiences they seem to remember, are flippantly and superficially prepared to transfer their deep love for their children to new people who, though just like them, are nonetheless distinct from them, and are slowly but systematically facilitating the extinction of their community. Though well aware that the Teletransporters don’t view things in this way, the Somataphiles just take this to illustrate the true depths of the Teletransporters’ delusions. And, when the Teletransporters offer the Somataphiles the teletransporter technology at bargain basement prices, the Somataphiles adamantly turn it down. The Teletransporters, having also been convinced by certain practice-independent theorists among them, shake their heads in utter disbelief. Here the Somataphiles are, generously offered the means of cost-efficient, environmentally friendly, superfast travel and eternal life at bargain basement prices, and they turn it down; and not because they wouldn’t like cheap, superfast travel or eternal life either!

Now are there any principled grounds upon which we can claim that one community is correct and the other mistaken in taking themselves to survive or not to survive the teletransportation procedure? Is there any principled way in which we can say, independently of the person-directed beliefs, attitudes, concerns, and practices of the people involved, that a person survives or fails to survive teletransportation? Practice-independent theorists, of course, answer ‘yes’: if there is a determinate fact of the matter as to whether or not persons survive teletransportation, then it is a fact which obtains or fails to obtain independently of the cognitive and affective responses and practices of the people involved. It is a fact served up by nature. But, as we saw in chapters 1 and 2, practice-independent theorists differ amongst themselves as to precisely what the fact of the matter is. Hence, physical continuity practice-independent theorists (and those psychological continuity practice-independent theorists for whom psychological continuity must have a normal cause) maintain that the Somataphiles were absolutely correct to turn down the offer of the teletransporter technology. For physical continuity is a necessary condition for a person’s survival, and in the teletransportation procedure there is no such continuity. For physical continuity practice-independent theorists, the Teletransporters are fundamentally mistaken. They are simply wrong about the necessary conditions for their survival. For most psychological continuity
practice-independent theorists, in contrast, it is not the Teletransporters, but the Somataphiles, who are deeply mistaken. For psychological continuity, even in the absence of underlying physical continuity, is what is necessary and sufficient for a person’s survival. And since, in teletransportation, there is such psychological continuity, the Somataphiles were deeply mistaken to turn down the offer of the teletransporter technology, and the chance of cheap, super-fast travel and eternal life which accompanied it.

But where are the facts served up by nature to choose between these two different object-level answers to the question of survival? In the absence of a community’s person-directed cognitive and affective responses, how can we really say that either the Somataphiles or the Teletransporters have the wrong concept of survival? For when challenged to explain and justify their person-directed actions, judgements and concerns, both communities point to a relation in the world which, though different, constitutes the relation of survival for them, and which can serve if needs be to justify these actions, judgements, and concerns. The Somataphiles point to a relation in the world, the relation of being a Somataphile survivor (where being a Somataphile survivor requires physical continuity); a relation which is for them the relation of survival, and which marks off the relations in the world that they happen to care about in survival. By the same token, the Teletransporters can point to a relation in the world, the relation of being a series-person (where being a series-person does not require strict physical continuity), a relation which for them constitutes the relation of survival, and which marks off the relations in the world which they happen to care about in survival.

And where are the natural facts to make it true that one community is right and the other wrong to take themselves to survive or not to survive the teletransportation procedure? For both communities are members of our biological kind, *homo sapiens*. The facts of teletransportation are exactly the same in both communities. Both communities are aware of all the relevant facts. Neither community has false beliefs about the facts. Both communities successfully organize their person-directed practices around those facts. Both communities can justify the organization of their person-directed practices in terms of these facts. It is just that the Teletransporters take themselves to survive teletransportation—they track survival with psychological continuity—whilst the Somataphiles do not. The Somataphiles track survival
with physical continuity. If, as nearly all practice-independent theorists nowadays maintain, there is no further fact of survival, and the only facts which are conceptually relevant to determining questions of survival are facts served up by nature, what natural facts could there be in this in case to make one community right and the other wrong?

Nor, in this case, does it seem right to say, as non-Humean, practice-independent theorists such as Parfit or Swinburne would do, that either community is rationally reprehensible to care about different relations; relations which they take to constitute survival? For, ex hypothesi, neither community’s preferences are based on cognitive ignorance or error—they know exactly what teletransportation involves. Nor are their preferences for physical and psychological continuity, respectively, instrumental ones—they do not care about these relations as a means to some other end. They just care about them. Perhaps the Somataphiles care about physical continuity because they just care an awful lot about their body’s unique causal history: that it is the same body that has been with them through trials and tribulations, triumphs and joys, sentimental moments and the like. They just wouldn’t, for example, be specially or intensely interested in what befell some future self with whom they were not physically continuous. Nor would they continue to love a person who was not physically continuous with the person they loved earlier. Nor do they feel it just to punish a person for a crime committed by some earlier person who was not physically continuous with the later person. Physical continuity just matters an awful lot to them in survival. The Teletransporters on the other hand, don’t care much at all about physical continuity in survival. They care about psychological continuity, irrespective of whether or not strict physical continuity is the causal mechanism that furthers their psychology. They would be specially and intensely interested in what befell some future self who was not physically continuous with them, so long as the future person was psychologically continuous with them. And they would continue to love a person who was not physically continuous with them, so long as the person was psychologically continuous.

Nor do either of the communities have preferences for devoting their lives to collecting bits of string, or anything weird like that. Of course, if the Teletransporters had the Somataphiles’ beliefs about survival—if they
believed that survival required physical continuity—then there might well, for the non-Humean, be grounds for calling them irrational. For then we would be imagining a society who, whilst fully acknowledging that teletransportation resulted in a person’s certain death, nonetheless persisted in allowing themselves and their loved ones to be teletransported, and in regarding teletransportation as a convenient means of travel. But neither the Teletransporters or the Somataphiles are such a society: the Teletransporters believe that they survive teletransportation, for psychological continuity is what survival involves; and the Somataphiles turn down the offer of teletransportation precisely because they believe that survival involves physical continuity and, in teletransportation, there is no such continuity. They both desire to survive (by their own lights), and their cognitive and reactive attitudes to people and survival are just like our own, were we to have—as, indeed, many amongst us may well have—their respective beliefs about survival. In virtue of what, then, could their respective concerns be irrational? If there are no further metaphysical facts to make one community right and the other wrong to take themselves to survive or not to survive the teletransportation procedure, what facts could there be to make one community right and the other rationally reprehensible to care deeply and non-instrumentally, as they do, about the different relations in survival?

Nor would it seem right simply to reply, as Unger and other theorists of his ilk would do, that the Teletransporters and the Somataphiles simply, though (for those who regard the facts of survival as inclusive justifiers) quite acceptably2, mean something different by survival. For the first part, the two communities clearly have something important in common. For, although they differ about exactly what survival involves, we (and they) nonetheless know exactly what it is that they disagree about, and moreover, know that they are disagreeing about the same thing: namely, survival, and not where is the best place to go for a holiday, or what sort of food is nice to eat. What they have in common is the role that survival plays in the vast and rich array

2Some rigidifying practice-independent theorists may be less ecumenical than Unger. According to these practice-independent theorists, the two communities mean something different by personal identity (in virtue of the fact that different properties play the survival role in their person-directed practices and concerns), but only one community means the right thing. For these theorists, the challenge simply reappears. For what could make it non-relatively true that one community rather than the other mean the right thing by survival?
of their person-directed practices. Indeed, it is this common role which gives point to their debate—which explains why they care one way or another about which facts are the facts of survival—in the first place. To say simply that the Teletransporters and the Somataphiles mean different things by survival is to deny this obvious and significant commonality. As Sydney Shoemaker writes of the similar case discussed in chapter 3 of the society who organize their person-directed practices around the assumption that the brain-state transfer procedure is survival-preserving, there is

a strong reason for saying that what they mean by personal identity is what we mean by it; they call the same things persons, offer the same sorts of characterizations of what persons are, and attach the same kinds of social consequences to judgements of personal identity—i.e., personal identity has with them the same connections with moral responsibility, ownership of property, etc., as it does with us.3

For the second part, the debate between the Teletransporters and the Somataphiles is just the personal identity debate between physical continuity theorists and psychological continuity theorists in this world writ large and practically significant. And few, if any, in the personal identity debate have been happy to grant that Parfit and Williams, for example, just mean different things by 'survival'. Here, again, the obvious explanation for this is that, although the realizer of survival for Parfit and Williams is different (in the former case, psychological continuity, in the latter case, physical continuity), the role that survival plays in their person-directed practices and concerns is the same; and it is the role, not the actual realizer, of survival that is its central defining feature. For, although survival has a different extension in their respective mouths, both Parfit and Williams, like the Teletransporters and the Somataphiles, have a special and intense interest in their own future selves and what befalls them; expect to remember directly their own experiences; ascribe to themselves and others beliefs, actions, thoughts, feelings, physical traits, certain ongoing rights and obligations, and so on. So the concept of survival plays the same role in their person-directed practices. It is just that, as a contingent matter of fact, they take different relations to instantiate that role.

If we take 'survival' to rigidly designate the relation that actually realizes the survival role for a community, we cannot explain why we think, as we do, that the two communities (like Parfit and Williams) have something in common, and that the disagreement between them is genuine disagreement about survival; disagreement, moreover, in which much is at stake. For, if we identity survival with actual realizer (physical continuity on the one hand, psychological continuity on the other) not only can we not explain what the two communities have in common, but moreover, it turns out that there is no point to their disagreement after all. For they simply mean different things. (Here we might well feel the same disquiet as we do for the meta-ethical relativist's claim that a community who happily assert that 'wanton murder is right' and a community who vehemently assert that 'wanton murder is not right', might nonetheless yet both be right, for, after all, they simply mean different things by 'right'). We need an account of survival which allows us to give voice to what the two communities (like Parfit and Williams) clearly have in common, as well as about what they differ. Where they differ, is with respect to what relation they take to play the survival role. What they have in common, as we might put it, is that they both care about survival, but where, and only where, 'survival' rigidly designates, not actual realizer (for then they do not both care about survival, one cares about survival*), but the role that physical continuity and psychological continuity, respectively, play in their person-directed practices and concerns.

For the third part, to insist, as Unger and other such rigidifying theorists do, that all people everywhere should track survival with the particular relation that we (or a limited group of us) now happen to track it with is unwarrantedly chauvinistic. (It is a lot akin to insisting that, no matter how apparently sophisticated and seemingly 'intelligent' the behaviour of other, differently neurophysiologically constituted creatures, only creatures with our particular carbon-based, neurophysiological make-up could have minds.) For what is so important and interesting about the particular relation with which we now happen to identify survival? What is so important about that relation that demands that all people, at all times, in all worlds (sufficiently similar to ours), organize their person-directed practices and concerns around the particular relation that we happen to organize ours now,
if they are to mean what we mean by survival? What reason have we to be so
dogmatically parochial about survival? Unger’s answer, of course, is that that
relation is significant because that is the relation in which we now most
deeply believe survival to consist; that is the relation which presently
organizes our deepest beliefs, concerns and values. But, granting that we
now believe in and care deeply about some particular relation, could we not
have cared deeply—might we not even actually come to care—about some
different relation?

As the Somataphiles most deeply believe, survival necessarily requires
physical continuity, and so they spurn teletransportation. For them,
teletransportation is a means only to a quick death. But suppose that, having
watched the teletransporters happily teletransporting themselves quickly and
efficiently from place to place, in the process ridding themselves of otherwise
incurable disease, they come to think that teletransportation carries with it
considerable advantages: quick, easy, cost-effective, environmentally-friendly
travel (with no jet lag), practical freedom from disease and eternal life. And
they come to wish that teletransportation could provide such advantages for
them too. It is not inconceivable that this higher-order desire might transfer
to a corresponding change in their brute concerns. After all, no matter how
‘deep’ a brute belief or desire, it is rarely immutable. (Ask many a reformed
smoker!) And so it is by no means impossible that, in the course of time, they
come rationally to believe that they, like the teletransporters, will survive
teletransportation; and proceed to organize their person-directed practices
around that belief.

But perhaps they were wrong to so change their person-directed
beliefs and concerns? Perhaps there was good reason to spurn
teletransportation, despite the considerable benefits it brought once they
came to believe they survived the process? Perhaps ‘survival’ is or is (in
Wiggins’s words) “akin to” a natural kind term; and physical continuity the
privileged natural kind? Perhaps, in so changing their person-directed
practices, the Somataphiles are now failing to take account of something
intrinsically very important and interesting about the way the world is?
Perhaps they are now missing a natural joint?
Chapter 4  Towards a pluralist account of survival

There are, I think, good reasons for thinking that 'survival' is not such a natural kind term; for thinking that there is no natural joint in the world which serves up survival independently of our interests and concerns, but rather a whole host of disparate competing natural candidate properties or relations, none of which, in the absence of our person-directed practices, have any privileged claim to be the relation of survival. Personal identity⁴ is, as Locke and others have stressed, not a natural, but a social, concept. For, in the case of personal identity, unlike the case of water or gold or atomic number, the phenomena to be explained are social not natural: they are the vast and rich array of person-directed practices. Indicative of the social nature of personal identity is the fact that, although we happily defer to scientists to tell us about what water is, or what gold is, or what atomic weight is, we do not so defer to scientists to tell us about who is the same person. Scientists might tell us some brute natural facts about what counts as 'clinically' dead, or what are the survival conditions for homo sapiens, but not about personal identity: not about who should be held responsible for some past crime, who should be attributed certain ongoing rights, entitlements and obligations, whose future pains I should anticipate, or whose past actions I should self-ascribe.

However, the critical point to note here is that, even if 'survival' were (as I think not) a natural kind term, then this supports a practice-dependent account of survival as well. Here's why. On the standard account of what a natural kind is, it is the grouping of things, or natures of things, that figures in the best scientific explanation of some phenomenon, that is, the scientific explanation with the best explanatory power and completeness. So, for example, water is a natural kind (H₂O) because H₂O is the grouping which best explains why water boils at one hundred degrees Celsius, freezes at zero degrees, rusts certain metals etc. What, then, is the phenomenon to be explained in the case of survival? Our person-directed practices. Consider, first, the Teletransporters. What is the natural grouping that would figure in the best explanation of the person-directed practices of the Teletransporters, that is, the explanation with the best explanatory power and completeness?

⁴In this context, I am especially unhappy about being forced to use the word 'survival', instead of 'personal identity'. For the term 'survival', unlike the term 'personal identity' carries with it overwhelmingly natural, not social, connotations—'survival of the fittest', for example. So, since nothing hinges on it here, for this paragraph I shall simply use the term 'personal identity'.
Answer: psychological continuity. Psychological continuity is the grouping which best explains why the teletransporters regard it as true that a person survives teletransportation; why they regard it as just that the person who emerges from the teletransporter should be attributed whatever ongoing rights, obligations and entitlements the person who entered the teletransporter had; why they are specially concerned for what will befall the person who emerges from the teletransporter; why they regard it as a means of quick, cost-effective travel, rather than death; and so on. Other groupings, such as physical continuity, cannot explain this well or completely. Indeed, positing physical continuity as the explanatory undergirdings of these practices, makes the teletransporters appear as a society of mad people. Consider, however, the Somataphiles. In this case, physical continuity, not psychological continuity, is the grouping which best explains their person-directed practices. Physical continuity is the grouping that best explains why the Somataphiles initially turn down the offer of the teletransportation technology; why they regard is not as a means of quick, cost-effective, environmentally-friendly travel, but as a means of quick death; why they regard it as entirely unjust that new people should inherit the rights, obligations and entitlements of old, dead people; and so on. For the case of the Somataphiles, psychological continuity does a very bad job of explaining their person-directed practices. Indeed, again, this grouping would make their person-directed practices close to unintelligible. A view of survival as a natural kind concept, then, far from being at odds with pluralism, actually supports it. For as the case of the Somataphiles and the Teletransporters illustrates, different natural groupings are required to figure in the best causal explanation of the respective communities' person-directed practices. This, I take it, is what Mark Johnston has in mind when he somewhat obliquely writes

...we need not oppose whatever is true in the doctrine that ‘human person’ is or is “akin to” a natural kind term. The import of...relativism...is that there may be several more or less natural kinds available to associate with a term like ‘human person’. Then our patterns of person-directed concern and expectation will be crucial.6

5Sydney Shoemaker (1984) regards psychological continuity as a functional notion. For him, the natural grouping would be whatever more basic natural grouping underpinned psychological continuity.
But this will not yet completely answer the natural kinds’ theorist who thinks that 'survival' rigidly names the particular kind that it names in the actual world. For, after all, we might say the exactly the same thing for water. On Earth H₂O is the natural grouping which best explains why water boils at one hundred degrees Celsius, freezes at zero degrees, rusts certain metals etc., but on Twin Earth XYZ is the nature of the substance that best explains these phenomena. Those who believe that 'water' rigidly names the natural kind that it names in the actual world can grant all of this. H₂O is the natural kind to which exemplars of water belong here on earth, but XYZ is the natural kind to which they belong on twin earth. But, they will say, ‘water’ nonetheless means something different in the mouths of people on Earth from what it means in the mouths of people on Twin Earth. It means something different in virtue of the fact that, although the role water plays is the same in both worlds, the chemical configuration which people on Twin Earth use the word ‘water’ to refer to is not the chemical configuration that ‘water’ names in the actual world. And since ‘water’ rigidly names the kind that it names in the actual world, despite the fact that XYZ plays exactly the same role on Twin Earth as H₂O plays on Earth, people on Twin Earth are talking about something else.

There is no doubt that the content of their utterances (like those of the Teletransporters and the Somataphiles) is different; and taking a term rigidly to designate its actual realizer, rather than role, is, in itself, a perfectly acceptable theory of reference. But which, if either, of these alternative analyses of meaning—rigidifying on actual realizer or rigidifying on role—is right or appropriate, will vary from term to term. No one, for example, even those who think that rigidifying on the actual realizer is right for the term ‘water’, thinks that such a rigidifying account is also right for terms like ‘food’. For what makes rigidifying on actual realizer attractive in the case of ‘water’ is notably absent in the case of ‘food’. Just so, I think, what makes this rigidifying analysis plausible in the case of ‘water’, is famously and notably absent in the case of ‘survival’ as well.

A large part of what makes the rigidifying account attractive in the case of water is that exemplars of water have a leading candidate—H₂O—to be the natural kind to which they all belong. This is what makes it plausible (for those who think that it is plausible) to think that ‘water’ is a natural kind
term which rigidly names the kind that it names in the actual world. Famously, however, in the case of 'survival', after centuries of searching by very smart philosophers of all ilks we still have not found anything remotely similarly resembling a leading candidate. Rather, there are a whole host of disparate and competing candidates in the natural world: bodily continuity, continuity of the brain, psychological continuity, psychological continuity with continuity of the human brain, psychological continuity without strict physical continuity, psychological continuity with a reliable cause...etc., and perhaps others yet uncanvassed.

The situation in the case of survival is rather like the following situation in the case of water. Imagine that the stuff which actually filled the oceans, fell from the sky, ran from taps, was colourless, odourless, good for drinking etc. was a mixed, disjunctive bag of natural properties, \( \text{H}_2\text{O} \) or \( \text{XYZ} \) or QRS or ABC and numerous other chemical properties regularly being newly discovered. If this were the case, it would be far less plausible, indeed extremely implausible, to maintain that 'water' names in all worlds the messy, non-principled disjunction of properties that play the role in the actual world. For, in this case, 'water' names no interesting principled natural joint, but a gerrymandered disjunction. In this case, then, I think, we would not say that 'water' named the long, messy, disparate disjunction of \( \text{H}_2\text{O} \) or \( \text{XYZ} \) or QRS or ABC... in all worlds. Nor, I think, would we embrace an error theory. We would say, rather, that 'water' names whatever stuff it is that plays the water role. We would give up on the idea that 'water' was a natural kind term rigidly to be identified with the long, messy disjunction of properties which actually played the water role, and instead treat 'water' as a functional term, in just the same way as we presently treat terms like 'food'. When asked to define 'food', for example, we do not say that food is apples, or oranges, or sausage rolls, or lettuces, or roast chicken, or biscuits, or bread, and the rest of the very long disjunction of materials that beings are disposed to eat when hungry to sustain themselves. Nor, for shorthand, do we even say that food is the carbon-based stuff which realizes the food role in the actual world. For go to a world in which beings who are differently physiologically constructed are disposed to eat silicon-based material when hungry to sustain themselves, and that material surely counts as food just as well. 'Food' rigidly refers, not to the messy, gerrymandered bunch of actual and possible, natural (and non-
natural!) properties that play the food role, but to the common role that this messy group of stuffs play. The more disparate, disjunctive and seemingly gerrymandered the natural properties to which a term refers in the actual world, the less plausible it is to think that the term rigidly names in all worlds the properties that it names in the actual world.

Just so, the fact that a significantly large, messy disjunction of properties (including newly discovered ones, such as Unger’s) play the survival role in the actual world makes it far less plausible, indeed extremely implausible, to continue to think that ‘survival’ rigidly names “some scientifically palpable real essence, some nature that is presupposed to a kind’s being nomologically founded, which not only underlies in fact but also must underlie the kind person [and persons’ survival over time]”⁷. Given the absence of a leading candidate to be the natural kind to which exemplars of survival in the actual world all belong, it is much more plausible to think that the term ‘survival’, like ‘food’, names the functional role that the disjuncts play in our person-directed practices and concerns, not the messy, gerrymandered bunch of relations that actually play the role for us.

However, all this may still not yet convince some particularly die-hard rigidifiers. In desperation, these theorists might maintain that the apparent absence of a leading candidate in the case of survival is just that, merely apparent. According to these philosophers, the fact that we have not yet uncovered the leading candidate to be the natural kind to which exemplars of survival belong is not because there is no such leading candidate to be uncovered, it is because we haven’t yet been sufficiently clever to discover it. We have not yet discovered the hidden natural unity beneath the facade of apparent disorder. To return to our scientific analogy, it is because we have not yet sufficiently fine-tuned our intuitive microscopes.

But this reply really is clutching at straws. For we know all about what the phenomena to be explained are. And we know all about physical continuity, psychological continuity and the various other competing natural candidates; candidates which, as we have seen, do an excellent job of explaining the phenomena. What reason, then, to posit some mystical hidden

---

natural essence (a haecceity or some such thing) which all these exemplars share? To insist on positing such a hidden essence is to render natural kinds utterly mysterious. As the case of Teletransporters and the Somataphiles illustrates, such unity as these exemplars of survival have, surely derives from the common role they play in our person-directed practices. Just as functional architecture not realizer is what is interesting as regards minds, so it is role not realizer that is the interesting and significant semantic feature in the case of 'survival'.

4.3 A different presentation of the challenge?

Mark Johnston offers the case of the Teletransporters and the Human Beings (my Somataphiles) in similarly arguing for a sort of practice-dependent account of survival (exactly what sort, we will discuss in chapter 6). However, Johnston makes different use of this case in presenting the challenge to practice-independent theorists. There is a reason why I have chosen not to follow Johnston here.

Johnston’s argument runs as follows. As nearly all practice-independent theorists (Johnston calls them 'absolutists') nowadays will happily grant, any plausible candidate relation for survival (i.e. physical and/or psychological continuity and connectedness) will admit of indeterminate cases: cases where, even though we know everything there is to know about the facts of some process p, the question ‘does a person survive process, p?’ has no determinate ‘yes’ or ‘no’ answer. Suppose we regard such vague cases as standard cases of semantic indeterminacy: where the indeterminacy is then to be explained, not by vagueness in the world, or lack of knowledge about the facts of the case, but by vagueness in the term ‘survives’. Suppose that there are ten different but acceptable ways of precisifying or sharpening the vague predicate ‘survives’: survives1, survives2, survives3...survives10. According to the super-valuationist analysis

---

8 Cartesian Dualists will not, of course, grant this premiss. For Cartesians, personal identity consists in the continued existence of the same soul, and the continued existence of the soul does not admit of degree or indeterminacy. There are many and famous independent arguments against the Cartesian view and few practice-independent theorists nowadays are explicitly Cartesian. In any case, the view I shall advocate is perfectly compatible with survival (contingently) consisting in the continued existence of the same soul, in a world where there are such immaterial entities. My disagreement is with practice-independence, the meta-theory, not with object-level views.
of indeterminacy, the statement ‘X survives p’ is super- or determinately true just in case, on all ten precisifications of ‘survives’, the predicate ‘survives’ applies; the statement ‘X survives p’ is determinately or super-false, just in case none of the ten precisifications of the predicate ‘survives’ apply; and the statement ‘X survives p’ is indeterminate just in case some of the precisifications of ‘survives’ (say, survives₁ to survives₅) apply, and some of the precisifications of the vague term ‘survives’ (say survives₆ to survives₁₀) do not.

It is, Johnston claims, if not actually the case, then certainly epistemically possible, that teletransportation be such an indeterminate case or process relative to our (or some) concept of survival i.e. a case where the concept of survival or survival neither determinately applies nor determinately fails to apply. So we can imagine a community, the ur-community, according to whose concept of survival teletransportation counts as an indeterminate case. Assuming a super-valuationist analysis of such semantic indeterminacy, there will be some acceptable ways of precisifying the term ‘survives’ (survives₁ to survives₅) according to which it comes out determinately true that a person survives teletransportation, and other acceptable ways of precisifying the term ‘survives’ (say, survives₆ to survives₁₀) according to which it will be determinately false that a person survives teletransportation. So we can imagine two communities who have simply internalized two different sharpenings of the ur-community’s vague concept, both of which, from the point of view of the ur-community, are equally acceptable ways of going. One community, the Teletransporters, have internalized a sharpening (say, survives₁) according to which a person determinately survives teletransportation. The other community, the Human Beings, have internalized a different, though equally acceptable, sharpening of the concept, (say, survives₆) according to which a person determinately fails to survive teletransportation. (Super-valuation then gives us a way of charting the range of acceptable sharpenings of the concept, the range, that is, of acceptable alternative concepts of survival.⁹)

---
⁹ This is one way of understanding what Johnston’s ‘Protean person’ is, discussed in more detail in chapter 6. The substance sortal Protean person is the sum of actual and possible cryptophase kinds, the boundaries of which are set by the range of possible acceptable sharpenings of the concept of personal identity.
As most practice-independent theorists should grant, given the indeterminacy and a super-valuationist view thereof, either precisification is an equally acceptable sharpening. So the practice-independent theorist must admit that the two communities have equally acceptable concepts of survival. As Johnston writes,

What most badly damages the absolutist cause is the historical discovery that the Teletransporters and the Human Beings had simply internalized different stipulations about Teletransportation, stipulations which from the point of view of the ur-community and its concept of personal identity were equally acceptable ways of going—equally acceptable precisifications. The Human Beings find it hard to believe that the accidents of conceptual history happened decisively to favour only them. And when they try, on the assumption that they are victims of conceptual history and have an erroneous concept of personal identity, they are at a loss as to how to discover what the error is. But then they see that the Teletransporters are in a parallel situation—plausibly taken as neither correct by absolutist standards nor incorrect by weaker standards. So certain relativists urge that the Teletransporters are just blamelessly different...\(^\text{10}\)

But exactly how is this supposed to be an argument against practice-independent theorists or, as Johnston calls them (rather misleadingly, in my view, since it suggests that the alternative view is relativism), 'absolutists'? How is the fact that their precise concept of survival can be reconstrued as just one of a number of possible, acceptable sharpenings of a vague concept supposed to undermine their commitment to their precise concept? Johnston’s answer is that it shows that their particular, precise concept is completely arbitrary. They might just as well, but for a twist of conceptual fate, have internalized some other sharpening. And, recognizing that there but for the grace of a conceptual god go they, ought to conclude first, that people who internalize a different sharpening are just blamelessly different, and second, that they have no good reason (at least, no good reason stemming from considerations of survival) not to go that way themselves.

However, it is hard to see how, as it stands, this argument will carry much force with practice-independent theorists. For the first part, note that we could run exactly the same strategic argument for any of our concepts. Are we then to conclude that all of our concepts are arbitrary, so that we have

\(^{10}\text{Johnston (1989c), p. 456.}\)
no good reason to be committed to any of them? For example, we could tell the following story about gold. Imagine a community, the ur-community, according to whose concept of gold, what we call 'fools gold' counts as an indeterminate case: the predicate 'is gold' neither determinately applies, nor determinately fails to apply to the stuff we call 'fools gold'. So we can imagine two communities, A and B, living in worlds in all physical respects just like our own, who have simply internalized different sharpenings of 'gold', both of which, from the point of view of the ur-community, are equally acceptable ways of going—equally acceptable precisifications. According to community A, the predicate 'is gold' determinately applies to the stuff we call 'fools gold'; according to community B the predicate 'is gold' determinately fails to apply to the stuff we call 'fools gold'. Given the indeterminacy, either precisification is equally acceptable, so we ought to conclude that the two communities are just blamelessly different—they have equally acceptable conceptual schemes. That is to say, we ought to conclude that practice-dependence is the right meta-account for gold. We could tell a similar story for atomic number too. Imagine a community, the ur-community, who are organizing the periodic table. They are organizing it in accordance with 'atomic number', however their concept of 'atomic number' is vague between what we call 'atomic number' and 'atomic weight'. So we can imagine two communities, living in worlds just like our own, who have simply internalized two different sharpenings of 'atomic number'—sharpenings which, from the point of view of the ur-community and their vague concept, are equally acceptable precisifications. One community, community A, have internalized a sharpening according to which 'atomic number' determinately means 'number of protons'. According to the other community, community B, 'atomic number' determinately means 'number of protons and neutrons' (our 'atomic weight'). Given the indeterminacy, either precisification is an equally acceptable way of going, so we ought to conclude that it is just as right for chemistry to organize the periodic table according to atomic weight. And that, moreover, we have no good reason for not organizing our periodic table according to atomic weight. That is to say, we ought to conclude that the atomic number of atoms is practice-dependent. And, having seen how the strategy goes, we can imagine many more similar examples (reductios!).
For the second part, Johnston’s argument will carry no weight with those practice-independent theorists who regard the our concept of survival as essentially vague with respect to teletransportation. These theorists will grant Johnston’s premiss—they will grant that given the indeterminacy there are a range of possible acceptable sharpenings—but they will deny that the concept of survival to which any one of these sharpenings gives rise are equally acceptable concepts of survival. For, the concept of survival is vague, not precise; and even though there are equally acceptable sharpenings, it is not true that any particular sharpening is acceptable.

For the third part, as it stands, Johnston’s argument will not convince even those practice-independent theorists who think that the concept of survival is precise. For these theorists, as we have already seen, will just deny that it is arbitrary and unfounded to go one way rather than the other. ‘Real essence’ theorists, for example, will claim that we have good reason to employ the precise concept of survival that the theorist in question claims we do—it is not true that either precisification, employed as a concrete concept, is equally acceptable. For this precisification, and not the other, latches onto the intrinsically interesting and important facts of survival in the world. For these theorists, recall, the semantics of survival tracks (or ought to track) the independently determined metaphysics of the matter. From the point of view of these theorists, Johnston’s argument will be akin to saying that we might just have well have employed a conceptual scheme that grouped elements in accordance with their atomic weight rather than atomic number. And, to group things in this way, they argue, would be to miss something very intrinsically important and interesting about the way the world is. We need a different sort of argument to show such ‘real essence’ theorists that, in the case of survival, there is no such intrinsically interesting and important natural kind; argument which, I hope, I have already provided.

Nor will Johnston’s argument carry much weight with rigidifying theorists, such as Unger. For, as we have already seen, these theorists will reply that our precise concept is far from arbitrary, for that sharpening—in Unger’s case, physical continuity—and not some other possible sharpening, captures the relation that we care about. That sharpening, and not some other possible sharpening, is the relation that organizes our deepest beliefs, concerns and values. That is the principled reason we have for internalizing
that sharpening rather some other, and that is what makes our commitment
to that sharpening entirely non-arbitrary. Moreover, a practice-dependent
theorist such as Johnston himself, ought to find this a pretty compelling line
of reply.

The objection to practice-independent accounts of survival in all their
various possible object-level forms, is not that they are arbitrary. For they are
not. The Somataphiles have pretty compelling reason, by their present lights,
for not going the way of the Teletransporters, for teletransportation fails to
preserve the relation that they care about in survival. Likewise, by their
lights, the Teletransporters have good reason for not going the way of the
Somataphiles, for teletransportation does preserve everything that they care
about in survival. The objection to practice-independent accounts of survival
is not, then that they arbitrary, but rather that, in claiming that 'survival'
rigidly designates actual realizer, rather than role, they misconstrue the
nature of survival. 'Survival' rigidly names not a natural, but a social or
functional kind—a kind determined by subjects' person-directed practices
and concerns, not by facts served up by nature.

In any case, as we have seen, the presentation of the challenge to
practice-independent theorists need not, and, I think, should not, go
Johnston's way. Some might think this comes at a cost. For Johnston's
presentation, it might be thought, has an advantage over mine in that, in
going by way of indeterminacy, something many practice-independent
theorists are happy to grant, it grants more up front to practice-independent
theorists and so may prove more congenial. However, Johnston's
presentation relies on an argument that will, I suspect, convince very few
practice-independent theorists. And, indeed, my presentation of the
challenge shares a widely-held practice-independent premiss too—namely,
that it is a requirement on an adequate account of survival that it explain and
render intelligible our person-directed practices and concerns. As the case of
the Teletransporters and the Somataphiles illustrates, practice-independent
accounts of survival, which rigidly identify 'survival' with actual realizer, not
role, fall foul of this requirement.
4.4 The lesson of the Teletransporters and the Somataphiles

To sum up. The case of the Teletransporters and the Somataphiles seems to me to vividly illustrate two, related things. The first is that, and as against the ‘no-further-question’ view, there is indeed the following (meta-) question to be asked, and indeed, answered, for the case of survival: in virtue of what does some relation count as the relation of survival? The case of the Teletransporters and the Somataphiles gives us good reason to doubt that practice-independence is the right answer to this question. For facts served up by nature leave the concept of survival crucially underdetermined. There are just too many relations served up by nature which are candidates for being the relation of survival. And, in the absence of a community’s person-directed practices, attitudes, beliefs, and concerns, there just seems no principled way of choosing among them.

But the case of the Teletransporters and the Somataphiles also vividly illustrates how, on practice-independent analyses, the concept of survival can come apart from what, inter alia, a community cares about in survival. Physical continuity practice-independent theorists, for example, insist that since a person’s survival necessarily involves physical continuity a community like the Teletransporters are simply mistaken to believe that they survive teletransportation and to organize their person-directed practices and concerns around this false belief. But the Teletransporters happen not to care about physical continuity. From their point of view, if survival requires physical continuity, then survival is not something they care about. If that’s what survival necessarily involves, they are indifferent as to whether they, their loved ones and other members of their community live or die.

But this seems fundamentally wrong. The intuition that whatever survival is, it is something that we care about, seems a particularly central one. (Witness the many and varied attempts to save it in the face of Parfit’s perceived challenge). The thought that we might be indifferent as to whether we live or die—that the concept of survival might be utterly inert as regards our person-directed practices and concerns—seems absolutely absurd. To maintain, as practice-independent theorists do, that the Teletransporters or the Somataphiles have the wrong concept of survival, or mean something
different by the survival, or do not have a concept of survival at all, in virtue of the fact that the relation which they care deeply about and take to constitute the concept of survival is not whatever relation it is that the particular practice-independent theorist in question take to be the relation of survival, is to make survival just such a potentially motivationally anaemic and inert notion. And whatever survival is, it is certainly neither anaemic nor inert. It plays a rich an integral role in our person-directed practices and concerns. An account of survival which so leaves it open that survival might turn out to consist in some relation that made us indifferent as to whether we, our loved ones, and other members of our community continued to live or die, would not only, as practice-independent theorists themselves admit, be inadequate, it would not be an account of the relation of survival. It would be an account of some other relation. (That is why the thought that physical continuity might count as the relation of survival for the Teletransporters strikes us as completely wrongheaded). And the only way to ensure that survival turns out to be something that a community cares about in the way which is definitive of our caring about survival is to incorporate our special concern for survival, and the rich role it plays in our practices, as an essential feature of the concept, and not something merely accidentally or contingently associated with it.

It is to developing such an account that I now turn.
A PLURALIST ACCOUNT OF SURVIVAL

The central lesson of the Teletransporters and the Somataphiles is that practice-dependence is the right meta-account of survival. There is no meta-fact served up by nature or God that independently demands that one socially enactable relation rather than another must determine the extension of survival, if we are not to fail to latch onto something intrinsically important about the way the world is. Which socially workable relation we use to determine our survival conditions is up to us, in particular, to the way in which we organize our person-directed practices and concerns. Relations in the world earn their right to count as the relation of survival for members of a community in virtue of the fact that those are the relations around which members of the community organize their person-directed practices and concerns. And since different communities might, neither in metaphysical nor motivational error, organize their person-directed practices and concerns around different relations, survival is multiply realizable. Survival for a Teletransporter-person involves psychological continuity, whilst survival for a member of the Somataphile community requires physical continuity. We have no reason, then, metaphysical or rational, to be chauvinistic about survival, in the way in which practice-independent theories of survival are chauvinistic.1

What impact will the truth of practice-dependence have on our object-level account of survival? What views about the properties of person-stages, in virtue of which different person-stages count as stages of a surviving person, will a practice-dependent meta-view yield? As we saw earlier, when discussing the case of value concepts in chapter 1, the practice-dependent theorist has two broad options at the object-level. Recall, the question then was whether or not the practice-dependent theorist (who thinks that what makes some properties of acts count as good depends on whether or not subjects in appropriate conditions are disposed to approve of them) will agree with the utilitarian about the

1There is, of course, a sense in which a practice-dependent account of survival is chauvinistic. It takes the reference of survival to be fixed by the role that survival plays for us. But we are, after all, concerned to analyse our concept of survival.
answer to the object-level question of what properties of acts are good, in circumstances where it turns out that all and only those acts which subjects are disposed to approve of are those which have the property of maximising happiness. Our answer then was that they may or may not, depending on which one of two more precise object-level views the practice-dependent theorist chooses to take: either they may (contingently) identify good with the particular property that the particular group of subjects in ideal conditions happen to approve of (in this case, maximising happiness), or they may identify it with a second-order property viz. the property of being a property that subjects in ideal conditions approve of. If the practice-dependent theorist takes the first, first-order, option, they will agree with the utilitarian about what property it is in virtue of which acts count as good (although, of course, they will disagree about the modal status of the property: utilitarians will identify good with maximising happiness necessarily, whereas practice-dependent theorists will identify good with the property of maximizing happiness only contingently). If, however, the practice-dependent theorist takes the second option, and views good as the second-order property, they will disagree completely with the utilitarian about the answer to this object-level question, as well as about the answer to the meta-question.

Practice-dependence about survival—the meta-view that what makes some relation count as the relation of survival is, not facts served up by nature, but rather that those are the facts around which subjects organize their person-directed practices—is itself agnostic as between these two sorts of object-level views. It tells us that two person-stages earn their right to count as stages of the same surviving person only in virtue of the fact that some relation which plays the survival role holds between those two stages. But practice-dependence does not tell us whether that relation is survival, or whether survival is a property of that relation. Practice-dependence tells us that survival is whatever relation between person-stages plays the survival role in the person-directed practices of the community involved, but it does not tell us whether that community’s survival conditions—the truth conditions for survival claims—are given by the relation that organizes subjects’ person-directed practices or by a higher-order property which that relation has viz., it being a relation which has the property of realizing the survival role.
I am more certain that practice-dependence is right for the case of survival than I am of exactly which of these two corresponding object-level analyses is right. Nonetheless, as I see it, there is at least some reason for preferring the first sort of view to the second. Before discussing these in more detail, however, let me sketch the two options.

5.1 Survival as a first-order property

If we view survival as a first-order property, as I shall advocate we do, we shall offer something like the following as an analysis of survival:

'A person-stage X at an earlier time survives as a person-stage Y at a later time' as uttered by us\(^2\) is true iff

(i) there is some property or set of properties \(p\) which X consistently organizes X's person-directed practices around, and

(ii) X and Y share \(p\).

(i) is a conceptual claim about the conceptual dependence and interdependence of the concept of personal identity on concepts of subjects' person-directed practices and concerns. And, as it stands, it is non-reductive. But, if we are reductively minded, we also want to know what property, as a matter of fact, \(p\) is. So add to our conceptual claim, the second substantive claim, (ii), that X and Y share \(p\). Though the conceptual claim is non-reductive, then, the substantive claim means that the analysis as a whole is thoroughly reductive. *Survival is whatever property it is that causes subjects, consistently, and without involvement in empirical error, to organize their person-directed practices around the assumption that X survives as Y.* Whereas (i) enshrines the lesson of the Teletransporters and the Somataphiles in the form of a conceptual claim—that survival is that relation, whatever it is, which plays the survival role in the person-directed practices and concerns of the community involved—(ii) is an *a posteriori* claim about what relation it is which in fact realizes that role for a given subject or community.

A first point to note about this analysis, then, is that (like the accounts of water and mental states I considered in Chapter 2) it proceeds in two stages. The first stage consists in *a priori* conceptual analysis to fix

---

\(^2\)Where 'us' refers to subjects or groups of subjects who share common person-directed practices. So 'A person survives teletransportation' is true in the mouths of the Teletransporters, demarcated by their common person-directed practices, but false in the mouths of the Somataphiles, defined in terms of their shared person-directed practices.
the role that the concept of survival plays in our network of person-directed practices; to tease out the conceptual connections between the concept of survival, and concepts of punishment, self-concern, self-ascription, identification and reidentification, moral responsibility, legal entitlement, memory, intention, reactive attitudes such as continuing love for a particular person over time, and so on. Some of these principles we may be able to make explicit just by reflecting on our everyday person-directed practices and the actual cases which hone them and noting the connections among them. But, more often than not, these reflections will involve thinking about what we would say of certain counter-factual situations. Here, thought experiments will have an important role to play. If we want to find out about the conceptual connections between punishment and survival, for example, (to simplify matters for the moment, supposing that we are deontologists about punishment) we might pose ourselves the following sort of thought experiment. Imagine two identical twins, X and Y, who are extraordinarily similar in physical and psychological make-up and background. X commits a crime, but manages to elude capture. Y is captured in his place. Should Y be punished for the crime? Intuitively, we answer ‘no’. This tells us something about the connections between our concepts of survival and punishment: it tells us that, ceterus paribus, we do not believe that a person who is not the same person as the person who committed a past crime should be punished for it. If we further ask ourselves—should X be punished for the crime when captured?—and we answer ‘yes’, then this tells us that we think a person should be punished for a crime only when, ceterus paribus, they are the same person as the person who committed the crime. (And we can design further thought experiments to tell us about what the ceterus paribus clause includes. For example, suppose that an evil neuro-surgeon had wired up X to a machine which caused X to commit the crime against his will, should X then be held responsible and punished for the crime?). Similar sorts of thought experiments will tell us about the connection between survival and the many other concepts with which it is connected. For example, the following thought experiment, mentioned earlier, tells us about the connection between survival and anticipatory concern. Imagine that your child is to undergo a root canal operation tomorrow without anaesthetic would you anticipate your child’s future pain? If we answer ‘no’, as I suspect we will, then this tells us that we do not anticipate the pain of another person, in this case our child. Imagine that we ourselves are to undergo the painful operation, and we do say
that we would anticipate the pain. In combination with the first case, this
tells us that we anticipate a future person’s pain when, and only when,
ceterus paribus, that future person is (a survivor of) us. Parfit’s case of
fission is another thought experiment of this sort. It shows us something
about the connections between the concept of survival and the concept of
self-concern; and, if Parfit is right about our intuitive reactions to this
case—if he is right in thinking that we believe that the correct description
of the case is that the original person ceases to exist, but that it would be
irrational of us, were we the original person in the case, to be concerned
about the impending fission in the same way in which we would be
fearful about our impending death—what it shows us is that identity
(where, for Parfit, that means simply uniqueness) and self-concern are not
as closely connected as we might have thought. Similarly, if we wanted to
elicit the conceptual connections between survival and memory we could
consider the following case: suppose you have vivid memories of strolling
around the Miro museum in Barcelona, are you the same person as the
person who strolled around the Miro museum in Barcelona? If we say
‘yes’, as I suspect we will, this tells us that we believe that, ceterus paribus,
if we remember some experience we are the same person as the person
who had that experience. (The ceterus paribus clause here includes such
things as the memories not having been implanted by an interfering
neuro-surgeon, not being hallucinations etc). And so on and so forth for
the manifold connections between survival and other concepts.

From thought experiments like these, emerge the following sorts of
platitudes about the connection between survival and other of our
concepts: ‘if a person today looks the same as a person yesterday then,
normally, they are the same person as the person yesterday’; ‘if a person
today remembers doing the things that a person three days ago did then,
normally, that person today is the same person as the person who did
those things three days ago’; ‘it is just to punish a person for a crime
committed by some past person when and only when, ceterus paribus, that
person is the same person as the person who committed the crime’;
‘normally, I will be specially and directly concerned about my going to the
dentist tomorrow in a way which I would not be about someone else
going to the dentist tomorrow’; ‘normally, if I love a particular person
today, then I will love that same person tomorrow’; and so on almost ad
infinitum. The conjunction of everyday, platitudeous principles like these
fix the role that the concept of survival plays in our person-directed practices.

To say that these principles are platitudinous is not, of course, to say that subjects, even those who are perfectly competent users of the concept, must be able to reel them off explicitly upon demand. For these principles are, for the most part, embedded implicitly or unself-consciously in our person-directed practices. To assemble these principles, and to make them explicit, informed and consistent is the job of the conceptual analyst, not of those who go about habitually and unproblematically employing the concept of survival in their day to day lives. Nevertheless, since these principles are supposed to be implicit in our practice, we would hope that these are principles that competent, informed and consistent users of the concept of survival would assent to when explicitly presented them. What these principles aim to articulate, then, are the inferential and judgemental dispositions of those who employ the concept. To possess and competently use the concept of survival is just to be disposed to make the kinds of judgements and inferences mentioned above. So that someone who systematically failed to be disposed to make these kinds of inferences would not, we would say, have mastery of the concept of survival. Someone who was systematically disposed to utter enough things of the following sort—'I believe that I shall surely be tortured tomorrow, but I see no reason to be specially worried about it'—would not have mastery of the concept of survival.

Of course, not all subjects, even when well-informed and consistent, will assent to all of the same platitudes about survival. Parfit, for example, would not agree with the widely held platitude that, normally, I should care specially and directly for some future person when and only when that future person will be identical with me. Recall, according to Parfit, what we should care about in our future-oriented self-concern are relations of psychological continuity and connectedness, not identity; and we should care about those relations even where, as in ordinary, everyday life, they happen to coincide with identity. And many consequentialists would not assent to the platitudes linking the concept of personal identity with ethical concepts of punishment, moral responsibility, attributions of ongoing rights and obligations, and so on. For, crudely, according to consequentialists, whether or not a person now
ought to be punished for some past crime depends, not upon whether the person now is the same person as the past person who committed the crime, but upon whether or not punishing the later person will have the best all-things-considered consequences. Of course, this is not to say that consequentialists will not have reason to employ the concept of survival in the context of punishment, for it may well be the case that punishing someone when and only when they are the same person as the person who committed the crime is the course of action which happens to yield the best consequences. But, for consequentialists, this principle, where it is an appropriate principle of punishment at all, will have contingent, not absolute or deontic, status.

Nor will all subjects assign all and every platitude equal weight or importance. For example, most of us assign very little importance to platitudes of the sort: ‘if someone owns a car today, typically, they will own the same car tomorrow’, or, ‘if someone has a particular haircut today, typically, they will have the same haircut tomorrow’. In contrast, we assign much greater weight to the platitudes about punishment, legal entitlement, moral responsibility, memory, self-concern, bodily continuity and the like. Moreover, even among the platitudes which most of us assign greater importance, there will be some variation in weighting. For example, physical continuity theorists such as Williams, Wiggins, Nagel and Unger (like the Somataphiles), will presumably assign greater weight to those platitudes involving physical continuity (for example, ‘if someone has the same body/brain/sufficiently continuous physical realizer then, normally, they are the same person) than they will to those platitudes involving psychological continuity (for example, ‘if someone remembers doing something earlier, then, normally, they are the same person as the person who did that thing’). Whilst psychological continuity theorists (like the Teletransporters), on the other hand, will assign less weight to those platitudes involving physical continuity and more to those involving psychological continuity.

However, survival, on this way of thinking, is a cluster concept. And what is important if people are to have the same concept of survival (or, indeed, to have a concept of survival at all), is that they share enough of the cluster: that they are disposed to make enough of the inferences and judgements regarding survival captured (for a very small part) in the aforementioned kinds of platitudes about survival. Exactly how much
counts as enough is, of course, going to be a vague matter, depending on exactly how we weight the various platitudes, and how many and how important we take the missing platitudes to be. But all those involved in the personal identity debate in this world clearly share enough of the cluster. For, though they disagree about what relations survival picks out, they nonetheless understand what it is they are disagreeing about, and know that they are disagreeing about the same thing. The obvious explanation for this is that the role that survival plays in their person-directed concerns is sufficiently similar. They share enough of the cluster of person-directed practices. What they disagree about is the relative importance or ranking of various of these platitudes.

This first stage of conceptual analysis—of teasing out the connections between survival and our person-directed practices, such as blame, punishment, self-concern etc.—by fixing the reference of survival, puts us in a position to answer the second, substantive question of what property or relation in fact underpins subjects' person-directed practices. It puts us in a position to answer, that is, the object-level question of what property or relation constitutes survival for them. But conceptual analysis alone will not answer this second question for us. For, to spell out the role that the concept of survival plays in our person-directed practices, is not yet to say anything substantial about what property or relations satisfy or instantiate that role. Merely teasing out the connections between our concept of survival and other concepts tells us very little about what relation as a matter of fact plays or realizes that role. It tells us very little about what property or properties make these principles come out true for a given subject or group of subjects. So we need a second stage of analysis: a posteriori investigation to discover what property or set of properties as a matter of fact cause members of a community to organize their person-directed practices around the assumption that X survives as

---

3Richard Swinburne, for example, claims that, if survival turns out not to consist in the continued existence of the same soul, he will cease to care about personal identity: he will be indifferent as to whether he lives or dies. Leaving aside doubts about the psychological plausibility of Swinburne's claim, if Swinburne, true to his word, really did cease to care about survival—about what befell him, his loved ones and everyone else—would he still have a concept of survival? The answer to this question depends on how central we take the platitude connecting survival to self-concern to be to survival. If we take it to be extremely important we may think that he no longer has a concept of survival; if we think it less important we may think he still has enough of the cluster to count as possessing a survival concept; or we may think it simply vague.
Y. And then, on this first-order view, that property or relation counts as the relation of survival for them.

In order for a relation to be properly said to play the survival role for a community (and so to constitute survival for members of that community) the connection between the relation and the person-directed practices cannot be merely accidental. The organization of a community's person-directed practices around a particular relation among person-stages must be counter-factually resilient in order for that relation to be said to constitute survival for members of that community. That is to say, members of the community must be disposed to organize their person-directed practices around that relation among person-stages, rather than any other, in counter-factual situations where they come apart. For example, for an immaterial entity to be properly said to play the survival role for members of a particular community (and so to constitute survival for members of that community) it has to be the case that members of that community would organize their person-directed practices around sameness of immaterial entity, rather than some other relation such as physical continuity or psychological continuity, were sameness of immaterial entity to happen not to coincide with relations of physical or psychological continuity among person-stages. In counter-factual situations where sameness of immaterial entity comes apart from relations of physical or psychological continuity among person-stages, members of a community must be disposed, for example, to hold the later person who possesses the same immaterial entity as an earlier person who did some deed responsible for the doing of that deed, not the later person who is physically or psychologically continuous with that earlier person. And similarly for enough of the rest of the person-directed practices. Likewise, in order for physical continuity to be properly said to play the survival role for members of a community (and so to constitute survival for them) it has to be the case that the person-directed practices of that community systematically covary with physical continuity, rather than some other relation. And so on for any other candidate survival relation. Where there is not such counter-factual dependence or systematic covariation between the organization of a community's person-directed practices and a particular relation, then that relation does not play the survival role for that community.
There are a number of complementary ways of going about the a posteriori task of discovering what relation plays the survival role for members of a particular community, and so constitutes survival for members of that community. We could, for example, simply sit and back and observe (a la the anthropologist or sociologist) the way in which a subject or group of subjects organize their person-directed practices. The conceptual work done at the first stage of the analysis enables us to get a fix on what sort of property we are looking for. The conceptual work tells us that survival for members of a community will be the relation (whatever it is) in accordance with which members of that community punish later people for earlier people’s crimes; in accordance with which earlier people anticipate the experiences of future people; in accordance with which they expect to remember experiences had by an earlier person; and so on. So, we might note, for example, that all and only those people that that group of subjects punish are people who are, say, strongly psychologically connected to the person who had earlier committed the crime. Similarly, we might note that subjects invariably anticipated the future experiences of persons who were strongly psychologically connected to them, but did not anticipate those future experiences otherwise. So then, we would initially and tentatively conclude that strong psychological connectedness was the survival relation for them. We would initially conclude so only tentatively, because we must also look to the world (a la scientists) to see that there is in fact such a property. Upon discovering that there is such a property, and upon observing a similar organizational pattern for a wide range of their person-directed practices, we could conclude more certainly that strong psychological connectedness was the relation which constituted survival for them.

But, unless various different, possible candidate relations for survival come apart in ordinary life, simply observing how people actually behave will not tell us much about their counter-factual dispositions. It will not tell us, for example, whether they would organize their person-directed practices around physical continuity, rather than psychological continuity, were the two relations to come apart. So here again, thought experiments which test for these sorts of counter-factual dispositions have an important role to play. We might present subjects with a variety of hypothetical puzzle cases, such as the case of teletransportation, and note which relation their person-directed
judgements and concerns systematically tracked. We would ask them, for example, should the post-teletransportation person be punished for crimes committed before being beamed up from the planet: if they say 'yes' we can conclude that they believe that a person survives teletransportation: if they say, 'no' we can infer that they believe that a person does not survive teletransportation. We are justified in so inferring, for the work done at the conceptual level tells us that we think that a person should be punished for a past crime, when, and only when, ceterus paribus, they are the same person as the person who committed the crime. Similarly, we could present them with familiar puzzle cases such as the following: a future person, who has your body, but whose psychology is very different from your own, will be tortured tomorrow, would you anticipate that future torture? If they answer 'yes' we can assume that they believe that a person survives loss of psychology, and since our conceptual work tells us that a person will anticipate future experiences when and only when they survive as that future person, we can conclude that physical, not psychological, continuity is the relation which underpins their person-directed practices and concerns, and so is the relation of survival for them. If, on the other hand, the person answers 'no' to this case, we can assume that they do not believe that a person survives loss of psychology, and so that psychological continuity is a necessary part of survival for them. Unger's 'future pain' test, another Williams-type case, would be similarly useful here. (Note, however, that, whereas Unger cautions that this test is a test for what matters in survival, not of what survival consists in, on my view, tests about what matters in survival just are tests for what survival consists in.) And so on and so forth.

In order for subjects' responses to puzzle cases to be a reliable guide to the relation that is survival for them, subjects need, first, to be reasonably informed. If someone, such as Swinburne, claims that survival consists in the continued existence of the same soul, in a world such as ours where is no such thing, he is simply mistaken about the relation which organizes his person-directed practices and concerns. For there is no such thing, and so the soul could not be the relation around which he has cause to organize his person-directed practices—it must be some other relation. Second, subjects' responses need to be internally or intra-personally consistent—they should not, for example, track survival with physical continuity in one case where physical and psychological
continuity come apart, but psychological continuity in another
descriptionally equivalent case. That is to say, they should not think that
they both survive and do not survive teletransportation, for example.
Hence the need for these constraints in the definition. These constraints
should be uncontroversial, for they are standard in the personal identity
debate.

An important point to note about this first-order account (and, as
we shall see, in contrast to the second-order view) is that it is by no means
at odds with a physical continuity account or a psychological continuity
account of survival. If the relation among two person-stages X and Y
which causes certain subjects consistently to organize their person-
directed practices around the assumption that X survives as Y is physical
continuity, and X and Y are physically continuous, then survival for those
subjects will consist in physical continuity. Likewise, if the relation
around which certain subjects have cause consistently to organize their
person-directed practices is psychological continuity, and X and Y are
psychologically continuous, then psychological continuity will constitute
the survival of those subjects. Nor, notably, is this analysis at odds with
the Cartesian view of personal identity or survival. For, if the Cartesian
view of persons as immaterial entities is coherent (and I think it is, or
certainly can be made to be4), then if sameness of immaterial entity is the
relation which causes a certain group of subjects consistently to organize
their person-directed practices around the assumption that a person
survives, and if those subjects are in a world in which there are such
things as immaterial entities, then the survival of those subjects will
involve the continued existence of the same immaterial entity. What a
posteriori investigation tells us, of course, is that in this world survival
does not consist in sameness of immaterial entity for anyone. For there is
no such thing.

But, on this way of analysing the concept of survival, these
properties or relations earn their right to count as the facts of personal
identity or survival for members of a community only in virtue of their
being the facts around which members of the community are disposed to
organize their person-directed practices; only in virtue of the fact, that is,
that subjects regard them as the facts of survival. They have no
independent claim to be the facts of survival.

4This issue will be discussed more fully in Chapter 8.
This sort of analysis enshrines the intuitive lesson of the Teletransporters and the Somataphiles. It makes subjects' person-directed practices and concerns in specified conditions not just contingently or accidentally connected to survival, but constitutive of survival. So survival cannot come apart from what members of a community care about in survival and the special and distinctive role that it plays in their person-directed practices and attitudes. For the special and distinctive role that survival plays in the person-directed concerns of a community is definitive of survival for that community. To give this analysis of what survival consists in, then, is ipso facto to give an account of what matters in survival.

But for all that subjects' responses are constitutive of the concept of survival, they do not float free of facts in the world. There are thoroughly objective, absolute properties out there in the world, properties which exist and have their character fixed entirely independently of us, it's just that which of those properties constitute survival for subjects depends on which of those properties cause those subjects consistently and resiliently to organize their person-directed practices around the assumption that a person survives.

Perhaps a good way to get absolutely clear about this account of survival is briefly to see how it is similar to (and different from) some better known accounts of other of our concepts that we discussed earlier. First, we can understand it as a particular sort (the 'contingent identity' sort) of first-order functionalist account of mental states, applied to the case of survival. Recall, analytic functionalism is the doctrine that mental state-types (such as pain) are defined in terms of their characteristic causal role in relation to inputs, outputs, and other mental states. This is a conceptual claim about what it takes for some mental state to count as pain. Then there is the a posteriori task of finding out which property (physical or non-physical) actually plays the pain role in a subject, and then that property counts as pain for them. Since, in us homo sapiens, C fibres-firing (say) is the (neurophysiological) property that plays the pain role—i.e. which is the brain-state causally responsible for us uttering 'ouch', and feeling pain, when pricked by a needle—pain is (contingently)

---

5Or, more precisely, as we shall see, a particular version of this first-order view, what I shall call 'temporal-phase pluralism', enshrines this lesson of the Teletransporters and the Somataphiles.
identical with C-fibres-firing for us. In a Martian, some other state, say, M-fibres-firing, may play the pain role; and then that state—M-fibres-firing—would count as pain for them. Pain is thus, on this account, multiply realizable. For different states may play or realize the pain role in different subjects, and those different states are all equally properly called 'pain'.

Similarly, on my account, survival is defined by the role that it plays, not in relation to inputs, outputs, and other mental states, but in relation to subjects’ person-directed practices viz., being the relation that they care about, in accordance with which they punish later people for crimes committed by earlier people, in virtue of the holding of which they treat certain people in particular ways and adopt particular reactive attitudes to certain people over time, and so on. This is a conceptual claim about what it takes for some relation to count as the survival relation for a community; and it is practice-dependent because the facts which determine which relation is the relation of survival for a community are person-directed practices. Then there is the second, a posteriori task of finding out what property in fact plays or realizes the survival role for a subject or given group of subjects, and then, by transitivity of identity, that property then counts as survival for them. But it does so only contingently. For, as the case of the Teletransporters and the Somataphiles illustrates, different properties may quite acceptably play the survival role for different subjects or groups of subjects. The survival role is multiply realizable. Survival for a Somataphile consists in physical continuity, for that is the relation which plays the survival role for members of that community; but the survival of a Teletransporter-person involves psychological continuity, for that is the relation that realizes the survival role for them. The essential structure of this account of survival might be put thus:

1. Whatever relation occupies (or realizes) the survival role for members of a community C is the relation which everyone (whether or not a member of community C) will properly take to constitute survival for members of community C (though not to constitute the survival of a person who is not a member of community C). (Conceptual claim)

---

6I borrow this way of putting the point from Jackson, who (following Armstrong and Lewis) explains first-order functionalist accounts of mental states analogously.
2. What occupies the survival role for C is relation R. (Empirical claim)

3. R is the relation which everyone (whether or not a member of community C) will properly take to constitute survival for members of community C (though not to constitute the survival of a person who is not a member of community C). (Transitivity of identity)

The case of the Teletransporters and the Somataphiles and the arguments of chapter 4, should, I hope, have already the plausibility of each of these claims, independently, and combined. But perhaps I ought briefly and schematically recapitulate.

In defining survival, as indeed any term or concept, we have a choice between two fundamental options: a choice between, as I have put it, defining survival in terms of the role that it plays in subjects' person-directed practices and concerns (or in terms of relations that happen to realize that role for a given community of subjects), and defining survival rigidly in terms of actual realizer. Practice independent accounts typically will, and indeed have, chosen the latter option. For, these accounts standardly have it, actual realizer, not role, is what is interesting, indeed essential, to survival. The arguments of chapter 4 aimed to establish, at the very least, that, however interesting actual realizer, the survival role—the role that survival plays in our person-directed practices and concerns—is just, if not more, as interesting and important. It is this role which, *inter alia*, allows us to know that we are still talking about survival, even where we may differ radically in our views about what relation realizes that role. The Teletransporters and the Somataphiles, like Parfit and Williams, are both talking about survival, even though they disagree about what survival consists in; and that is because, although different relations realize the survival role for them, the different relations are nonetheless both realizers of the common (survival) role. They are, as we might say, different species of the same genus. And just as genus, by specifying interesting and overarching commonalities between different

---

7Note that practice-independent accounts do not have to take the latter view. It is an unlikely, but perfectly coherent, position simultaneously to maintain that survival consists in the survival role (or being a relation that plays the survival role) and that this is a practice-independent truth—a fact served up by nature or God. This may indeed be Johnston's position when he claims that Protean Person is our substance sortal.
species, is important for biological classification, so role is important for classifying different realizers as survival concepts. It explains the significant commonality between the different species or realizers, in virtue of which all these realizers, though different, earn their right equally to count as survival relations nonetheless.

This 'contingent identity' first-order account differs crucially from the rigidifying view, also a first-order view. The rigidifier's analysis too, proceeds in two stages, a conceptual and an empirical stage, but the rigidifier differs in ruling out the possibility of multiple realizability (at least, the possibility of multiple realizability without change of meaning), a priori. Recall, the rigidifying view has it that pain, for example, is (necessarily) what property actually plays or realizes the pain role—pain is necessarily C-fibres firing. So that Martians, or some other differently physically constituted beings, however well they exhibit pain behaviour just like ours, could not be in pain. They will be in pain*. So too, for the case of water, the rigidifier has it 'water is the stuff which actually falls from the sky, runs from taps...etc' Since the stuff which actually satisfies this description here on earth is H2O, water is necessarily identical with H2O. But people on twin earth would mean something different by water, for water is, rigidly, the stuff which actually plays the water role. By rigidly identifying the meaning of terms with actual realizer, the rigidifying view thus rules out the possibility of multiple realizability without change of meaning a priori.

Recall, from the discussion of chapter 2, Unger is a personal identity theorist who fairly explicitly takes this view for the case of survival. Unger claims that survival is, rigidly or necessarily, whatever relation actually organizes our deepest beliefs about survival. So that a community of people who most deeply believed that survival consisted in some relation other than the one(s) which we actually most deeply believe it to consist in, would mean something different to what we mean by it. They would by talking about survival*. There is only one relation around which a community of people like us can organize our person-directed practices and it be properly called 'survival', and that is the particular relation that we as we actually are now most deeply believe to constitute survival. For Unger, the survival of all people in all worlds should be evaluated by our own present lights, whatever the person-directed practices and concerns of the other people in the other worlds. If we were
the Somataphiles, for example, we should not regard the Teletransporters as surviving teletransportation, even though we know that by their lights teletransportation preserves the survival of members of their community. For since survival is determined by our own person-directed practices and concerns, and we regard teletransportation as resulting in a person's certain death, no person survives teletransportation, whatever they may happen to think about things. At best, people with the concerns of the Teletransporters would survive* teletransportation.

Whereas Unger rigidly identifies 'survival' with the relation that actually organizes our deepest beliefs (thus ruling out the possibility of multiple realizability without change of meaning), on my account, although survival for a community is identical with the particular relation (whatever it happens to be) that realizes the survival role for that community, it is only *contingently* identical with that relation. Another relation might play that role for a community and that relation would count as survival (and not survival*, or survival**, or...) just as well. This is the crucial point of difference between a motivationally Humean, rigidifying account of survival, such as Unger's and a first-order, functionalist account of survival such as mine: it is what makes Unger's account practice-independent and mine practice-dependent. For survival, for Unger, can come apart from the relation that a community cares about—that organizes their person-directed practices and concerns. Person-directed practices, for Unger, are no essential part of survival, but only contingently associated with it. For go to a community who organize around some relation other than Unger's preferred variety of physical continuity, and physical continuity counts as survival round there nonetheless. (There is another difference between Unger's account and mine. Putting the point in my terms, whereas Unger takes the survival role to be fixed chiefly by our deepest beliefs about our survival conditions, mine takes it to be fixed by a much wider range—the totality—of our person-directed practices and concerns).

Note, before we move on, that there is a third sort of view to be had here. And, as I suggested in chapter 2, I suspect that it is the view that Wiggins holds. According to (first-order) pluralism, survival for members of a community is what plays the survival role for members of that community. So that we ought to evaluate the survival of members of different communities by their own lights, not by ours. We ought to say
that a Teletransporter-person survives teletransportation, whilst a Somataphile-person would not. On my view, survival for members of a given community is *whatever* relation plays the survival role for members of that community. In contrast, according to the rigidifying view, survival for members of any community is what plays the survival role in a *particular* community (typically the actual one). If what plays the survival role for us in the actual world is physical continuity, then physical continuity constitutes survival for all people everywhere. Consequently, no person (whether a Teletransporter-person or a Somataphile-person) survives teletransportation, for teletransportation fails to preserve the particular relation which actually constitutes survival for us. Whereas pluralism has it that role (or playing the role) is necessary and sufficient for survival, rigidifying views have it that actual realizer is what is necessary and sufficient for survival. The third view to be had would have it that both role and actual realizer are necessary (and jointly sufficient) for survival: survival, for members of any community, is what actually realizes the survival role, providing that it realizes the survival role for members of the community. Suppose that 'survival' rigidly designates physical continuity. Then, according to this third view, survival, for members of any community, is physical continuity, providing that members of the community organize their person-directed practices around physical continuity. If they do not organize around physical continuity, then physical continuity is not survival in that community. They do not survive, but, providing they organize around something else, they may survive*. I think that this 'mixed' view is implausible for the same reasons that the rigidifying view is implausible. If survival is given by what people care about, and some other community of people care about something other than what we care about, why isn't that just survival? All the more so, given that what we care about is scarcely a non-gerrymandered kind.

The first-order pluralist view is the sort of practice-dependent object-level view that I favour as against the second-order view, for reasons that I shall outline shortly. But it is worth noting that, if we take this first-order view, we are committed to a certain metaphysics, namely, to the existence of contingent property identities. That is to say, we are committed to the view that, although survival is identical with a certain first-order property (psychological continuity for the Somataphiles, say), it is only contingently identical with that property. For go somewhere
where a different property plays the survival role, (say, physical continuity) and that property counts as survival round there.

I do not think that being committed to contingent property identities is too great a cost. Indeed, I do not think it any cost at all. For the thesis, properly understood, is a pretty innocuous one. Consider the following example from Braddon-Mitchell and Jackson...

...consider the functional concept of assent. We can make true type-type identity statements about the (kind of) gesture that is a sign of assent in Western societies, and the (kind of) gesture of assent that is the sign of assent in certain Asian societies, despite the fact that the gestures are different. In the West nodding the head = the gesture of assent, whereas a different head motion—call it wiggling—is used in parts of Asia. So wiggling the head = the gesture of assent in parts of Asia.8

The ‘gesture of assent’ is thus multiply realisable. Different behaviours count as gesturing assent in different places, just as what counts as behaving politely i.e. behaving in a conventionally approved manner, may vary from place to place, culture to culture, in virtue of the fact that conventions for politeness may vary from place to place, culture to culture. Multiple realizability is thus just innocuous, old-fashioned liberal pluralism. (I will come back to this point in more detail shortly). Note particularly, in this context, that this view, by specifying the domain (for a community C) avoids intransitivities. Were the contingent identity thesis the claim that survival = psychological continuity and that survival = physical continuity we would be led, by transitivity of identity, to the claim that psychological continuity = physical continuity, which is clearly false. However, by making survival claims relative to a community (defined in terms of common person-directed practices), the contingent identity thesis avoids such false identity claims. For we say, survival for a Teletransporter-person= psychological continuity, but survival for a Somataphile-person= physical continuity. No intransitivity here. In any case, for those who may still baulk at talk of contingent identity, a weaker rephrasing would be enough for our purposes. It is enough, for our purposes, to say that ‘in the West nodding the head is properly called the gesture of assent (in virtue of the fact that it serves the function of a sign of agreement), whereas in parts of Asia a different head motion—wiggling—is properly called the gesture of assent (for it serves as a sign of agreement

8Braddon-Mitchell and Jackson, forthcoming.
round there).' These different gestures are both ‘properly called’ gestures of assent in virtue of the fact that they perform the same function—namely, gesturing assent—in the respective communities. The gesture of assent has the same meaning or character in the two communities.

Note again that if survival for a Teletransporter-person is psychological continuity, and survival for a Somataphile-person is physical continuity, then although a Somataphile-person does not survive teletransportation, a Teletransporter-person does. Thus the Somataphiles in the original case, as I presented it, were wrong to think that the Teletransporters did not survive teletransportation. For, if they have mastery of the concept of survival, they will realize that although they, the Somataphiles, do not survive teletransportation, people with appropriately different person-directed practices, such as the Teletransporters, do. For teletransportation preserves psychological continuity, the relation that is the survival relation for the Teletransporters. I will return to this issue in greater detail in chapter 9.

Perhaps not having noticed that the contingent identity thesis, by specifying the relevant domain, is not led to make false identity claims, many functionalists about mental states have retreated to the view that mental states are second-order properties. On this view, mental states, such as pain, are not to be type-identified with brain states (or any other realizer of the pain role), but with a second-order property that those realizers have, namely, being a state which typically plays the pain role. Thus, pain (for homo sapiens) is not C-fibres firing per se, rather it is being in a state (C-fibres firing) which typically plays the pain role.

An analogue of this view is available to practice-dependent theorists in the case of survival. Let me briefly sketch this second-order option, and then outline the reasons why I do not favour it for the case of survival.

5.2 Survival as a second-order property

A thought in favour of the second-order view, as against the first, runs as follows. Survival is multiply realisable. If we take the first-order view and identify survival with the particular, domain specific, realizer of the survival role, we will say that survival is physical continuity for the Somataphiles; psychological continuity for the Teletransporters. But what
is it, contra rigidifiers, that makes these different relations properly called *survival* relations nonetheless? Answer: the second-order property. It is in virtue of the fact that these relations are both realizers of the survival role for the community in question that unifies them as survival concepts. What I earlier called 'genus' and 'family resemblance' is just the survival role. So, since it is the second-order property which is doing all the unifying work, why not just identify survival with the survival role, rather than with the particular, domain specific, realizer(s) of that role? Why not just view survival as the second-order property?

In the case of mental states, there is a very strong reason for resisting this move to the second-order. For it is a pretty firm intuition there first, that mental states cause behaviour and second, that the brain-state type, and not the second-order property of having a state playing a role in one, is what causes behaviour. It is the pain caused in me by the hot potato that causes me to loosen my grip on it, not the second-order property of having in me a brain state of the sort that typically causes me to loosen my grip. That is to say, it is a fairly central intuition in the case of mental states that mental states are brutely causally efficacious. But the second-order properties—being in a state such as typically to be disposed to drop the hot potato—are causally impotent.

In the case of survival, however, it is not so clear that the intuition that survival, the first-order property, must be causally efficacious, at least in this brute sort of way, is such a firm and central one. Of course, the belief that survival involves psychological continuity must causally impact on other beliefs and behaviour. It is the Somataphiles’ belief that teletransportation results in their death which causes them to refuse to enter a Teletransporter, for example. And, no doubt, there will be some sort of causal story to be told about the properties of person-stages in virtue of which we judge that a person is the same person. For when we judge that someone is the same person we are presumably responding to some property that the person-stages share. Nonetheless, the causal connection between the properties and our judgements does not seem brute in the way that the causal connection between the properties of the hot potato and my loosening my grip seem to be. In the case of survival, then, *prima facie*, it does not seem wildly counter-intuitive to suppose that what causes subjects’ beliefs about survival is not psychological continuity, but rather the fact that psychological continuity is the realizer...
of the survival role. That is to say, there seems nothing particularly counter-intuitive about supposing that survival might consist in a second-order property in the way in which it seems highly counter-intuitive to suppose that pain, for example, might consist in such a second-order property. This is, of course, in the final analysis, a question for empirical psychology. But it is enough to note that the central intuition that mental states are brute causes of behaviour which seems to block the move to second-order identification in the case of mental states may well be absent in the case of survival.

The second-order analysis of survival will proceed in much the same way as the first. However, whereas the first-order analysis identifies survival with the particular relation that organizes a community's person-directed practices, the second-order view will identify survival with a higher-order property that these relations have. A person X at an earlier time survives as a person Y at a later time, they will say, iff some realizer of the survival role holds between the earlier person X and the later person Y. Theorists who take this second-order view will regard the first-order analysis as an analysis, not of survival, but rather, as an analysis of what it takes for a certain relation to count as a realizer of survival—as a candidate survival concept—for a given community. The first-order analysis is an analysis of the particular parochial realizer; not survival, but survival$_{s}$, survival$_{t}$, and so on. Survival, on this second-order view, in contrast to the first, does not consist in physical continuity or psychological continuity or sameness of soul... per se, rather it consists in a higher-order property which all these relations share, viz. being realizers of the survival role.

Whereas the first-order view is not necessarily at odds with any existing object-level account of survival—it is perfectly compatible with a view of survival as consisting in physical continuity, or a view of survival as consisting in psychological continuity, or a view of survival as consisting in sameness of soul, or a view of survival as consisting in some other socially workable relation—the second-order view is at odds with all these accounts of survival which have dominated the personal identity debate. For, on the second-order view, survival does not consist in any of these particular candidate realizers of the survival role. Survival consists, rather, in the second-order property that these relations possess when
Chapter 5  A pluralist account of survival

subjects organize their person-directed practices around them: namely, being a realizer of the survival role.

I mention this second-order option, because it is a different sort of view to the first-order view, and one which some may find attractive. However, although different from the first, the second-order account is not as different as one might initially suppose. It differs about the metaphysics of survival—about what property survival consists in—but, as we shall soon see, the motivational upshots of the two accounts are substantially the same.

5.3 Survival: a first- or a second-order property?

What, then, is there to choose between the first- and second-order views? Which of these views about the metaphysics of survival is right as analysis of our concept of survival boils down, in the final analysis, to the question of which jibes best with the platitudes about survival. (This is true, of course, not just as between the first- and second-order practice-dependent views, but for any account of survival, including the various practice-independent ones.) Which view of survival jibes best with the platitudes about survival is really a question for empirical psychology. However, there is an important conceptual issue to be cleared up first. For not all of the platitudes are pertinent to deciding which view is right. Here’s an example. Suppose that it turned out that people were disposed to say things of the following sort: ‘You should care about some future person when and only when some realizer of the survival role holds between you and that future person’. Would this show that the second-order analysis of survival is right? Prima facie, it might seem that it would. But actually, depending crucially on the epistemic status of this platitude, it may well not.

To see why, we need to distinguish between three sorts of platitudes which, to avoid complicating things, I glossed over in the previous discussion. There are, on the one hand, two sorts of what we might call ‘analytic platitudes’—platitudes about what we mean by survival. First, there are meta-analytic platitudes about in virtue of what some relation gets to count as the relation of survival; and, I have claimed, the platitudes here are practice-dependent ones (this is what, if I am right in assuming that my intuitive reactions are typical, the case of the Teletransporters and the Somataphiles shows). Second, there are analytic
Chapter 5
A pluralist account of survival

platitudes about which practices are relevant to questions of survival; as we might put it, which practices are the person-directed practices? So for example, there are the following sorts of platitudes that we discussed at length before; platitudes which fix the reference or determine the role of survival: ‘You should punish someone when and only when the person is the same person who committed the crime’; ‘You are entitled to hold someone to an earlier promise, when and only when they are the same person as the person who made the promise’; and so on. On the other hand, there are what we might (somewhat tentatively) call ‘contingent platitudes’ about exactly how we fill out the occurrences of ‘is the same person’ in the platitudes that fix the role: that is to say, there are platitudes about what relation actually realizes the survival role for us. So, for example, there are platitudes of the following sort: ‘You should punish someone when and only when they are psychologically continuous with the person who committed the crime’ or ‘You should punish someone when and only when the are physically continuous with the person who committed the crime’; and so on. These are the platitudes which are relevant to deciding as between physical continuity accounts, psychological continuity accounts, or any other object-level account of what relation realizes the survival role for a community. Calling platitudes of this third sort ‘contingent’ is potentially a little misleading (hence my tentativeness in using that label), since there is a sense in which the analytic platitudes are also contingent—we could have meant something different by the word ‘survival’ or ‘personal identity’, in just the same way as we could have meant something other than unmarried man by the word ‘bachelor’. However, the sense in which platitudes of the second sort are ‘contingent’ is that, consistent with the analytic platitudes remaining exactly as they are, we might have cared about (i.e. organized our person-directed practices around) some other property. Just as, for example, holding what we mean by ‘cream cakes’ fixed, we might not have found them tasty.

Now we are in a position to see why certain sorts of platitudes which might appear to support the second-order view are actually perfectly compatible with the first-order view. Suppose that we think that the right analysis of what we mean by ‘survival’ is the first-order view. That is to say, that ‘survival is whatever plays the survival role for a community’ is an analytic truth. Consistent with this, there are a number of properties or relations people might, contingently, care about or
organize around in survival. Contingently, different properties might realize the survival role for different subjects. (We saw this with the Teletransporters and the Somataphiles.) So, for example, people might say ‘survival is whatever property plays the survival role for a community, and (contingently) what plays the survival role for me or my community is psychological continuity’. A Teletransporter appraised of the first-order view would say that. A Somataphile committed to the first-order view as an analytic truth would say the following: survival is whatever plays the survival role for a community, and (contingently) what plays the survival role for me or for my community is physical continuity. Just so, and obviously similarly consistent with the first-order view, someone might say the following: survival is whatever realizes the survival role for a community, and what realizes the survival role for me or my community is the property of having the property that realizes the survival role. The property of being a property that realizes the survival role is thus just another candidate realizer of the survival role for a community—just another property, like physical continuity or psychological continuity, around which a community could organize their person-directed practices and concerns. The second-order view is thus a first-order view relative to being a relation around which a community may organize their person-directed practices and concerns. That is to say, relative to the first-order account of survival—viz., being a relation around which a community might organize their person-directed practices and concerns—that we have been calling the second-order view is just another competing first-order view. The issue emerges clearly in the following sort of case.

Suppose that you are going to undergo some process X. You don’t know what X involves: that is to say, you don’t know whether the person who will emerge from X will still be physically continuous with you, or remain psychologically continuous with you, or be related to you-now by some other candidate survival realizer about which you might now care. But you do know this: before you undergo X you will come to care about (i.e. organize your person-directed practices around) the relation (whatever it is, for you do not know yet) that obtains between you-now and the post-X person. How would you feel about the prospect of X? If you feel perturbed, it is probably because you organize your person-directed practices and concerns around physical continuity, or psychological continuity, or some other, what I have been calling ‘first-
order’, property. If you feel not in the least perturbed at the prospect of X, it is probably because you organize your person-directed practices around what I have been calling the ‘second-order’ property. You do not care that physical continuity or psychological continuity or some other socially workable relation per se holds between you and a future person, you just care that some relation that realizes the survival role for you holds between you and a future person. Suppose that some of us were to have this latter reaction (this would be equivalent to us being disposed to utter the platitude mentioned initially which seemed like it straightforwardly supported the second-order view). This would not show that the first-order view was false. It would simply show that, consistent with the first-order view (that survival is whatever property plays the survival role), we contingently care about the second-order property. The second-order property, rather than physical continuity per se, or psychological continuity per se, is the property that realizes the survival role for us. If it turned out that the ‘second-order’ view was implicit in these contingent platitudes, then, it would not turn out that the first-order view was false. It would simply go to show that the property which we (or some of us) organize our person-directed practices around is the property of being a property which realizes the survival role in a community. That survival is reflexive in this way—that role can be its own realizer—is a very interesting feature of the case of survival. And it means that the first-order view is not a simple hostage to empirical fortune. For even if it turns out that the second-order view is implicit in the platitudes (at least the contingent ones), the first-order view can nonetheless remain true. (Note that the same manoeuvre is not available to the theorist who takes the second-order view—that survival, as an analytic truth, consists in the property of being the property that plays the survival role for a community. For, if it turns out that the first-order view is the one embedded in the contingent platitudes—if it turns out, that is, that people do not organize their person-directed practices around the second-order property—this second-order view will simply be false).

[9]This said, we can notice an interesting difference between the case of water, and the case of survival. The case of survival, unlike the case of water, is reflexive: role can be its own realizer. In the case of water, recall, the role is ‘the stuff that falls from the sky, runs from taps, fills the seas...’. This is the conceptual claim. Then there is the empirical question of what stuff (or what properties of stuff) realizes that role. But, in the case of case of water, the realizer cannot be ‘the stuff that fills the ...’ for that would be an a priori, not an a posteriori, truth. Role cannot be its own realizer.
Indeed this, I suspect, may go a fair way to explaining the appeal (insofar as there is appeal) of the second-order account of survival. For, if you organize your person-directed practices around the second-order property, it is all too easy to make the mistake of taking the claim ‘survival is the second-order property’, ambiguous as it stands, to be an analytic, rather than a contingent truth—to be a truth about what survival means, rather than a contingent truth about what property realizes the survival role for you. The fact that a given community organizes their person-directed practices around the second-order property is no reason for making survival itself, as a matter of analytic truth, a second-order property. Rather, survival is the first-order property, and its particular realizer (for those who think this way, anyway) the second-order property. Their survival concept is $S_{so}$.

However, the first- and second-order views are real competitors, and not all sorts of platitudes will necessarily be consistent with the first-order view. Only the contingent platitudes which support a second-order view (or a view of survival as consisting in any other socially workable relation) will be compatible with the first-order view. If the analytic platitudes—the platitudes about what we mean by survival—support a second-order view, this would show that the first-order view was wrong, at least as an analysis of what we presently mean by survival. To reiterate, I am not particularly fussied about whether or not the first or second-order view is right, for both are compatible with practice-dependence and the morals I wish to draw from it. Nonetheless, as I see it, the analytic platitudes support the first order view.

Return to the case of the Teletransporters and the Somataphiles. Here we have two communities who initially thought that they were absolutely right about survival—they, and not the other community, had latched onto the privileged fact of the matter. Then, however, they look more closely at the other community—a community who, although having an apparently different concept of survival, don’t seem to be adversely affected by it in any way. And so, naturally, they ask themselves: what entitles us to think that we, and no-one else who differs from us, have the right concept of survival? Looking around, they can’t seem to find the privileged independent facts which would make them right and the others wrong. And so they conclude, that practice-dependence is the right meta-story for survival—person-directed
practices, and not a privileged set of facts served up by nature or God, are what determine what determines the extension of survival. They determine that person-directed practices are what determine what determines the extension of survival because it is these practices that explain why they identify the other community as still talking about survival, despite the fact that the community takes a different view about what survival involves.

This, I take it, is our train of thought when we think about the case of the Teletransporters and the Somataphiles. And it tells us something about our survival concept—it tells us something about what we mean by survival. And, I think, inter alia, what it tells us is that the first-order, and not the second-order view, is implicit in the analytic platitudes. For, although there is a sense in which the Teletransporters and the Somataphiles are both talking about survival (survival plays the same role in their respective person-directed practices), there also seems to be a sense in which they differ. Survival has a different nature in the two communities. For one community take themselves to survive teletransportation, whilst the other do not. They seem to differ about their survival conditions. However, if the second-order view really were implicit in the analytic platitudes, we would not think, as we seem to do, that there is a sense in which the two communities have different survival concepts. For, on the second-order view, since the role that survival plays in their person-directed practice and concerns is exactly the same, and since survival is that role, they are talking about exactly the same thing. Survival consists in exactly the same thing in the two communities. The first-order view does a better (indeed, an excellent) job of explaining our intuitive reactions to this case. For the first-order view explains both why we think there is something in common between the two communities and why we think there is something different. What they have in common is that survival plays the same role in their person-directed practices and concerns; where they differ is respect to what relation they take to realize that role. Since, on the first-order view, survival, is identified with realizer, not role, we can say, what seems the right thing to say, that the two communities differ about survival.

Moreover, and perhaps more tellingly, it is a central platitude about survival that, whatever survival is, it is something that we value. But when we value survival, what we value is the relevant first-order
property, not the second-order property of being a property which has the property, \textit{inter alia}, of being valued. What we care about when we care about survival is that there be a future person who is related to us by physical continuity or psychological continuity or whatever, not that there be a future person who has the property of having the property that we care about. The second-order property is typically motivationally inert. Just as, for example, when we want to do the right thing, what we typically care about is a first-order property (saving a life, making someone happy, giving to the poor...). When we care about some action being the right action we do not care that that action has the property of being the right action, we care about saving a life. There is something pretty strange about someone who does the right thing (say, saves a life), not because they care about saving a life, but because saving a life has the higher-order property of being a good act. In both cases, the second-order property is not typically psychologically motivational. Since survival \textit{is} psychologically motivational, we had better identify survival with a first-order property.

This is not to suggest that the second-order view is by any means incoherent. It seems perfectly possible that someone might care about the second-order property, in just the same way as it seems perfectly possible that someone might just desire that their desires be fulfilled, irrespective of the content of those desires (although it might be practically problematic if that were their \textit{only} desire). But it is to say that the onus of proof seems firmly on the side of the second-order theorist who must explain how, contrary to these initial appearances, the second-order view is somehow really buried deep in the platitudes.

I, then, shall rest my hand with the first-order view. There are, however, number of different ways of understanding or fleshing-out the first-order view. I favour a specific one of these. Perhaps the best way of understanding the motivation for this more specific first-order view is to approach it by way of setting out the metaphysical and motivational consequences of another sort of first-order approach to survival.
TEMPORAL-PHASE PLURALISM

Of the those few philosophers who have mooted what I have called a 'practice-dependent' approach to survival, only one, Mark Johnston, has gone on to develop an object-level account of exactly what properties survival consists in in any great detail. Johnston does not (or not obviously, at any rate) draw the distinction between the first- and second-order views; and his own account can be read either way (although, in what follows, I will read it as a first-order account). Rather, he draws a distinction between 'cryptophase kinds' and 'substance sortals'. Johnston writes,

A plausible principle drawn from the theory of sortals is that each individual is classifiable under one and only one substance sortal. A substance sortal is a term for a sort or kind of thing, a term which is the substance sortal for a given individual, just in case it is the most specific kind of term such that necessarily there is no time at which the individual exists without satisfying it at that time. Thus 'person or dog' is not a substance sortal for persons, since there is at least the more specific sortal 'person'. And 'child' is not a substance sortal for individual Human Beings; it is a mere phase sortal [cryptophase kind] which applies to a Human Being during what need be only a part of his total existence.1

What Johnston terms a person's 'phase-sortal' or 'cryptophase kind' is the first-order view of survival that I have endorsed above. A person's cryptophase kind is given by the relation that organizes their person-directed practices at a particular time. So, to translate into Johnston-ease, the Somataphiles' cryptophase kind is physical continuity; the Teletransporters' cryptophase kind is psychological continuity. However, whereas on the first-order view I have advocated, the first-order or 'cryptophase kind' view straightforwardly gives a person's survival conditions, on Johnston's view it gives their survival conditions only for a given phase in their continuing existence. For, Johnston's central object-level claim is that the substance sortal for persons is, not cryptophase kind, but Protean person: "We are inevitably and always Protean persons, that is, we are inevitably and always such that

the cryptophase kinds under which we fall at particular times are determined by our concerns and expectations at the respective times."^2 Protean person is the overarching or, as Johnston puts it, "higher-order", kind in virtue of which the different cryptophase kinds which organize an individual's person-directed practices at respective times nonetheless count as stages of the same persisting person; much as Human Being is the overarching kind which unites different chronological phases in a Human Beings' development—childhood, adolescence, middle-age, old-age—into developmental phases of a continuing Human Being; phases which human beings can live through. Protean person is the disjunction of all the realizers that might play the survival role for a subject or community at different times in their continuing existence; or, as Johnston would probably prefer to put it, the disjunction of all the acceptable sharpenings of the (vague) concept of survival.^3 That Protean person is our substance sortal is, Johnston thinks, an a priori truth, delivered by philosophical reflection on the sort of thing that we are.

6.1 The problem of refiguration.

Johnston is unhappy with the sort of first-order, 'cryptophase kind' view of survival I have advocated because he wants to draw from practice-dependence a 'radical' and 'revisionary' upshot—"we all now have reason to aim to internalize a socially enactable concept of personal identity which makes available and contributes to a better life"^4; "If Modified Relativism is true, then the obvious upshot is that we should all think of ourselves as Protean persons and should seek a cryptophase conception of ourselves which fits a life which is best or maximally good"^5. That is to say, we may all now have reason to reorganize our person-directed practices around

---

^2 ibid., p.466. This is also just one example of a claim from Johnston that can equally well, if not better, be read as advocating the second-order view—survival is the property which all the cryptophase kinds have in common, viz being a realizer of the survival role for a given community at a given time. For the overarching property which unifies different cryptophase kinds is just the second-order property; and, Johnston seems to be suggesting here, survival (Protean person) is that unifying property.

^3 On the second-order understanding of Protean person, Protean person is the disjunction of all the relations which have the property of realizing the survival role for a given community at different times in their continuing existence.

^4 ibid., p.468.

^5 ibid., p.467.
whatever relation it is that leaves us able to survive such things as teletransportation, which will give us access to such advantages as superfast travel and freedom from disease.

The motivation for this thought, I take it, is this. When we were practice-independent we thought that our survival conditions were carved immutably in stone. To organize our person-directed practices and concerns in terms of some relation other than the privileged relation served up by nature or God would be to make a serious mistake—it would be to miss or misdescribe something intrinsically important about the way the world is. Once we become practice-dependent, we realize that there is no such privileged relation served up by nature or God which demands that we organize our person-directed practices and concerns around one socially workable relation rather than another if we are not to be seriously mistaken. Rather, what makes some relation count as the relation of survival is that that is the relation around which subject's organize their person-directed practices and concerns. So communities, such as the Teletransporters and the Somataphiles, who successfully organize their person-directed practices and concerns around different relations, have blamelessly different survival concepts. And once we come to think that another community might, quite blamelessly, organize their person-directed practices and concerns around a relation different from the one around which we organize ours, it looks as if we can admit that we ourselves would be blameless had we organized, or were we to come to organize, our person-directed practices and concerns around a relation different from the one around which we presently organize them. By analogy with the inter-communal case, practice-dependence tells us something about the intra-communal case. It tells us that, were we to organize our person-directed practices around some relation other than the one around which we presently organize them (were the Somataphiles to become Teletransporters, for example), we would come to have a blamelessly different survival concept—we would have a blamelessly different view of the sorts of events that we could survive. Once we admit multiple realizability for different communities, it looks as if we can similarly admit it for our own.

Practice-dependence thus seems to make for a radical new possibility in survival. It seems to pave the way for us to change our survival conditions
by refiguring our person-directed practices and concerns in terms of some new, more advantageous, relation. It seems to pave the way for us to internalize a concept of personal identity which best suits our needs and wants. It seems to free us to organize our person-directed practices around whatever relation it is that makes for the good life. In so doing, practice-dependence places the question of personal identity firmly in the moral and political sphere.

But, Johnston thinks, practice-dependence makes for an even more revisionary upshot than this. According to Johnston, practice-dependence makes for the possibility for the Somataphiles, not merely to view teletransportation as providing great benefits for other people with a blamelessly different set of person-directed practices and concerns such as the Teletransporters, but as able to provide great benefits for them, the Somataphiles. Practice-dependence, Johnston thinks, paves the way for the Somataphiles to rightly view teletransportation as a means to realize a better life for them. It makes for the possibility of surviving events which, by our present lights, we determinately fail to survive. This is the revisionary moral that Johnston wants to draw from practice-dependence.

However, my account of survival (in Johnston’s terms, a ‘cryptophase-kind’ account) poses a problem for Johnston’s revisionary aspirations. The problem is that, on my view, the statement ‘I will survive teletransportation’ uttered by a Somataphile is simply, straightforwardly false. For teletransportation fails to preserve physical continuity, the relation that they care about. And so the Somataphiles cannot view teletransportation as providing great benefits for them. For they will die upon teletransportation. Moreover, even the statement ‘Were I to refigure in the appropriate way, I would survive teletransportation’ is, from their perspective, false as well. For they stand to the thoughts and doings of that later refigured self as they presently stand to the Teletransporters. Once appraised of pluralism, they

---

6 Refiguration is a very interesting issue in itself, and raises many questions. For example, might individuals refigure on their own, or must it be a community-wide change of mind and concern? Interesting though they are, I shall not address these issues in any detail here. Moreover, in what follows, I will speak of ‘refiguration’ as if it were the easiest and most off-hand of things to do. But, of course, it may often involve a quite radical re-orienting of perspective and practice. Perhaps some individuals could refigure overnight, but I suspect that, for most of us, it would be a gradual process.
regard the Teletransporters as having a blamelessly different survival concept, but as having a different survival concept nonetheless. For whereas the Teletransporters regard teletransportation as a quick and efficient means of transport, they regard it as certain death. At most, what can be true for a Somataphile appraised of pluralism, is that, although they would not now survive teletransportation, some future person-stage, were they to have the appropriate person-directed practices, would blamelessly regard teletransportation as survival-preserving. But such a person-stage, post-teletransportation, would not be a stage of them. For, by their lights, they will not survive such an event.

Moreover, on my analysis of survival, the initial act of refiguration (i.e. the reorganization of person-directed practices around some new, more advantageous relation) is entirely unmotivated from the point of view of the participants. For think of things again from the point of view of a Somataphile-person convinced of pluralism. They know that their survival conditions are not, of natural necessity, carved immutably in stone, but are given by the relation that organizes their person-directed practices and concerns. However, that relation presently happens to be physical continuity; and, from their present point of view, refiguration will result in their doing things (such as teletransportation) which, far from bringing them great benefits, will simply result in their death. From their present point of view refiguration looks like an absolutely crazy, suicidal (or, at any rate, suicide-inducing) thing to do. And so it looks as if, by their present lights, the Somataphiles have pretty compelling reason not to refigure; to self-bind to prevent such refiguration which will result in their doing such foolhardy things as teletransportation, if needs be. From their point of view, the question posed by such refiguration is not the question, ‘will teletransportation make for a better life for me?’ but rather, ‘am I prepared to die so that some future person very like me in psychological respects (about which I don’t much care) can reap great benefits?’ As Johnston himself describes the problem,

The...objection is that it is simply not correct for our representative Human Being to think that having survived reculturation, he will then survive Teletransportation. The objection is that since at the time at which he is deliberating his concept of personal identity is the concept of being the same human being, he ought to employ that concept in thinking about and
evaluating all possible future situations, even possible future situations which involve his adopting a different concept of personal identity. According to the objector this is a general point about concept use with respect to possible situations. We must always clearly distinguish the concepts we presently use to describe, think about, and evaluate the situation from the concepts used in the situation. Thus it is not correct to say that if we come to mean by “green” what we now mean by “blue”, then the concept green would apply to the sky, so that then the sky would be green. The sentence “The concept green applies to the sky” would then be taken by us as true, but nevertheless from our present perspective we must we must say that the concept green does not apply to the sky. Similarly, our representative Human Being at the time of considering the upcoming reculturation and Teletransportation should consistently employ his present concept of personal identity. From his present perspective he would not survive Teletransportation even if he were recultured first. Reculturation would simply give him false views about what he could survive. So he has no reason to refigure his pattern of concerns and expectations, and hence his concept of personal identity. Teletransportation can never provide benefits for him...It is important for my purposes that this objection can be got around since the moral I wish to draw from relativism is that we are all in a situation in some ways parallel to our representative Human Being. We may all have reason to refigure our concerns and expectations.7

The problem for Johnston, then, who wants to draw from pluralism this ‘revisionary moral’, is how to make it true, and motivationally salient, that the Somataphiles can view teletransportation as providing great benefits for them, and not just for some other future person or persons distinct from them.

Note that it will not help Johnston here to move to a second-order account. This is what I had in mind when I earlier said that, although the first and second-order views differ about the metaphysics of survival, they do not have substantially different motivational upshots. For teletransportation, because it does not preserve physical continuity, does not preserve the higher-order property possessed by physical continuity either—namely, realizing the survival role for the Somataphiles.8 For, whether survival is

7 *ibid.* pp. 461-2.
8 There is a different sort of second-order view which will allow for the possibility of refiguration. It is the view which says that survival consists in the second-order property of being a relation that might play the survival role for a community i.e. being a relation around which, compatible with their cultural identity, they could organize their person-directed practices. No problem about refiguration here, for psychological continuity, just as much as
physical continuity or the higher-order property possessed by physical continuity (namely, being the realizer of the survival role for the Somataphiles), the extension of the concept of survival is exactly the same—a person will survive or fail to survive exactly the same set of events in either case.

6.2 Are we Protean People?

Johnston’s solution to the problem of refiguration is, in effect, just to deny that the application conditions for the use of the term ‘survival’ in a person’s mouth are given by the relation that presently organizes that person’s person-directed practices and concerns. According to Johnston, our survival conditions are given not just by the relation that organizes our person-directed practices and concerns at the present time, but also by the relations that will come to organize them at future times. It turns out that we can survive events which by our present lights we determinately fail to survive. For, Johnston claims, being a Somataphile-person (regarding teletransportation as certain death) is a mere phase sortal, real survival—our substance sortal—is Protean person. According to Johnston’s Protean person view, a person X survives some future event e just in case e is not an event at odds with X’s cryptophase kind at the time of the event e. That is to say, you survive an event just in case, by your lights at the time of that event (not by your lights now), you survive that event. By refiguring their person-directed practices and concerns in terms of some new relation before an event, a person can bring it about that they survive that event, even if, by their present, pre-refigured lights, they determinately do not survive that event. For, from their perspective then, at the time of that survival-threatening event, they would survive that event; and so, according to the Protean person view, they would survive it.

According to the Protean person view, then, a Somataphile-person who reorganized their person-directed practices around psychological continuity, and was then teletransported, would survive teletransportation, despite the fact that, by their present lights, they fail to survive. For the Protean person view tells a Somataphile-person that, were they so to refigure

physical continuity, has the property of being a relation around which a community could organize their person-directed practices and concerns.
before the event of teletransportation, they would survive teletransportation, despite what they happen to think about things now. For, in so reorganizing their person-directed practices, they will have changed their cryptophase kind: they will have become a Teletransporter-person. And by their lights then, at the time of the event, teletransportation would be survival-preserving. But a Somataphile-person, forcibly teletransported before reorganizing their person-directed practices around psychological continuity, would not survive teletransportation. For their person-directed practices would still be organized around physical continuity—they would still be of cryptophase kind, Somataphile—and, by their lights then, as now, teletransportation would result in their death. According to the Protean person view, then, reorganizing your person-directed practices in terms of a relation according to which you survive teletransportation (if you have to), before you get into the teletransporter is the difference between life and death. (Note that a problem for Johnston here is what to say about a community who have a concept according to which refiguration—the initial act of reorganizing their person-directed practices around another relation—results in their death. It is certainly possible to imagine a community who have a concept of survival according to which you do not survive such a change of mind. In fact, I think, there are actual communities—various religious sects in the United States, for example—who believe something strikingly akin to this.)

If Johnston’s account is right, then we must revise our initial opinion that the Somataphiles in our original story were just blamelessly different from the Teletransporters to regard teletransportation as resulting in their certain death. For it turns out that they are not blamelessly different after all. They are not blameless to care deeply and non-instrumentally as they do about physical continuity in survival; so deeply that they turn down the offer of the teletransportation technology with all the benefits it could bring. Indeed, if we take Johnston’s (apparently non-Humean!) revisionary normative claim at face value, not only were they mistaken to regard teletransportation as certain death, but moreover, they were defective in choosing not to become Teletransporters themselves. They were defective not to refigure their person-directed practices and concerns in terms of the Teletransporters’ concept of survival which, were they so to refigure, could
then ‘make available and contribute to a better life’ for them. As Johnston writes,

The relativist should say that kinds like Human Being and Series Human Being are, in fact, cryptophase kinds, associated with phases which persons can live through given, but only given, special circumstances that is refuguration... Nonetheless—and this is the crucial point—although these concepts were correct concepts of personal identity modulo the restriction that no refuguration takes place, they were strictly defective. For they were concepts of identity restricted to (unobvious) phase kinds.9

The Somataphiles were defective not to refigure their person-directed practices and concerns in terms of psychological continuity because the statement ‘I will survive teletransportation, providing I refigure in the appropriate way beforehand’, uttered by an individual of the cryptophase kind Somataphile, is true. It is true because the occurrence of ‘I’ in this statement picks out, not an individual of cryptophase kind Somataphile (for, of course, from their perspective as a Somataphile, they will die at teletransportation, whatever some later person-stage may think about things) but an individual of the overarching kind Protean person—an individual who can survive refuguration and, having survived refuguration, will then survive teletransportation. For they will then, at the time of that survival-threatening event, be of cryptophase kind Teletransporter, and from their perspective then, teletransportation will be survival-preserving. And, since whether or not a person survives an event such as teletransportation is determined by their having the appropriate person-directed practices and concerns at the time of that event, they will survive that event. As Johnston writes,

Once our representative Human Being is converted to Modified Relativism about personal identity, he should think that his previous conviction that he could not survive any process, however complex, which brought it about that nothing after the process is the same human being as him was a mistake, the mistake of taking a phase kind to be a substance kind. He should conclude that his substance kind is the kind Protean person. Then he will be able to trace himself through the reculturation and even through the Teletransportation by thinking, "I will survive reculturation, and then I will survive Teletransportation". Both occurrences of "I" in this thought pick out an individual of the substance kind Protean person, and so an

9ibid., p.464.
individual which will survive reculturation and then Teletransportation.\textsuperscript{10}

The Somataphiles were then mistaken to attach so much importance to physical continuity in questions of survival. For such concern hinged on a false view of the metaphysics of survival. In caring so deeply and non-instrumentally about physical continuity in survival, the Somataphiles were not caring about \textit{survival}, but instead about a mere phase-sortal; about just one parochial possible realizer of the survival role. According to Johnston, to think that our survival conditions are straightforwardly given by the relation about which we \textit{presently} care would be to make a big mistake about the metaphysics of survival. It would be to confuse our cryptophase kind for our substance sortal. Our substance sortal is Protean person; and what is important for survival for Protean persons is not what a person presently thinks about survival, but what that person will think about survival at future times: most saliently, at the future times where they undergo survival-threatening events.

But why is it such a mistake for a person to care about surviving by their present lights? Why, in Johnston's terms, is our substance sortal Protean person rather than 'Cryptophase Kind'? Certainly, nothing in practice-dependence tells us that our survival conditions are determined by our lights at future times (the times of survival-threatening events) rather than by our present lights. Nothing in practice-dependence tells us that when we wonder now whether we will survive a survival-threatening event in ten years time we must predict what a future person in ten years time will think about that survival-threatening event. Of course, nothing in practice-dependence obviously tells us that our survival conditions are determined by our present lights either. Practice-dependence simply tells us that a relation earns its right to count as the relation of survival for members of a community in virtue of its being the relation about which members of that community care. It does not, in itself, tell us which practices at which times determine a person's survival conditions in circumstances where a person's person-directed practices change (although it does tell us that all socially workable survival concepts are on a normative par). That is to say, it does not in itself

\textsuperscript{10}\textit{Ibid.}, p.467.
tell us whether our practices now determine our survival conditions or whether it is our practices at the time of some future survival-threatening event (or our practices at any other time for that matter) which determine whether or not we survive any given upcoming survival-threatening event.

But Johnston's way of going comes at some considerable cost. For the first part, and I shall come back to this point, it seems completely arbitrary to pick on a person's person-directed practices at the time of a survival-threatening event, rather than their practices at some other time, as the privileged perspective that determines whether or not that person survives that event.

For the second part, on Johnston's account, if you are now wondering whether or not you will survive some future survival-threatening event in ten years time, what you should be thinking on is not what you think about things now, but whether a future person in ten years time will regard that event as survival-preserving. In order to know whether the statement 'I will survive teletransportation' uttered by you now is true or not, you need to predict what the opinions of a future self ten years on just before teletransportation will be. If that future self has the Somataphiles' beliefs about teletransportation, then you will not survive teletransportation, and the statement 'I will survive teletransportation' uttered by you ten years earlier will turn out to be false (even if you now think it counts as determinately true). If, on the other hand, that future person has the Teletransporters' beliefs about survival at the time at which they are teletransported, the statement 'I will survive teletransportation' will turn out to be true (even if you now think it counts as determinately false). The counter-intuitive thing here is not that our predictions about what a future person will think, and so whether or not we will survive future events, might be mistaken. For we can have mistaken beliefs about what relation underpins our person-directed practices at the present time (as, for example, does Swinburne, who believes that he organizes his person-directed practices around the soul, when in fact there is no such thing) let alone at future times. The counter-intuitive thing about Johnston's view is that we have to make predictions at all in order to know the application conditions for the term 'survival' in our mouths now.
It is no accident that we typically apply aesthetic concepts, value concepts, the concept of survival, and the like—concepts which connect up with our concerns, sentiments, and dispositions to act in important and intimate ways—in accordance with our present lights, rather than our lights at other times. For it is our present perspective on these things, not our predictions about the perspective of some future self or anyone else, that is typically psychologically and motivationally salient for us.\footnote{This is exactly the worry that so many have about 'ideal-observer' theories in meta-ethics: theories which say that what you ought to do in a situation is what an ideal observer would do. But why, you might well ask, should I care about that; why should I be in any way motivated to do what some 'ideal observer' would do?} If we believe that we will shortly die, for example, it just doesn’t seem to be much consolation to be told that at the moment just before our death we will take a different view of things. We just don’t seem to care much about surviving by some future person’s lights; we care about surviving by our present lights.

A central intuition driving practice-dependence, I earlier argued, is the intuition that survival should not be able to come apart from what a community cares about. But, on Johnston’s account, after all that effort, it yet again does. As both I and Johnston set up the case of the Teletransporters and the Somataphiles, for example, the Somataphiles care about physical continuity. They do not care about surviving Protean-style—that there may be some future refigured person who will regard teletransportation as survival-preserving just before that event—they care about surviving by their own present, situated lights. They realize that this concern may seem parochial from the Protean person view of things, but that is their concern. They do not care about surviving Protean-style, they care about surviving by their present, situated lights. Insofar as a community may care about surviving by their present lights, survival, on Johnston’s account, comes apart from what a community care about. (Note that even on Johnston’s account, you should care a little about the relation that presently organizes your person-directed practices, for if there is no relation that organizes your person-directed practices at any given time, you do not survive at all (at least by your own lights)! But, if you care about survival, you should not care too much about the relation that presently organizes your person-directed practices. For the Protean person view tells you that you will survive reconfiguration and, insofar as you care about survival, you must also care about
the relations that, post-refiguration, will come to organize your person-directed practices and concerns.)

Of course, the sense in which survival can come apart from what a community cares about on Johnston’s account is different from the way in which, for example, Parfit claims that survival can come apart from what matters. Parfit’s claim is that, although where you have survival (identity) you always have what matters, the reverse is not true—sometimes you can have what matters, where you do not have identity. That is to say, for Parfit, survival (identity) is sufficient, but not necessary, for the preservation of what matters. On Johnston’s view, where survival comes apart from what a community cares about, it comes apart, not because survival ever fails to guarantee the holding of what matters, but because a community may care about something else more than they care about survival. They might care more about the relation that presently organizes their person-directed practices than they do about the relations that will organize their person-directed practices at future times. They may prefer to die, than to refigure and survive. That is to say, whereas on Parfit’s view, identity is sufficient, but not necessary for the preservation of what matters, on Johnston’s account, survival is necessary, but not sufficient, for the preservation of what matters.

Nonetheless, on Johnston’s account, to care greatly about the relation that presently organizes your person-directed practices and concerns is instrumentally irrational, if not irrational simpliciter (for I suspect that Johnston is all of a sudden non-Humean when it comes to his revisionary upshot). For, insofar as you care about survival, you should care not just about the relation that presently organizes your person-directed practices, but also about the relations that will organize them at future times. The Somataphiles, for example, were instrumentally irrational to care so much about physical continuity that they turned down the offer of the teletransporter technology. For such concern prevented them from refiguring to take advantage of the teletransporter technology which, the Protean person view tells them, were they only so to refigure, would provide them with great advantages.

Moreover, on Johnston’s account, when a person is confronted with a survival-threatenning event—an event which, by their present lights, they do
not survive—the instrumentally rational thing for that person to do is to hasten to refigure in whatever way is required before they undergo the survival-threatening event. For, on Johnston’s view, then, but only then, will that person survive it. We might imagine, for example, that the Teletransporters decide that the way to convince the Somataphiles to take up the offer of the teletransporter technology is forcibly to teletransport them a few times. For then they will come to see that teletransportation does them no harm. On Johnston’s account, in such a situation, if the Somataphiles care about surviving, the instrumentally rational thing for them to do is to refigure post-haste before they are teletransported. Suppose, moreover, (as is plausible given the depth of the Somataphiles’ concern for physical continuity), they find that, try as they might, they cannot cease greatly to care about physical continuity. Increasingly desperate, they beg the Teletransporters not to teletransport them. The Teletransporters try to help out. They begin to hold special brain-washing clinics, after which Somataphile subjects emerge unharmed, except for the fact that they no longer care much about physical continuity in survival. To speed things up in time for the impending teletransportations, Teletransporter scientists develop a tablet which, when taken, causes the subject to cease to care about physical continuity. Signing up for the brainwashing and/or taking the mind-altering tablet doesn’t seem like a particularly rational thing to do (especially if you now think that it will only lead you happily to later suicide), but, on Johnston’s account, it is exactly the rational thing to do. If you care about surviving, you ought to take the mind-altering tablet. But it just doesn’t seem that, if you care about surviving, brainwashing and/or taking the mind-altering tablet would be the rational thing to do. Nor would it seem much consolation to know, as you raised the tablet to your lips, that although (as you now most deeply believe) teletransportation will result in your death, as you step into the teletransporter you will have an entirely different view of things. Needless to say, it seems as if something has gone wrong here. True enough, there are cases where actively seeking brainwashing might be the rational thing to do in order to survive—in China during the cultural revolution, for example—but that we may all have reason to be such cultural fugitives seems hard to believe.
But perhaps Johnston takes his account to be revisionary of our ordinary way of thinking about survival. Perhaps Johnston’s claim is that we should cease to think of survival as we do. We should cease to care about surviving by our present lights, and instead care about surviving by some future person’s lights: most notably, by the lights of a future self who will undergo some survival-threatening event. But Johnston offers no argument as to why we should think of survival in this revised way rather than in the way we presently do, except that it allows for his ‘revisionary moral’. And, moreover, it is hard to see what, compatible with practice-dependence, such an argument could be.

Johnston himself feels the pull of the intuition that our present concerns, parochial though they may seem from the Protean person view of things, are nonetheless our concerns; and we ought to respect them as such. Why, for example, should it be a person’s person-directed practices just before the survival-threatening event, rather than just after it, that determine whether or not they survive that event? Johnston’s answer, I take it, would be that we need to respect “the basic relativistic idea...that what a person considered at a time can survive depends on his concerns and expectations at that time.”\(^\text{12}\) The problem is that, if we respect the perspective of the Somataphiles as giving their survival conditions, then pluralism will not allow for Johnston’s revisionary upshot. But the next best thing—making their survival conditions determined, not by their present perspective, but by their perspective at the time at which they undergo a survival-threatening event—will. The problem is that the next best thing fails to respect the Somataphiles’ perspective on survival, according to which teletransportation will result in their death, whatever they may later come to believe. However, as Johnston sees it, it is impossible both to accommodate the intuition that a person’s survival conditions are straightforwardly determined by the relation that presently organizes their person-directed practices and to allow for the ‘revisionary upshot’. And so, in the end, Johnston trades off the intuition that a person’s survival is determined by their present perspective on the matter for the possibility of refiguration. This is what happens in the move from cryptophase kind to Protean person. Apparently, for Johnston, the revisionary benefits are worth the cost. If I had to choose, I would say that

\(^{12}\text{Johnston}(1989c), \text{p.464.}\)
the benefits were not worth the cost. But, fortunately, I do not have to. For there is a way of fleshing out first-order practice-dependence which will allow for the sense in which practice-dependence makes for a revisionary upshot, without denying that there is an important sense in which, when the Somataphiles say 'whether I refigure or not, I will not survive teletransportation', they really do utter a truth.\textsuperscript{13} Call this account 'temporal-phase' pluralism. Exactly how it is a pluralist, rather than a relativist, view will emerge in section 6.5.

\textbf{6.3 Temporal-Phase Pluralism}\textsuperscript{14}

It is perhaps easiest, to begin with, to explain this view in four-dimensionalist terms (although there is, I think, a three-dimensionalist paraphrase\textsuperscript{15}). Four-dimensionalists regard persisting things, including persons, as made up of distinct temporal parts. We can think of these parts as like the dots on a join-the-dots picture, before they have been joined up to form a picture. (Or as like the stars in the sky, before they have been joined up into the constellations). From the four-dimensionalist's point of view, the question of personal identity is the question of how to put the temporal parts or person-stages (where here 'person-stage' is being used as a synonym for temporal part of a person, and thus not agnostic between three- and four-dimensionalism) together into parts of a single, whole, continuing person. Or, to use the join-the-dot analogy, how do we join the dots—through which dots do we draw the survival analogy. Which temporal parts are parts of a single continuent person?\textsuperscript{16}

Practice-dependence gives us an answer to this question or, at least, it tells us how to go about finding the answer to this question. It tells us that different person-stages count as stages of the same person when, and only when, those person-stages are related by the relation (whatever it is) around which a person (or set or aggregate of person-stages) organize their person-

\textsuperscript{13}Temporal-phase pluralism is also perfectly compatible with the second-order view, if the second-order view turned out to be right after all.

\textsuperscript{14}This section owes a special debt to many stimulating and helpful conversations with David Braddon-Mitchell.

\textsuperscript{15}I will briefly discuss how such a paraphrase might go in Chapter 7.

\textsuperscript{16}From the three-dimensionalist's point of view, the question of personal identity is the question of what makes a person wholly present at one time the same person as a person wholly present at another.
directed practices. The problem posed by refiguration is the problem of what to do when different person-stages organize their person-directed practices around different relations—when there is conflict between the sorts of events that an earlier set of person-stages views as survival-preserving and those that a later set of person-stages views as survival-preserving. That is to say, when the prospective attitudes of an earlier set of person-stages conflict with the retrospective attitudes of a later set of person-stages. When an earlier set of person-stages views an upcoming event (say, teletransportation) as certain death, while a later set of person-stages take themselves to survive it. What we are imagining in refiguration is the Somataphiles' and the Teletransporters' perspectives on survival rolled up into the perspective of a single, deeply conflicted, fragmented self.

One solution to the problem would be to dissolve the conflict of perspectives by abstracting away from the perspectives of person-stages altogether and taking a detached, unsituated, god's-eye perspective on the conflict. (This is the solution which Johnston adopts. Where refiguration is in the offing we should think of our survival conditions in the following abstract, schematic way—I will survive an event just in case I have a concept according to which I survive that event at the time at which I undergo that event). Another would be to resolve the conflict by privileging the situated perspective of one set of person-stages over another—to claim that one situated perspective rather another is the situated perspective which determines a person's survival conditions. The problem is that both of these solutions go against the grain of practice-dependence. For, according to practice-dependence, and in vivid contrast to practice-independent views, survival is importantly perspectival—relative to the perspective and concerns of individuals located in time at times. And, according to practice-dependence, no one (socially enactable) perspective on survival has any more privileged claim to give a subject's survival conditions than any other.

In the spirit of practice-dependence, then, we should claim that survival is relative to the perspective of individuals in time at times. But, in circumstances where different person-stages have different perspectives, which practices at which times? It would seem completely arbitrary to privilege the perspective of any one person-stage over another as determining whether or not a person survives some event. Strangely enough,
Johnston himself notes this point as against the claim that a person's survival conditions are determined by a person's present perspective on survival. He writes,

...[I]f the basic relativistic idea is that what a person considered at a time can survive depends on his concerns and expectations at that time, then when considering someone who has refigured his core concerns and expectations, it will seem arbitrary for the relativist to privilege this person's earlier concerns and expectations and say that they alone determine what he can survive, so that, for example, a Human Being, even after reculturation as a Teletransporter remains able to survive all and only what a Human Being can survive. There should be a certain symmetry here.17

Johnston, I think, is absolutely right: there should be a symmetry here. But if privileging the earlier perspective of an individual is arbitrary, so too is privileging the later perspective of the individual at the time of a survival-threatening event. If we go the former route, we arbitrarily privilege the earlier perspective; if we go the latter route, we arbitrarily privilege the later perspective. But there is one perspective which it is not arbitrary to privilege. And that is the perspective of each and every person-stage. According to temporal-phase pluralism (TPP), the perspective of each and every person-stage equally determines a person's survival conditions. How?

Each and every person-stage regards certain other person-stages as related to them in the appropriate way. Practice-dependence tells us what that appropriate way is or, at least, it tells us how to find out what that appropriate way is. It tells us that a person-stage A1 will regard another person-stage A2 (or A3 or A4 or...) as being related to them in the appropriate way when, and only when, A2 (or A3 or A4 or...) share the relation around which person-stage A1 organizes their person-directed practices and concerns. Each person-stage constructs a person out of all and only those person-stages which are related to them in the appropriate way, i.e. by the relation about which that person-stage cares. A person stage, A1 has a relation Ri which constructs a person out of all and only those person-stages which are Ri related to A1; a person-stage A2 has a relation Ri which constructs a person of all those person-stages which are Ri related to A2; a person-stage A3 constructs a person made up of all and only those person-

17loc.cit.
stages which are Ri related to A3; and so on for each and every person-stage. We can represent the situation diagrammatically like this (Figure 1):

In the first line of the diagram, a person-stage A1 constructs a person out of two other person-stages, A2 and A3, which are all and only the person-stages which are related to A1 by the relation around which she cares. A4, A5, A6 and A7 are not, from A1's perspective, related to her in the appropriate way, and so not parts of her, the same whole or continuent person. In the second line, A2 constructs a person out of all and only those person-stages which are related to A2 by the relation around which she cares, namely, in this case, A1 and A3. A1 and A2, in this case (the first two lines of the diagram), construct the same person: both person-stages regard all and only person-stages A1, A2 and A3 as being related to them in the appropriate way. A3 (the third line in Figure 1), however, constructs a different person from the person that person-stages A1 and A2 construct. A3 constructs a person out of person-stages A2, (A3), A4 and A5 which are all and only those person-stages which, from her perspective, are related to her in the appropriate way. In Figure 1, although A1 regards A3 as being related to her (A1) in the appropriate way, A3 does not regard A1 as being related to her (A3) in the appropriate way. A1 regards A3 as being a part of her, but A3 regards A1 as being a part of a different person. For, from A3's perspective, A1 is not related to her (A3) in the
appropriately way, i.e. by the relation around which she organizes her person-directed practices and concerns. Perhaps, for example, A1 is not related to her (A3) by physical continuity, the relation about which she (A3) cares. A1 cares about some different relation, say, psychological continuity, and person-stages A2 and A3 are related to her in that way. Survival is thus relative to the perspective of each and every person-stage in time at a time: which person-stages are parts of the same person depends on which person-stages each person-stage regards as being related to them in the appropriate way.

However, if each and every person-stage constructs a person out of all and only those person-stages that are related to them by the relation about which that person-stage cares, won’t this create too many people? What is there to stop the arbitrary generation of more and more people: a new person for each and every person-stage? Any two person-stages which have all and only the same stages as parts will be the same person. And, although each person-stage constructs a person out of other person-stages related to them in the appropriate way, many person-stages construct the same person—they regard themselves as being identical with exactly the same set or aggregate of person-stages. Thus, for example, in Figure 1, person-stages A1 and A2 construct the same person—from both of their perspectives all and only person-stages A1, A2 and A3 are parts of them, the same whole person. A3 constructs a different person from the person that A1 constructs. For the person constructed by A1 and the person constructed by A3 have different stages as parts.

What happens when a person-stage refigures their person-directed practices and concerns is that they change their view about what counts as being related to them in the appropriate way. They change their view about exactly which other person-stages are parts of them, the same whole continuing person. For they change their view about which relation it is that unites different stages into stages of the same person. They thus construct a person out of different stages to the stages out of which an earlier person-stage or set of person-stages constructed a person of which they may be a part. That is to say, they construct a new and different person. We saw this is Figure 1 where, although person-stage A1 constructed a person of which A3 was a part, A3 constructs a person of which A1 is no part. A1 and A3 thus
differ about which person-stages are parts of them, the same whole continuing person. Refiguration, but refiguration alone, thus leads to the construction of new people. Insofar as different person-stages take a common perspective on survival, insofar as they regard the same set of person-stages as being related to them in the appropriate way, they will construct the same person. If, on the other hand, every person-stage refigured to take a new perspective on which person-stages were related to them in the appropriate way, there would be a new person for every person-stage. TPP leaves it open that there might be a new person for every person-stage (in the somewhat unlikely event that every person-stage refigured), but it gives us what is important—a principled practice-dependent principle for counting persons.

Let me put some flesh on these somewhat schematic bones, by considering what TPP has to say of Johnston's case: the case of the Somataphile who refignures to come to regard teletransportation as survival preserving, and is then teletransported. Every person-stage which has the Somataphile perspective on survival regards all and only those person-stages which are related to them by physical continuity as being related to them in the appropriate way, and so parts of them, the same whole or single continuent person. But a person-stage who takes the Teletransporter's perspective on survival (a Somataphile post-refiguration) will regard different person-stages as related to them in the appropriate way, because they take a different view on what the appropriate way is. Whereas a Somataphile will regard no post-teletransportation person-stages as parts of them (for after teletransportation there will be no person-stages which are related to them by physical continuity), a Teletransporter will regard those person-stages as related to them in the appropriate way, and so parts of them. In short, the two perspectives give rise to two different people. The situation can be represented as follows (Figure 2):
The large, filled circles in the diagram (some of which are labelled A, B, C and D) are distinct time-slices of persons or four-dimensional person-stages. A is the first time-slice, the first person-stage; D the last. B is the person-stage who refigures to become a Teletransporter, and C marks the time-slice or part of a person who undergoes the potentially survival-threatening event, in this case, teletransportation. The person-stages A to B (or just before B) all take the Somataphile perspective on survival—they regard all and only person-stages who are related to them by physical continuity as being parts of them. From their perspective, the post-teletransportation person-stages, C to D, are not parts of them, because not related them in the appropriate way. From the perspective of all those person-stages, who collectively construct a person let’s call Alpha, they will die at teletransportation. After teletransportation, there will be no person-stages who are stages of them, Alpha. Person-stage B, however, refigures to take a new and different perspective on survival: she regards all and only those person-stages that are related to her by psychological continuity as being parts of her. She decides she doesn’t much care about physical continuity. And, let’s say, all person-stages B to D take a
similar perspective on matters. Person-stage B thus constructs a new person, call her Beta, made up of all and only those parts which are psychologically continuous with her. Because person-stages B to D all take the same perspective on matters they too regard themselves as parts of Beta. From Beta's perspective, she will survive teletransportation. For teletransportation preserves the relation among person-stages that she cares about.

As I have described the situation in Figure 2, the person-stages which make up Alpha are also parts of Beta. This is because, as I am assuming, the person-stages which comprise Beta retrospectively regard the person-stages which make up Alpha as equally parts of them. They expect to remember the experiences had by those earlier person-stages; they self-ascribe the actions of those early, pre-refigured person-stages, feeling remorse, guilt, pride, shame etc for things that those earlier person-stages have done; and so on. In short, they have retrospective concern for those early person-stages. Suppose, however, that Beta did not have this sort of retrospective concern for the prerefigured person-stages—that is to say, suppose they regard those prerefigured person-stages as stages of a different person. If that were so, the situation would look like this (Figure 3):
The point to note is that even if Beta did not have retrospective concern for past person-stages—if the person-stages which collectively comprise Beta did not regard the pre-refigured person-stages as stages of them—Alpha and Beta would still share stages or parts (though fewer than if Beta did have retrospective concern for those pre-figured person-stages). For Beta exists at the point of refuguration B—B is the person-stage which first constructs a person out of the person-stages which are related by psychological continuity—but Alpha does not die at refuguration. For teletransportation (C), not refuguration (B), is the event that Alpha does not survive.

As should be more than apparent, then, (even from Figure 1) TPP allows for stage-sharing: the possibility that a person-stage might be a part of more than one person. In both Figure 2 and Figure 3 Alpha and Beta share parts or person-stages. There are some person-stages which are both a part of Alpha and a part of Beta. Some find the idea of stage-sharing counter-intuitive. And, indeed, conceived (or perhaps misconceived!) in certain ways, it probably is—if we think that what is going on is that where we usually thought there was only one person there are actually two people mysteriously floating round in a person’s head with no access to each others mental states, for example. But persons can share stages without having some sort of Multiple Personality Disorder. Consider the following diagram (Figure 4):

![Diagram of roads sharing parts](image-url)

Figure 4. Roads that share parts.
The Goulburn Road and the Hume Highway merge for a stretch before again diverging and going their spatially separate ways. How should we describe what is going on, metaphysically speaking, here? In particular, what should we say about the stretch of road in the middle where the Goulburn Road and the Hume Highway join up for a stretch? One thing we might say is that there are two roads which merge to into one road for a bit and then become two roads again. This is the way we sometimes do talk, but the problem is that, strictly speaking, it is incorrect. For it violates the logic of identity. Identity is transitive—if A is identical with B, and B is identical with C, then C is identical with A. But the above way of speaking violates the transitive character of identity. Strictly speaking, and the way we sometimes actually do talk, we might say that the section of road where the Goulburn Road and the Hume Highway merge is a part of more than one road. It is part of the Hume Highway and it is a part of the Goulburn Road. The Hume Highway and the Goulburn Road have parts in common, namely, the parts in the middle where they merge. This latter description is all that the idea of stage-sharing amounts to. It is the idea that a part (a part of a road, for example) can be part of more than one thing (part of more than one road), or, to put it conversely, that two different things can have parts in common.

This is a description of the case that both three- and four-dimensionalists about persisting things can agree with. For three-dimensionalists, like four-dimensionalists, admit that objects have spatial parts, they simply deny that objects also have temporal parts. For the three-dimensionalist, persisting things are not made up of distinct temporal parts extended over time, rather things exist wholly or ‘fully present’ at a time. Three-dimensionalists will then grant the explanation of stage-sharing in terms of the shared spatial parts of roads, but they will deny that there is any analogy to be drawn between the sharing of spatial parts and the sharing of temporal parts. For there are no temporal parts to be shared, only things wholly present at times. For the three-dimensionalist, a better spatial, stage-sharing analogy for the temporal case would need to show how two things might be wholly present in the same place. For example, where a building society and an insurance company are both located in the same building: they jointly inhabit, or are both fully present in, one building. For the three-

---

18David Lewis (1986) also explains the idea of stage-sharing in terms of intersecting roads.
dimensionalist, the temporal case will be exactly the same—two things are both fully present, not in the same place now, but at the same time.

Famously, David Lewis drew on the idea of stage-sharing to save common sense from the problem of fission. Recall, Parfit takes the case of fission to show that identity does not matter in personal identity. According to Parfit, what matters in survival is psychological continuity and connectedness. The problem is that identity is a relation with a strict formal character it is one-one (a relation that everything bears to itself and nothing else) and all-or nothing (it does not admit of degree). But relations of psychological continuity and connectedness (relation-R) need not have that formal character: they can be one-many (as in the case of fission) or many-one (as in the case of fusion), and they can hold to degrees (I may be more or less psychological connected to a future person). Lewis agrees with Parfit that what matters in survival is psychological continuity and connectedness. But he denies that this can come apart from the identity relation—the relation that holds between different person-stages when they are stages of a single continuent person. How could this be so? After all, in fission it looks like there is one person R-related to two distinct future people. Lewis’s solution is to say that there were two people there all along: that is to say, two continuent persons who shared earlier person-stages, but who, post-fission, have become spatially distinct. Let me explain.

Lewis is a four-dimensionalist. For four-dimensionalists, as we have noted, the question of personal identity is the question of what makes different temporal-stages stages of a single continuent person. Following Lewis, call the relation (whatever it is) that unites different person-stages into stages of the same person, the I-relation. The relation which holds between person-stages when and only when they are stages of the same person. Any two person-stages will be stages of the same person when they have all parts in common, no parts distinct, and there is no other person of whom they are a part. Now, Lewis has an object-level story about exactly what relation it is that unites different person-stages into stages of the same person i.e. what

\[19\text{Note that the}\ prima facie \text{problem fission poses for the claim that identity matters is equally a problem for physical continuity theorists. For physical continuity does not have the formal character of identity either. Indeed, it is a problem which besets anyone who claims that what matters is a relation that does not have the formal character of identity.}\]
makes different person-stages I-related. Different person-stages will be stages of the same person when and only when those person-stages are maximally R-related i.e. when the relation of psychological continuity and/or connectedness holds between those person-stages. According to Lewis, then, there is a non-contingent conceptual connection between the I-relation and the R-relation: wherever you have the I-relation you have a maximally R-related set of person-stages; wherever you have a maximally R-related set of person-stages you have the I-relation holding amongst those person-stages.

What, however, about the case of fission? For, in fission, isn’t there one earlier person R-related to two distinct future people (who are not R-related to each other)? Counting persons the way Lewis does, the answer is ‘no’. If a single, continuent person (or an I-related set of person-stages) is a maximally R-related aggregate or set of person-stages there were two people prior to fission as well. For, in fission, there are two, not one, maximally R-related aggregates of person-stages: two people all along, who, pre-fission, shared parts. To Lewis’s mind, fission looks like this (Figure 5):
In fission, one earlier person-stage S at t₀ is maximally r-related to two future person-stages S₁ and S₂ at t₁. (S₁ and S₂ are not, of course, R-related to each other). There are thus two continuent persons (two maximally R-related sets of person-stages) in fission, S and S₁ (C₁) and S and S₂ (C₂). The pre-fission person-stage, S at t₀ is thus a part of more than one person: it is a part of C₁ and it is a part of C₂. For Lewis, then, fission does not violate identity (or, better, the I-relation) for in fission one person does not become two, rather there were two people there all along: two continuent persons, C₁ and C₂, who, at t₀, shared parts, but who, at t₁, have become spatially distinct.

Similarly, on my account people share-stages: a particular time-slice of a person or person-stage can be a part of more than one person. The parts A to C in Figure 3 are parts of Alpha and parts of Beta. Like Lewis, I make new
people—a new person for every refiguration—but, unlike Lewis, I make them without fission.

There is a worry here, however. The worry is what is there to stop arbitrary generation of people—as many people as there are members of the power set? Lewis too has this worry. And he answers it as follows. According to Lewis, a set of person-stages will constitute a single whole person just if those stages are maximally I-related. That is to say, when they have all parts in common, no parts distinct, and there is no other person of which they are a part. My criteria for distinctness, however, is, and must be, different from Lewis's. For, in Figure 3, for example, I claim that Alpha is a different person from Beta, even though Alpha is wholly contained within Beta—even though, that is to say, Alpha has all parts in common with Beta, and no parts distinct from Beta. If Lewis's criteria for distinctness were mine, then Alpha would not be a distinct person from Beta. Moreover, *prima facie*, it look as if my criteria for distinctness must not merely be different from Lewis's, but actively at odds with it. For not only does Alpha have all parts in common with Beta and no parts distinct, but it looks as if there is another person of whom Alpha is a part, namely Beta.

My criteria for distinctness is indeed different from Lewis's, but, contrary to initial appearances, it is not at odds with it. For Lewis's claim is that you cannot have a maximally I-related person wholly contained within another maximally I-related person. But that is not the case with Alpha and Beta in Figure 3. For there we have, not an I-related person wholly contained within another I-related person, but rather an S(Somataphile)-related person, wholly contained within a T(Teletransporter)-related person. And although, on Lewis's account, you cannot have an I-related person wholly contained within another I-related person, you *can* have an S-related person wholly contained within a T-related person.

So what, then, is the criteria for distinctness I need? What is there on my account to stop the arbitrary generation of people? How can I claim, as I do, that Alpha and Beta are two distinct people, rather than merely different parts of one continuent person? Fortunately, as already noted, such a principled criteria for distinctness drops straight out of practice-dependence. Practice-dependence gives us a principled criteria for determining which
person-stages are part of the same person (and which are not): different person-stages will be part of the same person when, and only when, they share the property around which the person of whom they are parts organizes their person-directed practices and concerns. Alpha is a different person from Beta because Alpha organizes her person-directed practices and concerns around a different relation from the relation around which Beta organizes them. Alpha and Beta differ about which person-stages are stages of them, a single continuent person.

This claim—that wholly contained things can be distinct from the things in which they are wholly contained—is, I think, neither difficult to understand, nor peculiar to practice-dependence. Return to roads (Figure 6):

![Diagram of the Alpine Way and Kahncoban Rd](image)

Fig. 6. A wholly contained road.

The Kahncoban Road is a section of road which is wholly contained within the Alpine Way. Although the Alpine Way has parts which are not contained within, or coincident with, the Kahncoban Road, the Kahncoban Road is fully contained within the Alpine Way. All parts of the Kahncoban road are equally parts of the Alpine Way. There are no parts of the Kahncoban Road which are not also parts of the Alpine Way. By Lewis’s criteria for distinctness, then, the Kahncoban road is not a distinct road from the Alpine Way—for it has all parts in common with the Alpine Way and no parts distinct. Despite Lewis, locals (and some Australian road maps) talk of the Kahncoban Road, although fully contained within the Alpine way, as nonetheless a distinct and different road from the Alpine Way. They regard it as distinct because it has special historical significance and interest. It was the original road along which many famous settlers once travelled and to which,
at either end, the Alpine Way was later added. In keeping with its great historical significance are various other salient differences: the Kahncoiban Road has a gravel surface, whereas the Alpine Way (or, more precisely, all the parts of it which are not also parts of the Kahncoiban Road) is tarmac. Just as our interests and concerns make it the case that the Kahncoiban Road, although wholly contained within the Alpine Way, is nonetheless a distinct road, so the interests and concerns of the person-stages which make up Alpha make it true that, Alpha, although wholly contained within Beta, is nonetheless a distinct person. Just as our interests and concerns make it true that for a certain stretch of the Alpine Road there are two roads rather than one, so the interests and concerns of person-stages make it true that in Figure 3 there are two people rather than one.

6.4 Refiguration revisited.

How, then, does temporal-phase pluralism allow us to accommodate the intuition that there is an important sense in which, when a Somataphile (in Figure 3, Alpha) says "even were I to refigure in the appropriate way, I will not survive teletransportation", they utter a truth? It does so quite literally. For when Alpha says "even if I refigure, I will not survive teletransportation" Alpha unambiguously utters a truth. For both occurrences of 'I' in this sentence pick out the person Alpha, from whose perspective refiguration will result in her doing things, such as teletransportation, which will result in her certain death. And, when Alpha speaks, that is absolutely true—Alpha really will die at C. For after C there will be no person-stages who are related to Alpha in the appropriate way; after C there will be no person-stages who are parts of Alpha. Contra Johnston, then, Alpha is not making any mistake in thinking that she will not survive an event, such as teletransportation, which brings it about that there is a future person who is not physically continuous with her. Alpha’s life, determined by there holding between person-stages the relation which presently organizes her person-directed practices, really does end with teletransportation. For teletransportation fails to preserve physical continuity, the relation which, from Alpha’s perspective, determines which person-stages are rightfully part of her.

If Alpha determinately dies at teletransportation, however, how can temporal-phase pluralism allow for the possibility of refiguration and
subsequent teletransportation without death? Well, the crucial question is death for who? For, when a pre-refuguration (pre-B in Figure 3) person-stage, say A, utters the sentence “I will not survive teletransportation”, two people speak: one, Alpha, utters a truth; the other, Beta, utters a falsehood. For person-stage A is, timelessly, a part of two persons. It is a part of Alpha and it is a part of Beta. If we’re talking about whether or not Alpha can regard teletransportation as providing great benefits for her, the answer is ‘no’. Alpha will die at teletransportation, for the relation which Alpha cares about does not hold between the pre- and post-teletransportation person-stages. If we’re talking about Beta, on the other hand, the answer is ‘yes’. For, given that the relation that organizes Beta’s person-directed practices and concerns holds across teletransportation, Beta will survive teletransportation; and so rightly view teletransportation as providing great benefits for her.

But it might seem that this does not get us all of what Johnston wants from the possibility of refiguration. However, it seems to me it is not entirely clear exactly what Johnston wants. Sometimes Johnston writes as if what he wants is a sense in which a Somataphile, qua Somataphile, can view teletransportation as providing great benefits for them, one and the same Somataphile. But this cannot be what Johnston wants, even by his own lights. For, as Johnston himself points out, there is no way that a Somataphile, qua Somataphile, can view teletransportation as survival-preserving. Only a Somataphile, qua protean person, can view teletransportation as survival-preserving. That was the whole point of the distinction between phase sortal/cryptophase kind and substance sortal. Only if an individual’s substance sortal is protean person, not cryptophase kind, can they view teletransportation as providing great benefits for them.

Johnston doesn’t frame things in these terms but, in the light of TPP, we can understand his account in this way. When an individual utters the word ‘I’ it is ambiguous: ambiguous as between picking out an individual of a certain cryptophase kind and picking out an individual of the overarching kind protean person. For individuals are simultaneously members of a

---

20 We might, if we liked, explain things in another, perhaps less counter-intuitive sounding, way. We could say that, strictly speaking, time-slices do not say anything—they just make noise. Only persons say things, and the noise made by a time-slice serves as the truth maker for the utterances of persons.
particular cryptophase kind and members of the overarching kind protean person. So, when an individual presently of cryptophase kind Somataphile utters the sentence “were I to refigure in the appropriate way I would survive teletransportation” the word ‘I’ in this sentence is ambiguous. Disambiguated one way, the statement “were I to refigure, I would survive teletransportation” uttered by an individual of cryptophase kind Somataphile is false. It is false where the occurrence of ‘I’ is taken to pick an individual qua cryptophase kind Somataphile, whose survival conditions are given by the relation that presently organizes her person-directed practices. But an individual Somataphile is also and simultaneously a member of the overarching kind Protean person. And, when the occurrence of ‘I’ in “were I to refigure, I would survive teletransportation” uttered by a Somataphile picks out an individual of the kind Protean person, the statement is true: true, at least, on the assumption that the individual refigures before teletransportation (false, if they do not). So, that Alpha (individual qua cryptophase kind Somataphile) cannot regard teletransportation as survival-preserving is common ground between me and Johnston.

I too think that there is a sense in which—a disambiguation according to which—when person-stage A utters the statement ‘I will survive teletransportation’ they utter a truth. For when person-stage A speaks for Beta, they speak as part of person who survives teletransportation. The crucial difference is that what Johnston wants from the possibility of refiguration is not merely a sense in which a time-slice of a person is part of some person who will survive teletransportation, but a sense in which a whole single continuent Somataphile person can survive teletransportation. Johnston wants to say that there is a sense in which Alpha can survive teletransportation: namely, Alpha qua Protean person can survive teletransportation, providing she refigures before that event. I deny that there is any sense in which Alpha, the person made up of the person-stages who care about physical continuity, can survive teletransportation, whether a part of Alpha refigures or not. If a person-stage who is a part of Alpha refigures, they construct a new person; a person who will rightly regard themselves as surviving teletransportation and who will view parts of Alpha as equally parts of them. Such a person would not be Alpha, for from Alpha’s perspective Alpha will determinately die at teletransportation, but they
would regard the person-stages which are parts of Alpha as equally parts of them. According to TPP, if a person-stage refigures their person-directed practices and concerns, someone can survive teletransportation, but that someone is not Alpha. At best, that someone is a different person who regards parts of Alpha as parts of them.

And this seems exactly to capture the sense in which practice-dependence makes for a revisionary upshot. Contra Johnston, it seems to me no part of practice-dependence that a person can survive a future event that, as they now most deeply believe, they determinately fail to survive. To claim otherwise fails properly to respect the central practice-dependent tenet that a person’s survival conditions are determined by their person-directed practices and concerns, so that a person cannot survive events which by their lights they determinately fail to survive. For if survival is relative to the perspective of person-stages, and if the person-stages which collectively comprise a Somataphile person view teletransportation as certain death, there should not be any sense in which a Somataphile, by whose lights teletransportation is certain death, can nonetheless survive teletransportation. On the other hand, there does seem to be a sense in which practice-dependence really does make for a new possibility in survival. For if there is nothing served up by nature or God which constrains us to organize our person-directed practices and concerns in terms of that privileged natural relation if we are not to miss or misdescribe something intrinsically important about the way the world is—if survival is determined by our person-directed practices and concerns—then there does seem to be a sense in which we are free to organize or reorganize those concerns in terms of whatever (socially enactable) relation we think best. TPP seems to me exactly to capture the sense in which practice-dependence makes that possible, without either disrespecting or arbitrarily privileging the perspective of one person-stage or set of person-stages over another. It makes for it because, when a person-stage or set of person-stages reorganize their person-directed practices in terms of a new relation, they construct a new person from whose new perspective a new and different set of events are survival-preserving. That person—the person made up of the parts with that new perspective on survival—can then survive all and only those (new) set of events that, by their lights, they survive. But there is no reason to impose that new
perspective on other (earlier) person-stages who take a different view, and claim that it was true all along, despite what those earlier person-stages thought, that they would survive all and only those events that, according to the ‘new’ perspective, they survive.

6.5 Organizing our person-directed concerns to survive

According to Johnston, if we care about survival, at least where refiguration is in the offing we shouldn’t care very much about the relation that presently organizes our person-directed practices and concerns. For in caring about that relation we are not caring about survival, but instead about just one parochial realizer of the survival role. On my account, if we care about survival we will care a great deal about the relation that presently organizes our person-directed practices and concerns; for that relation is survival.

On Johnston’s account, if you care about survival—if you want to survive—it doesn’t make much difference which relation you presently organize around; for you will survive reorganizing your person-directed practices in terms of some new relation. On my account, however, if you really care about surviving, which relation you presently organize around makes a big difference to how easy it is for you to survive. If you presently organize around physical continuity, for example, you (the person made up of the parts who organize around physical continuity) will not survive some future events, such as teletransportation, whether you refigure or not. If you refigure, someone else will survive those events; a person who will regard parts of you as equally parts of them. But that may not be much consolation. For presumably, when you want to survive, you want you, the numerically same person, to survive: you want as many future person-stages as possible to be stages of you, not merely stages of someone else who regards you as part of them.

Sometimes we can’t have everything want we want. But this time we can have quite a lot of we want. For, for many of us (those of us who have yet to organize their person-directed practices and concerns in terms of any particular relation), even if our survival conditions are determined by the relation that presently organizes our person-directed practices and concerns,
we can get all that Johnston's Protean person view offers in the way of survival, and more. It all depends on which relation we organize around.

If we care a lot about surviving on my account (given that whether we refigure or not we will not survive events which by our present lights we do not survive) it becomes very tempting (although not, of course, obligatory) to organize our person-directed practices and concerns around whatever relation it is that makes it easiest for us to survive. That is to say, if we care about surviving, the instrumentally rational thing for us do given Temporal-Phase Pluralism is to become 'minimalists' about survival—to be very generous about the sorts of events that we view as survival-preserving.

For some of us this news will have come too late. For some (most notably, philosophers who have thought a lot about personal identity) have already organized their person-directed practices and concerns in terms of some or other particular relation which makes it quite hard for them to survive. Unfortunately, if you are one of those people, there is nothing that you can do: you will die at the events at which by your present lights you die, however generous the relation in terms of which a future person of which you may be a part organizes their person-directed practices and concerns. But, for those many philosophers and lay people who haven't as yet organized their person-directed practices and concerns in terms of any particular relation, there is still time. On the reasonable assumption that these people care about surviving, we ought to encourage them to move swiftly to internalize a minimalist conception of survival before they develop particular, restrictive commitments.

Exactly what commitments should we encourage them to take on, however? What is the relation that would, were we to organize around it, make it easiest for us to survive? Clearly, if you care about surviving, having a concept according to which you do not survive being depressed, for example, would be a bad idea. For it would probably make it quite difficult for you to survive. Organizing your person-directed practices and concerns around physical continuity would plausibly make it easier to survive—but not all that easy. You wouldn't survive teletransportation-like events. Organizing in terms of psychological continuity is better than organizing around depression, but not all that much better than organizing around
physical continuity. For you won’t survive having your psychology wiped from your brain or otherwise removed—a possibility which, as things stand, is more likely to obtain than teletransportation or such-like events. Organizing around the second-order property of there holding between you and some future person a relation which has the property of realizing the survival role for you (as discussed in Chapter 5) would not fare you much better than organizing around a first-order property such as physical or psychological continuity. For, as already noted, whether you organize around physical continuity or the higher-order property that physical has (viz. realizing the survival role for you), the extension of the concept of survival will be exactly the same—you will survive or fail to survive exactly the same set of events in either case. Just as the second-order property is a candidate relation for survival i.e. a relation among person-stages that we might organize our person-directed practices and concerns around, so Johnston’s protean person view is another object-level candidate. Johnston’s protean person view would tell you that you survive a survival-threatening event just in case you have a concept according to which you survive that event at the time of that event. However, organizing around Johnston’s protean person view—thinking that you survive a survival-threatening event just in case you have a concept according to which you survive that event at the time of that event—won’t fare you much better than any other of the above. For this concept actually makes it quite hard for you to survive. For if you haven’t refigured in time (if, unforewarned, you undergo or are made to undergo a procedure which, by your lights at the time of that procedure, you don’t survive) you die, whatever you later think about things.

Rather, I think, if you care about surviving, you should (instrumental rational ‘should’) organize around the largest possible disjunction of relations you can manage: You survive just in case there is a future person or person-stage who is physical continuous with you or psychological continuous with you or ... For if any one of the disjuncts obtain, you survive. Moreover, apropos of Johnston’s view, were you to organize around such a disjunction in the first place, you would probably never need to bother refiguring. For if you have chosen the disjuncts carefully enough, you will survive most survival-threatening events by your present lights.
Choosing to organize around such a disjunction, however, has one potential drawback. Suppose a new and radically different sort of survival-threatening comes along—an event which, although you thought long and hard to try to include a disjunct to deal with every survival-threatening event you could think of—you hadn’t thought of; a survival-threatening event which doesn’t preserve any one of the disjuncts you organize around. You just hadn’t thought of this survival-threatening event. Then, on my account, you would die. But there is another, perhaps better, object-level candidate which doesn’t make surviving so dependent on how good we are at thinking up possible survival-threatening events and including disjuncts to deal with them. I mentioned this view earlier in a footnote, as a second-order view which, unlike the second-order view mentioned above and discussed in chapter 5, is not extensionally equivalent to the first-order views. This view has it that you survive a survival-threatening event just in case there holds between you-now and the post-event person-stages a relation which might realize the survival role for a community. If you organized around this view you would, in effect, be organizing around a disjunction. But a disjunction, not merely of all the relations you can presently think of, but of all the possible realizers of the survival role. That is to say, a disjunction of all the possible survival relations. If you care about surviving, it seems to me, you couldn’t do better than organizing around this. (In chapter 8, we will get a better idea of what the class of possible survival relations for us includes.)

For those who have yet to organize their person-directed practices and concerns in terms of any particular relation, then, my account offers much more in the way of survival than Johnston’s. For, although Johnston tells you will survive refiguring your person-directed concerns in terms of a relation which allows you to survive some upcoming survival-threatening event, you have to refigure before that event, otherwise it will be too late. Forewarned is certainly forearmed. But if we organize around a minimalist concept—in particular, if we organize around this latter second-order concept—we don’t even need to be forewarned. For we will already have a concept according to which we survive all the survival-threatening events it is possible for us to survive. That is the radical, revisionary upshot of temporal-phase pluralism.
6.6 Pluralism distinguished from relativism

Johnston calls his account a ‘relativist’ account of the self. So too, Stephen White, labels his view—the view that “the personal facts about a given subject do not supervene on any set of facts which does not include facts about the attitudes and feelings of others, the conventions and practices of the subject’s society, and in some cases even the society’s level of technological development”\(^{21}\)—‘metapsychological relativism’. These labels, however, are potentially quite misleading, and it is important to avoid confusion. For pluralism, as I have explained it at any rate, is a quite different sort of view from other views which often also go under the broad heading of ‘relativist’.

Relativists standardly claim that there is no objective, absolute or universal fact of the matter. Relativists admit that there are facts of the matter. For the relativist, however, these are relative not absolute facts. For the relativist, things can be good, true, right, justified or wrong, for example, but only relative to particular individual or community standards. The relativist then agrees with the absolutist, universalist or objectivist (as against the nihilist or the error theorist) in maintaining that there are facts of the matter, but disagrees with the absolutist in maintaining that these facts are not absolute, but relative facts: facts which obtain or fail to obtain only relative to particular individual or community standards.

However, this broad characterization of relativism obscures a crucial distinction within relativism between two importantly different kinds of relativism. For there are two conceptual levels at which one might be a relativist. One might be a relativist at the meta-level—at the level of what makes it the case that a certain property or relation determines the extension of a concept—or one might be a relativist at the object-level of what properties of things a concept picks out.\(^{22}\)

Meta-ethical relativists, then, count as paradigm (meta-)relativists on this way of thinking. Meta-ethical relativists claim that what makes an act good or bad, right or wrong, is that that act counts as good or bad, right or

\(^{22}\)One could, of course, be a relativist at both levels or about both kinds of facts.
wrong by the moral standards of the particular community making the ethical judgement. There is no absolute or non-relative fact of the matter or to what makes things good or bad, right or wrong; individual and/or community customs or moral standards determine what is good or bad, right or wrong. For meta-ethical relativists, then, what makes the statement ‘abortion is wrong’ uttered by X true (if it can be properly called true or false at all\textsuperscript{23}) is that, relative to X’s moral code, abortion is wrong.

Meta-ethical relativism comes in a variety of more and less sophisticated forms. In its crudest form, appropriately known as simple subjectivism, it is the view that moral attitudes are mere expressions of personal tastes. Statements such as ‘abortion is wrong’ express nothing more than the fact that the speaker dislikes abortion and prefers that abortions not occur. According to simple subjectivism, there is no fact of the matter as to what makes abortion right or wrong, independent of the personal desires or feelings of the individual or society concerned. What makes it the case that an act is right or wrong is entirely relative to, or determined by, the preferences and desires of the individual or society concerned. Moral statements are made true or false by, or relative to, the desires and preferences of those who express them. Simple subjectivism then has the consequence that the two contradictory ethical statements might nonetheless both be true. They might both be true in virtue of the fact that both are true, because made true, by the (different) moral standards of the two individuals or communities making the statement. So, for example, the statement ‘abortion is wrong’ uttered by Ronald Reagan is true because, relative to Reagan’s moral code and the code of the conservative community to which he belongs, abortion is wrong. But, according to the simple subjectivist, the statement ‘abortion is not wrong’ uttered by a pro-chooser is also true, because true, and made true, relative to the particular moral code of the pro-chooser making the judgement according to which abortion is not wrong.

Meta-ethical relativisms, such as simple subjectivism, contrast with the broad view that variously goes under the headings of absolutism,

\textsuperscript{23} For emotivists, for example, moral statements are neither true nor false they are, literally, meaningless; expressing no more than the moral sentiments, approval or disapproval, of the speaker.
universalism or objectivism. Despite their differences\textsuperscript{24}, what unites these views is a common denial of relativism at the meta-ethical level. According to absolutists, objectivists and universalists moral statements are made true or false by facts quite independent of the personal or social preferences and moral codes of the people involved. What makes it the case that a moral term (say, ‘good’) picks out acts with a certain property, is not that that is the property which counts as good relative to an individual’s or community’s moral code, but because that is the property in which good independently and universally consists; and would continue to consist even if it were the case that no individual or community regarded it as good. ‘Good’ picks out the property of maximising happiness, for example, because maximising happiness is a fundamental, absolute human value.

However, one might be a relativist of a very different kind. One might be a relativist (i.e. maintain that there is no absolute or non-relative fact of the matter) not at the meta-level of what makes it the case that a concept picks out certain properties, but at the object-level: at the level of what properties of things a concept picks out. This kind of relativism I want to call \textit{pluralism}, to signal this important difference from relativisms at the meta-level. The pluralist maintains that there \textit{is} an absolute or non-relative fact of the matter at the meta-level, but denies that there is, of conceptual necessity, any absolute, non-relative, objective or universal fact of the matter at the object level of what properties the concept picks out. Pluralism is, if you like, a third option or a half-way house between relativism and absolutism. In the spirit of relativism, the pluralist maintains that there is an important sense in which morality, for example, at the meta-level is relative to, or, as I prefer to put it, dependent upon, a community’s moral practices. These practices—such as approving of things that are good, disapproving of things that are bad; being motivated to pursue the good, and being reluctant to pursue the bad—are what make it true that certain properties of things (the properties of things which organize these practices) get to count as the properties of things in virtue of which things with those properties count as good or bad, right or wrong, just or unjust. However, although this third option maintains, with

\textsuperscript{24}For example, in addition to asserting that there is a single, universal and independent fact of the matter as to what makes something good or bad, right or wrong, just or unjust, absolutists often make a further claim with which many objectivists and universalists disagree, viz. that some moral rules or duties are absolutely without exception.
the relativist that moral truths constitutively involve our practices in this way, it denies that this is a relative truth. With the absolutist, this third option regards it as an absolute (not a relative) truth that good, for example is that property (whatever it is) which subjects approve of, are motivated to pursue and praise, and so on. As we have seen, this absolute meta-fact is what enables us to know that communities who differ about the object-level facts are nonetheless disagreeing about the same thing or concept; that they have not changed the subject.

The pluralist about survival, then, claims that survival is, in an important sense, relative to the person-directed practices of individuals and/or communities—survival is whatever relation it is that plays the survival role for a given subject or group of subjects. But the pluralist denies that this fact—the fact that survival is whatever relation it is around which a community organizes their person-directed practices and concerns—is a relative fact. The pluralist takes this analysis to be an absolute, a priori truth about survival. There is no further question usefully to be asked about in virtue of what this is the right analysis of what determines what determines the extension of survival, in the same way as there is (plausibly) no further question to be asked about in virtue of what things with three angles and straight sides get to count as triangles. The buck stops here, for this is just what we mean by survival.

What often goes by the name of ‘normative relativism’ is a good example of such object-level relativism or, what I am calling, pluralism. The normative relativist believes that there are absolute normative facts or imperatives at the meta-level, it is just that what counts as realizing or instantiating these absolute imperatives may vary from moral context to context. It is an absolute, non-relative normative imperative, for example, that one should always be polite. It is just that what counts as being polite—the criteria for politeness—may vary from community to community, culture to culture. In some middle eastern countries, for example, burping after a meal counts as polite behaviour. But burping after a meal in most western countries counts as exceedingly impolite. Acting in accordance with the absolute norm of always being polite, then, requires burping after a meal when eating with middle easterners, but refraining from burping when eating with westerners. Always acting politely requires acting in accordance
with the criteria for politeness, and this is an absolute imperative. But, given that criteria for politeness vary, acting in accordance with the absolute norm of always acting politely requires us to act differently in different cultures in accordance with their different criteria for politeness. It requires that, when in Rome we do as the Romans do, when in the middle east we do as the middle easterners do, when in the west we do as the westerners do, and so on.

As well as the relative innocuousness of the view, there is something very important to note about normative relativism. And that is that, though it is often characterized as a form of relativism, it is not really a kind of relativism at all. For the so-called relativism embodied in normative relativism is not *apriori* or essential. It is not embodied in the normative relativist’s conceptual claim, for the normative imperative is absolute. Such relativism as may be the product of normative relativism, is simply the contingent upshot of obeying an absolute *apriori* normative imperative in light of *a posteriori* information about the criteria for politeness, or whatever, in the community in which one happens to find oneself. It is simply a contingent matter of fact that different communities have different criteria for politeness. Were all communities to have the same criteria for politeness, normative relativism would not be a relativism at all: always acting politely would require acting in acting in accordance with the universally-shared criteria for politeness, and one should act in exactly the same way wherever one happened to be. There is nothing, then, intrinsically or essentially relativist about normative relativism. That acting politely may require acting differently in accordance with the criteria for politeness of the particular community in which one finds oneself is simply the upshot of the utterly contingent, *aposteriori* fact that different communities happen to have different criteria for politeness.

The first-order account of survival I have argued for is analogous to normative relativism. It, like normative relativism, is a pluralist account. If it must be called a kind of relativism at all, it is relativism at the object, not at the meta, level. True, the account is practice-dependent in that it claims that, within certain constraints (to be further discussed in chapter 7), subjects’ person-directed practices determine what relation(s) count as the survival relation(s). That is to say, subjects’ person-directed practices make it the case
that one relation rather some other counts as the relation of survival. But this, I claim, is a necessary or absolute, not a relative, meta-fact: absolute in the sense of being a meta-fact which holds true for all people, at all times, in all worlds, irrespective of whether or not they believe it to be true. That survival is whatever property it is around which subjects have cause, consistently and without involvement in empirical error, to organize their person-directed practices, is, on the pluralist account I advocate, an absolute fact. It is just that, within certain constraints, different properties may happen to realize that role for different subjects. The survival role is, if you like, multiply realizable. But that the survival role is, if it is, multiply realized is just the a posteriori, contingent upshot of the fact that different subject's organize their person-directed practices around different properties.

I do not want to quibble about names. Call pluralism a sort of relativism if you like. But, if you call it a sort of relativism, it is important to bear in mind exactly what sort of relativism it is—it is object-level, not meta-level, relativism. And, as such, is not subject to the sorts of worries and difficulties with which the label ‘relativism’ has (rightly or wrongly) been associated. A pluralist account does not admit that two contradictory statements about survival might both be true. Nor would it have communities who organize their person-directed practices around different relations simply talking about different things. On a pluralist account, the statement ‘A Teletransporter-person survives teletransportation’, for example, is true, no matter who utters it. And the contrary statement, ‘A Teletransporter-person does not survive teletransportation’ is false simpliciter, since teletransportation preserves the relation which members of the Teletransporter community organize their person-directed practices around. A pluralist account is relativist only in the sense that it allows for multiple realizability: that what relation counts as the survival relation for communities may differ from community to community depending on what relation it is that (as a thoroughly a posteriori matter of fact) communities organize their person-directed practices and concerns around, just as what counts as behaving politely may vary from community to community in accordance with the (different) norms governing polite behaviour.
SOME POSSIBLE OBJECTIONS

The preceding discussion has presupposed a number of broader methodological issues. To discuss these issues in any great detail is well beyond the scope of this thesis. Nonetheless, in this chapter, I want briefly to discuss how these background issues might give rise to three sorts of objections to my analysis of survival, and to sketch the lines along which I think that these objections can be answered. The first two of these objections, in different ways, cluster around the accusation that this sort of account of survival is viciously circular. The third objection is that it is at odds with a reasonably popular view of identity.

7.1 The Paradox of Analysis.

There is a famous worry about the tenability of the whole project of conceptual analysis itself. As Simon Blackburn articulates the worry, attempts at conceptual analysis face an insoluble dilemma: either conceptual analyses are reductionist—they aim to analyse a concept in independent, naturalistic terms—and false (Blackburn calls this horn Scylla), or they are non-reductionist—the term to be analysed (the analysandum) reappears on the right-hand side of the biconditional as part of the analysans—hence circular, and therefore trivial (Blackburn calls this horn Charybdis). This is the 'paradox of analysis'. Since all conceptual analyses are either reductive or non-reductive, and so fall prey either to Scylla or Charybdis, Blackburn claims, we can, at best, engage only in conceptual explication. (Exactly how explication is supposed to differ from analysis is not entirely clear!)

Scylla is G.E. Moore's famous 'open-question' argument applied more broadly than merely to naturalistic or descriptive definitions of the term 'good'. Recall, the thrust of the open-question argument is that given any (reductive) analysis, Y, of a (normative) term, X, it still remains a substantial or 'open' question as to whether X is Y. So, for example, given the following analysis—an act X is good iff X maximizes happiness—it is still possible to ask in a substantial, non-trivial way, whether maximising happiness is good.
The point is that it is importantly platitudinous that good is motivational. That some act has the property of being a good act doesn’t just give us a reason to pursue it, it *motivates* us to pursue it. But why would the fact that it is an analytic truth that a certain naturalistic property is good motivate us in the least?

The worry behind Charybdis, on the other hand, might be threefold. First, it might be the worry that circular analyses, because circular, are unilluminating and uninteresting. They tell us nothing new about the concept. The analysis, X is red iff X is red, for example, imparts no new information about the concept red. Second, the worry might be that circular analyses do not enable us to get an independent fix on the concept. Suppose we want to know if it is right to call a particular object X square—we want to know, that is, whether X is square. Then, we might think, it will be little help to be told that X is square iff we judge X to be square. For we need to have some independent fix on what square is before we are in a position to know whether or not to judge an object square. Third, the worry might be that circular analyses are question-begging: that the *analysans* presupposes a prior understanding of the *analysandum*. Consider the following analysis: X is red iff agents think X is red. Clearly, however, not any old agents thoughts will do—not the thoughts of people who systematically think that green things are red, for example. So we must place constraints on the sorts of subjects whose thoughts are allowed to count. However, these constraints cannot be things of the sort: the thoughts of people who understand the concept red. For that would make the analysis trivially true. (Perhaps this third circularity worry is just a more sophisticated version of the first).

Scylla is the horn that, not in so many words, has worried many proponents of the different sorts of practice-independent accounts of survival to date. For these accounts typically define survival in straightforwardly independent, naturalistic terms. Recall, for example, physical continuity theorists offer the following kind of analysis:

A person X today survives as a person Y at some time in the future if, and only if, enough of X’s brain/body continues to exist for it to be the brain/body of a living person, and is now Y’s brain/body.

Or, for psychological continuity theorists,
A person X today survives as a person Y at some time in the future iff between X now and Y in the future there is sufficient psychological continuity and/or connectedness.

These accounts of survival makes no mention of our person-directed practices and concerns; and so, for accounts of these sort, it has been a further and open question why we should care about physical continuity or psychological continuity or some other candidate relation. For accounts of this sort leave it an open conceptual possibility that survival can come apart from what we care about. The attempts to close this conceptual gulf between survival and what we care about in survival that I discussed in Chapter 3 were attempts to answer Scylla—and pretty desperate ones at that! But there are two points to note about Scylla. First, Scylla does not seem to infect all naturalistic conceptual analyses. For example, the analysis X is square iff X has four straight sides of equal length doesn’t seem to leave any question open. Scylla seems to infect only those of analyses of normative terms: terms where there is a further question to be asked about in virtue of what do those properties—in this case, the property of having four straight sides of equal length—get to count as the properties in virtue of which a thing counts as square; in particular, terms where the answer to this third question seems to require reference to our practices and concerns. (It may be that proponents of Scylla take rule-following considerations to show that all concepts, including squareness, are normative.) Second, and like the case of squareness, it may be a quick but fair response to Scylla to say that, given the right analysis of a concept X, however naturalistic that analysis, there will simply be no question left open. If there is an open question, it simply goes to show that we haven’t got the analysis quite right yet (not that we never could get the analysis right). Nonetheless, despite the availability of these quick responses, Scylla has worried practice-independent theorists; and it has worried them because they take it to be importantly platitudinous that we care about survival—that survival is, in this important sense, a normative concept.

Scylla might infect my analysis too. For like practice-independent accounts, mine too aims at conceptual (and ontological) reduction. I too want to reduce survival to quite independent, naturalistic properties and relations among person-stages. On the other hand, Charybdis might also seem to be a worry. For, unlike practice-independent theorists, I define survival in terms
of its being a relation among person-stages which plays a certain role in our
person-directed practices; and mightn’t that be viciously circular or trivially
true if, for example, our person-directed practices presuppose survival?
However, I think, my pluralist account of survival avoids both Scylla and
Charybdis. For survival consists in a quite independent, naturalistic set of
relations among person-stages (answer to Charybdis), but it consists in those
relations just because those are the relations that we care about (answer to
Scylla). But I need to say a bit more.

7.2 Charting a course between Scylla and Charybdis

Recall, I defined survival as follows:

‘A person X at an earlier time survives as a person Y at a later time’ as uttered
by us is true iff (i) there is some property or set of properties p which cause
us, consistently, and without involvement in empirical error, to organize our
person-directed practices around the assumption that X survives as Y, and (ii)
X and Y share p. (Where ‘us’ refers to subjects or groups of subjects with
common person-directed practices).

In other words, survival for a community is whatever property or relation it
is between two person-stages, X and Y, that plays the survival role in that
community’s person-directed practices. The role that survival plays is fixed
by the platitudes about the conceptual relationship between survival and the
various other concepts—punishment, self-concern, moral responsibility,
agency and the like—with which it is connected. So the first stage in
analysing survival consists in teasing out the platitudes—in making explicit
the conceptual dependence and interdependence of survival with the other
concepts. So, for example, the platitudes listed included ‘You should punish
a person for a past crime when, and only when, ceterus paribus, they are the
same person as the person who committed the crime’; ‘normally, if someone
remembers some experience, they are the same person as the person who had
that experience’; ‘if the person before you now is the same person as your
loved partner, then, normally, you will love that person before you now’;
‘hold someone to a prior promise when, and only when, they are the same
person as the person who made the promise’; and so on and so forth.

Note that even if the analysis finished at this first stage (and it does
not) we would have an answer to one of the worries about Charybdis—the
worry about how it could be that an analysis, although circular in the sense that the term to be analysed appears in the platitudes that form the \textit{analysans}, might yet tell us something new and interesting about the concept of survival. For this sort of analysis, by teasing out the connections between survival and the concepts with which it is intertwined, makes explicit what was previously merely implicit in our use of the concept. It tells us something new and interesting about the concept of survival because it makes us consciously or explicitly aware of the conceptual dependence and inter-dependence of the concept of survival with various other of our concepts. It gives us ‘knowledge that’, where previously we had only ‘knowledge how’\textsuperscript{1}; or, to borrow a distinction from Philip Pettit\textsuperscript{2}, it gives us ‘intellectual’ belief where previously we only had ‘practical’ belief. It is not, however, as the term ‘intellectual’ may misleadingly suggest, that this analysis gives us a special new sort of knowledge or belief about survival. Rather, it points explicitly to an interesting connection among certain of our concepts: a connection which was previously implicit and unnoticed in our practice.

But the analysis of survival I outlined did not remain conceptually non-reductive for long. For, recall, I went on to point out that once we have seen how the concept of survival is connected to other concepts—once we have fixed the role that survival plays in our person-directed practices—we can dispense of talk of ‘survival’ in our analysis. For this conceptual work at the first stage tells us that $X$ at $t_1$ and $Y$ at $t_2$ will be stages of the same person (or $X$ will survive as $Y$) for a community, just in case a community punishes $Y$ for crimes committed by $X$; $Y$ is held to promises made by $X$; $Y$ expects to remember the experience of $X$; $Y$ is the object of certain persisting reactive attitudes—trust, love and the like; $X$ anticipates $Y$’s future experiences; and so on. This second stage of conceptual analysis thus removes all occurrences of the term ‘survives as’ or ‘is the same person as’ from the \textit{analysans}. We can get a fix on what relation is the survival relation for a community without making any mention of ‘survival’ on the right-hand side of the analysis. The

\textsuperscript{1}This is a slightly unhappy way of putting the point, since some of our person-directed practices, however unconsciously embedded in our practice, seem to involve propositional knowledge or ‘knowledge that’—for example, our practices of making judgements that $X$ is the same person as $Y$.

\textsuperscript{2}Pettit, forthcoming.
analysis as I described it proceeding in Chapter 5 was thus rendered non-reductive at the conceptual level. (Of course, if the concepts of punishment, self-concern and the like which figure on the right-hand analysis themselves somehow inextricably presuppose survival then this reduction would fail. The next section suggests why this is not the case, and why it might not be so bad even if it were).

The final stage of the analysis aims at ontological reduction. We want to find out what relation survival is for a given community. And we find this out by observing communities and/or presenting them with various puzzle cases and a posteriori ascertaining which relation among person-stages their person-directed practices track. That is, by noting which relation it is in accordance with which they punish Y for crimes committed by X; in accordance with which Y is held to promises made by X; in accordance with which X and Y are treated in certain ways—trusted, loved and the like; in accordance with which X anticipates Y’s future experiences; and so on. And then that relation counts as survival for them. Survival is psychological continuity for the Teletransporters, for that is the relation which their person-directed practices track; physical continuity for the Somataphiles, for that is the relation around which their person-directed practices are organized.

How, though, does this sort of reduction avoid Scylla? It does so because, on this analysis, there simply is no open question about why we should care about physical continuity or psychological continuity or whatever other relation is the survival relation for us. For there is nothing left open. My analysis mentions (or would mention when complete) all of the platitudes about survival. What more could there be? So it is, on this analysis, conceptually impossible that survival might not be a relation that we care about, for it is a priori true that survival is whatever it is that we care about. Physical continuity or psychological continuity or some other independent relation out there in the world gets to count as the relation of survival for a community just because that is the relation that they care about. By thus closing the conceptual gulf between survival and what we care about in survival, pluralism shuts the door on the open-question argument. Survival (whatever property it, a posteriori, turns out to be) is ipso facto what

---

3In effect, this second stage of the analysis consists in ramsifying over the platitudes.
we care about. This normative component is enshrined in the conceptual claim.

However, perhaps I haven't yet fully answered Charybdis. For I haven't said much directly to the worry that our person-directed practices might themselves presuppose survival—that we cannot get a fix on our person-directed practices independently of, or without first having a fix on, survival—so that, pluralism, although true, would be trivially so.

7.3 Do our person-directed practices presuppose survival?

Charybdis is a worry about the possibility for conceptual reduction. The worry is that we cannot understand the various concepts of person-directed practices without first understanding the concept of survival. Survival is conceptually prior to concepts of our person-directed practices. We cannot know whether or not to have self-concern for some future person without first knowing whether or not the future person is us. The claim here will be that my second-stage of conceptual analysis is not really reductive. This amounts to insisting that the direction of conceptual determination goes the practice-independent way. It is because someone is the same person as us that we anticipate their future experiences. It cannot be that someone is the same person as us because we anticipate their future experiences. For personal identity is conceptually prior to self-concern.

It is not clear to me why this sort of conceptual circularity, in itself, should be particularly worrisome. If it turned out to be \textit{a priori} true that the concepts of self-concern and personal identity were so inextricably interdefined, so that we couldn't understand one without understanding the other, why would this be such a problem? It would have just turned that the right account of some concept has it conceptually interdependent on others. And why would that be so bad? For, as we earlier noted, such analyses, though circular, nonetheless tell us something new and interesting about the concept in question—they tell us about the conceptual relations between the concept to be analysed and various other concepts. They make explicit what was previously only implicit in our employment of the concept. Of course, conceptual circularity might well be a worry if there were no way to get outside the circle. If there were no way, that is, to get some independent fix
on the concept. But, in the case of survival, even if it turned out that it was impossible to define self-concern independently of survival, there are platitudes which lead outside the circle—platitudes, for example, about the behavioural dispositions associated with self-concern (bracing yourself, rehearsing stoic platitudes, and the like). These platitudes enable us to get an independent fix on the concept. Here is an analogy. Suppose we define ‘husband’ as follows: a husband is someone who is married to a wife; and a wife is someone who is married to a husband. The terms ‘husband’ and ‘wife’ are interdefinable. This is circular, but not viciously so. For we can get an independent fix on the meaning of the terms 'husband', ‘wife' and 'is married to'. A husband is (at least traditionally) a male, a wife is a female, and we define marriage in terms of a certain ritual ceremony—two people of opposite sexes uttering certain words before an authorized celebrant and other witnesses.

However, for those who think that this sort of conceptual circularity is a problem, and who accept the standard response to this worry by proponents of psychological continuity accounts of survival in response to this sort of objection, an analogous line of response is available in this case: a response which ought to prove congenial to those many who accept it in the case of memory and the like.

For, famously, the same objection besets psychological continuity accounts of survival which analyse personal identity, *inter alia*, in terms of experiential memory. Of such accounts, Bishop Butler once objected that it is “self-evident that consciousness of personal identity presupposes, and therefore cannot constitute personal identity, any more than knowledge in any other case, can constitute truth, which it presupposes.”4 Butler’s objection was that it was a conceptual truth that a person could only ‘remember’ past experiences when they were the same person as the person who had those experiences. Analysing survival in terms of memory would then true, but trivially so, since in order to know whether or not someone remembers something we need first to know whether or not they are the same person as the person who had the remembered experience, and that is exactly the question that the memory criterion was supposed to answer.

---

4Butler (1736), p. 100.
Following Sydney Shoemaker, psychological continuity theorists have typically responded to this objection by defining a wider concept of which our ordinary experiential memories are a subset. They think as follows. In ordinary, everyday life when we have a genuine memory of some past experience i.e. a memory that is causally dependent on the past experience in the right kind of way, we just assume that the memory is of something that we ourselves experienced. For in ordinary life, we don’t, as a matter of fact, genuinely remember from the inside the experiences had by others. So in ordinary, everyday life memory does, as a matter of fact, presuppose personal identity. But it need not do so. We can define a wider concept, quasi-memory (q-memory, for short), of which our ordinary memories are a subset. Quasi-memory is memory which has exactly the same phenomenology as our ordinary memory, it is just that in quasi-remembering some experience we cannot automatically assume, as we do in ordinary-remembering, that the remembered experience is of something that we experienced ourselves. Quasi-memory feels exactly like ordinary memory from the inside, it is simply that in quasi-remembering things seen or done or experienced we cannot assume, as we do ordinarily, that we ourselves saw, or did, or experienced, the remembered things. Parfit offers the following example.

Venetian Memories. Jane has agreed to have copied in her brain some of Paul’s memory-traces. After she recovers consciousness in the post-surgery room, she has a new set of vivid apparent memories. She seems to remember walking on the marble paving of a square, hearing the flapping of flying pigeons and the cries of gulls, and seeing light sparkling on green water. One apparent memory is very clear. She seems to remember looking across the water to an island, where a white Paladian Church stood out brilliantly against a dark thundercloud...What should Jane believe about these apparent memories? Suppose that, because she has seen this church in photographs, she knows it to be San Giorgio, in Venice. She also knows that she has never been to Italy, while Paul goes to Venice often. Since she knows that she has received copies of some of Paul’s memory-traces, she could justifiably assume that she may be quasi-remembering some of Paul’s experiences in Venice.5

Q-memory solves the circularity problem for psychological continuity theories by severing the ordinary conceptual connection between memory

and personal identity. Ordinary memory thus consists in two, in principle, separable components: q-memory and personal identity. Ordinary memory is q-memory plus personal identity—memory of things which we happen ourselves to have seen or done. But personal identity is defined in terms of q-memory. According to psychological continuity theorists who adopt this line of response (and that is most), there is no necessary or essential interdependence between the concept of q-memory and the concept of personal identity—we might q-remember from the inside experiences had by someone else.

The pluralist might borrow this leaf from the psychological continuity theorist’s book in reply to the worry that self-concern presupposes survival. Just as it is no necessary part of (q-)memory that we can remember only our own experiences, so we may think it is no necessary part of self-concern that we can have this concern only for our own future self. The pluralist might similarly answer the objection that self-concern presupposes survival by defining a wider concept—call it q-concern—which, although it has the same distinctive phenomenology as our anticipatory self-concern, does not presuppose that only our own future person-stages can be the object of such concern. Q-concern, we can say, feels just like the anticipatory concern we ordinarily have for our own past and future selves alone—we anticipate a person’s future actions and experiences—but we cannot automatically assume that when we have that sort of concern for some future person-stage that future person-stage is our own. Parfit, Shoemaker and others who claim that identity is not what matters in personal identity, must themselves believe in something like q-concern. For they themselves claim that we can have concern for future people who are not us—as, for example, when we are ‘self-concerned’ for what befalls both of the two fission products.

We might even aim for a thoroughgoing ontological reduction of q-concern: a naturalistic explanation of the special, anticipatory phenomenology of q-concern. As John Perry has noted, there is direct connection between indexical beliefs (beliefs about ourself) and action; a connection which is not there for the connection between non-indexical beliefs and action. Perry offers the following as just one example:
Chapter 7  Pre-empting some possible objections

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch...I believed at the outset that the shopper with the torn sack was making a mess. And I was right. But I didn’t believe I was making a mess. That seemed to be something I came to believe. And when I came to believe that, I stopped following the trail around the counter, and rearranged the torn sack in the cart. My change in beliefs seems to explain my change in behaviour.6

Presumably, then, there will be certain distinctive sorts of processes in the brain which underlie this peculiarly direct connection between indexical belief and action. So, in principle, we could get a thoroughgoing reduction of the phenomenology of self-concern to neurophysiological processes in the brain.

However, although q-concern (by characterizing self-concern in terms of its distinctive phenomenology) is a fair response to one sort of circularity objection—the worry that we cannot get a fix on what self-concern is independently of, or without first, having a fix on what survival is—q-concern may not answer another sort of circularity worry—the objection that q-concern is insufficient for self-concern. According to these objectors, self-concern is not simply q-concern, for, by definition, we may have q-concern for future person-stages who are not stages of us. Self-concern, the sort of concern that is relevant to survival, is q-concern plus (independently determined) personal identity. Call this concept, s-concern. Then a person now survives as a future person if and only if subjects have s-concern for that future person. But, the objection will go, s-concern presupposes survival, and the analysis, though true, is trivially so. (The same objection may equally go for the q-memory manoeuvre.)

But I can agree that q-concern is insufficient for self-concern. I could even grant that perhaps self-concern does require for its application the prior belief that the future person is us (although in fact I think it does not). For, survival is a cluster concept, and although self-concern is part of the cluster, it

may, on its own, not be enough. On my view, we come to acquire the belief that someone is the same person by looking not just at the relation in accordance with which whose future experiences we anticipate, but also at who gets our legal entitlements, who is identified and reidentified as who, who we punish, and so on. Which future person is us is not wholly transparent to us a priori, as on the Cartesian view of things. Rather, we come to believe that someone is the same person as us when they have the relation that our cluster of person-directed practices (or enough of them, anyway) collectively track.

7.4 Identity and personal identity

For the most part, the personal identity debate has been conducted quite independently of debates about the identity of artifacts or things over time in general. Most philosophers believe that strict identity is absolute: that the identity conditions of an individual are not relative to the sortal under which we classify it. The central intuition behind the absolute identity thesis is that no matter how we classify an individual, say a table—whether we classify it as a table, or a lump of wood, or a middle-sized object—there is still only one thing there, and so only one set of identity conditions. The alternative, but much less widely-endorsed view, is that identity is sortal-relative. A particular table may have different identity conditions qua table, than it does qua middle-sized object, or qua lump of wood. A table might be the same table, without being the same lump of wood or the same middle-sized object. Depending on how we classify it, an individual can have more than one set of identity conditions.

Some may think that the case of personal identity is just a special case of identity. Personal identity is in exactly the same conceptual boat as the identity of artifacts over time. These philosophers might be bemused by the way I have gone about addressing the question of personal identity: shouldn’t I just have discussed the issue of identity, and then cranked my favourite theory of identity through for the case of personal identity? If identity is absolute, so is personal identity. If identity is relative, so is personal identity relative. Whatever goes for identity, similarly goes for personal identity.
The reason I have not addressed the question of personal identity in this way is because, like many others in the personal identity debate, I think that personal identity is not a special case of strict identity. From my point of view, the question of personal identity is not the question of what makes two different person-stages or temporal-parts identical. For the answer to that question is nothing—for different person-stages are not identical. Identity is a relation that a thing bears to itself, and to nothing else. Aggregates of person-stages (or whole persons) can be identical to themselves: one and the same aggregate of person-stages. But different temporal parts cannot be identical, for they are different—different things cannot be one and the same thing (although they can be parts of one and the same thing). The question of personal identity is thus, strictly speaking, not the question of identity.

Although I think that the question of personal identity is not the question of what makes different person-stages identical, but rather of what unifies different temporal-parts or person-stages into stages of the same aggregate person, it is worth noting that often the question of personal identity is posed in a way that leaves it ambiguous as to which of these questions the question of personal identity is. The question I posed at the beginning of Chapter 1—what makes a person identified at one time the same person as a person identified at another—was ambiguous in this way. Some philosophers who have talked about personal identity, such as Saul Kripke and David Wiggins, are clearly interested in the question of identity: in the question of what makes two person-stages identical. For them, the answer to the question of personal identity drops straight out of a general theory of identity. But these philosophers seem to be in a minority. Most philosophers in the personal identity debate either explicitly reject the idea that the question of personal identity is the question of what makes different person-stages identical (for personal identity is not identity) or address the question of personal identity without making any mention of identity in general, as if it were a different sort of question altogether. I take it then that most philosophers of personal identity are interested in the question in which I am interested: namely, what makes different person-stages stages of the same continuing person or, as four-dimensionalists prefer to put it, what makes different temporal parts of a person parts of the same whole. And, for the
reasons noted, this is not the question of identity, but rather of similarity or some such looser relation among person-stages.

Of course, this is a four-dimensionalist way of thinking about the question of personal identity. From the three-dimensionalist's point of view, the question of personal identity is the question of what makes a person fully present at one time the same person as a person fully present at another. But three-dimensionalists might similarly think that the question of personal identity is not the question of identity. They might think that there is a person fully present at one time and a person fully present at another, but that the person fully present at one of those times is not, strictly speaking, identical with the person fully present at the other time. Identity does not hold between these two people fully present at different times, rather some sort of similarity or counterpart relation holds between them.

However, although, three-dimensionalists can say this, many of them I suspect will not. Although there is no necessary connection between being a three-dimensionalist and believing that personal identity is a special case of identity, those who believe that personal identity is identity are typically three-dimensionalists. For, as we have seen, if we are four-dimensionalists there is good reason to think that, strictly speaking, personal identity is not identity. We think of persons as aggregates of different temporal parts (or as different temporal parts linked by a similarity or counterpart relation); and different things cannot be identical. On the three-dimensionalist picture there is no such simple reason, for there are no different temporal parts, but only people wholly present at times.

Now, nothing in practice-dependence—the thesis that what makes a relation among person-stages earn the name 'survival' is that that is the relation about which subjects care—is strictly incompatible either with three-dimensionalism or with thinking of personal identity as a special case of absolute identity. The thesis will then be that what makes a person fully present at one time strictly identical with a person fully present at another time is the person-directed practices of the community involved. Indeed this, I think, is exactly Mark Johnston's view. However, although Johnston seems happy to embrace this combination of views, it should be noted that practice-dependence in combination with the view that personal identity is absolute
identity commits one to a fairly strong brand of anti-realism that others may find less savoury. For then strict numerical identity turns out to be determined by cultural practices.

What about temporal-phase pluralism (TPP)? Could, contrary to initial appearances, an absolute identity theorist about personal identity endorse that? The short answer is that they could, but they would have to think of person as a phase kind, rather than as a substance sortal. An individual's substance sortal gives the unique set of identity conditions for that individual. An individual can survive all and only those changes that are not at odds with its substance sortal. So, for example, if the substance sortal to which persons belong is *child*, then persons would cease to exist on their thirteenth birthday (assuming that is when childhood ends). It seems fairly implausible to think that *child* is the substance sortal to which persons belong. We do not think that persons literally cease to exist upon turning thirteen (although Johnston claims to be prone to a recurrent scepticism about this!) *Child* then seems not to be a substance-sortal, but a phase kind. A stage in a person's continuing existence which that person can live through. Childhood, adolescence, middle-age, old-age and so one are phase sortals, changes which, compatible with their substance sortal, a person can live through.

In order to endorse TPP, absolute identity theorists about personal identity need some way of understanding stage-sharing—of understanding how stages of individuals can overlap. Substance sortals cannot overlap—at any given time an individual can be a member of only one substance sortal, and that substance sortal gives the unique set of identity conditions for that individual. Since, according to TPP, persons overlap, persons cannot be a substance sortal. But they can be a phase-kind, for phase-kinds can overlap. An individual can simultaneously be a member of more than one phase-kind. This might be so, for example, if childhood did not end until age thirteen, and adolescence began at age eleven. Then a person aged twelve would simultaneously be a member of both the phase-kind, *child*, and the phase-kind, *adolescent*. For absolute-identity theorists to endorse TPP, then, they would need to think of *person* as a phase-sortal (like childhood and adolescence). The substance sortal to which persons belong would be whatever overarching kind it is that strings persons (or phase-kinds)
together; some over-arching property which the different phase-kinds have in common which explains how an individual can continue to exist despite changing its phase kind. It might, for example, be human animal. Or, more in tune with practice dependence, it might be the second-order property which I earlier claimed we should all be organizing around: there holding between individuals at different times a relation which could play the survival role. Whatever exactly our substance sortal should be, however, those who think that personal identity is a special case of absolute identity can endorse TPP by conceiving of person as a phase-kind. Despite initial appearances, then, neither TPP nor practice-dependence about survival are necessarily at odds with viewing personal identity as a special case of absolute identity.
A CONSTRAINED PLURALISM?

On a pluralist account of survival, a central and interesting analytic task becomes charting the limits on the kinds of properties or relations which could play the survival role; on charting, that is, the boundaries of the range of possible natures that survival could have.

Some may view this task as all the more crucial because, although pluralism avoids the worries many have about relativism at the meta-level—at the level of what makes it the case that survival consists in certain properties—some may still have reservations about such ‘relativism’ as is embodied in multiple realizability. For, they may think, if pluralism is to be plausible, it cannot allow that the concept of survival might quite correctly pick out just any old bizarre property. My crushed and broken body lying lifeless at the bottom of a thousand foot cliff, for example, surely does not count as my survival, whatever I or anyone else am inclined to belief, or judge, or desire about it.

Indeed, one of Wiggins’ chief arguments for conceiving of persons as a natural kind is founded on exactly this sort of worry. For, Wiggins thinks, if persons are not a natural kind, then they must be a social kind; and this would be to pave the way for unacceptably authoritarian and totalitarian conceptions of what is and what is not a person. As Wiggins writes,

Suppose that we dispense with the animal component in the animal attribute elucidation of person and assert cheerfully that specification by some finite number of attributes, fixed \textit{without} essential reference to human beings just as they are, captures every mark that is essential to the concept of person. What we then have to imagine is a state of affairs in which all conceptual constraints whatever (including moral or normative constraints of conceptual provenance and important discriminations of importance) will be founded in the finite systemic specification of person; and a state of affairs in which nothing prevents the interpretations of speech, conduct and the thoughts that lie behind the desires that lie behind conduct, from being founded in a conception of personhood perceptibly narrower and certainly
simpler than that implicit in our present (however conservative and obscurantist) mode of interpretation of these things.¹

Wiggins' worry is that if what counts as a person is entirely up to us, then there is nothing to stop us developing a set of criteria for personhood which arbitrarily excludes certain individuals (Jews, blacks, the old, etc) from counting as persons. Wiggins, it seems to me, is wrong about this. For his functional or systemic requirement alone (i.e. without the animal component) is sufficient to rule out such arbitrary conceptions as acceptable. (Jews, blacks and old people 'perceive, feel, remember, imagine, desire, make projects, move themselves at will, speak, carry out projects, acquire a character as they age, are happy or miserable, are susceptible to concern for members of their own or like species...[note carefully these and subsequent dots], conceive of themselves as perceiving, feeling, remembering, imagining, desiring, making projects, speaking...have and conceive of themselves as having, a past accessible in experience-memory and a future accessible in intention..., etc.'; or they do more than enough of these things to count as persons at any rate). But even if a pluralist analysis were not to rule out such concepts as metaphysically possible person concepts (or, more precisely, in our case, survival concepts), how bad would that be for a pluralist analysis? If it turned out that pluralism were to allow for some weird and wonderful survival concepts as possible survival concepts, would that be so bad? Pluralism is a semantic-cum-metaphysical thesis—a thesis about what relation survival consists in for a community—not a moral thesis.²

²Of course, what relation counts as survival for a community will typically depend, inter alia, on that community's moral theory. For some of the platitudes that fix the reference of 'survival' in a community are moral ones—platitudes about moral responsibility and the like. And it may be that members of one community regard the moral views held by another community as mistaken. That is to say, they are practice-independent about morality. Compatible with that, however, they may regard the (from their point of view, mistaken) moral views held by that other community as nonetheless serving to fix the reference of survival for that other community. For example, a community may take it to be an absolute, practice-independent moral truth that you should only punish a person for a past crime when they are psychologically continuous with the person who committed the crime. Such a community will regard another community who punish a person for a past crime only when they are physically continuous with the person who committed the crime as having a mistaken view about responsibility. Nonetheless, they can admit that that mistaken moral view fixes the reference of survival in that other community. What is important as regards survival is not that other communities have the right moral view, but what others regard as being the moral truth.
tells us what survival concepts are possible; moral theory tells us what survival concepts are morally permissible. And it may be that pluralism allows some survival concepts as possible, which our preferred moral theory tells us we ought not organize around. So that pluralism may allow for some weird and wonderful, perhaps even morally repellent survival concepts, would not be an objection to pluralism. (What would be bad for the pluralist is if it were to turn out that there were actual concepts which the pluralist said were (metaphysically) unacceptable, but which we thought were perfectly valid survival concepts.) In any case, how constrained is pluralism?

If pluralism is to be an interesting, plausible, and substantive thesis about a given area of discourse, the constraints on pluralism must satisfy three conditions: they must be non-trivial; they must be able to be fixed upon a priori; and they must be principled or non-arbitrary.

They must be non-trivial because otherwise practice-dependence would reveal nothing interesting about the nature of the concept in question. For it could be made true of just about any concept, even of those where subjects responses seem only contingently, if at all, associated with the concept. For example, given a trivial specification of the constraints on subjects' responses, we could be practice-dependent about square: X is square iff X would be judged to be square by observers who are infallible in identifying square things in circumstances which are such that they will never fail so to do. But this analysis captures nothing distinctive about the concept of squareness. For if the analysis X is $\theta$ iff X... holds true for absolutely any $\theta$, it can scarcely tell us something interesting about any particular $\theta$. And indeed, for many, the appeal of practice-dependence for certain of our concepts—the idea that certain of subjects' practices are constitutive of these concepts—is grounded in precisely the intuition that these concepts seem to implicate the subject and subject's responses in a way in which primary quality concepts, like squareness, do not. So, if pluralism is to be interesting and substantial thesis about a given concept the analysis of the concept must implicate subjects in a non-trivial way.

And the constraining conditions must be principled. They must fall directly out of the pluralist's conceptual claim itself. The pluralist can only help herself to constraints which drop out of the claim that for a property or
relation to be the relation of personal identity or survival for a given group of subjects is for it to be the property around which those subjects have cause to organize their person-directed practices in the specified conditions. She cannot appeal to a view of persons as natural kinds, or any other such independent facts.3

As it stands, the proffered analysis is already explicitly somewhat constrained: on one side, by facts in the world, and on the other, by constraints placed on subjects practices. The analysis stipulates that in order for some property to count as a property of personal identity it must exist, it must (causally) underpin our person-directed practices (in the right kind of way), and two person-stages must share it. And it requires that we may only take subject's person-directed practices to determine the property which is personal identity or survival for them when those subjects have a consistent and empirically informed set of beliefs about personal identity. It is also constrained by pluralism itself. In order for a property or relation to count as an acceptable candidate relation for personal identity it must be one around which subjects could organize their person-directed practices. As it stands then, the analysis is already constrained. But exactly how constrained is it? What sort of concepts are acceptable concepts of personal identity? More to the point, what concepts of personal identity are possible concepts for us? If we are free to organize our person-directed practices and concerns in terms of a relation that makes for a better life, what are the relations in terms of which we are free to organize them?

3 As, for example, Mark Johnston (1989c) appears to do. Johnston stipulates as part of his relativist analysis that “The members of a given community C would be correct to take identity restricted to K's to be the relation of personal identity for them iff (i) K is a more or less natural person-kind whose non-defective members never exist without the capacity for reflective mental life...” (p.457). According to Johnston, clause (i) is required to ensure “that the members of a given community could not be correct in taking personal identity for them to be absolute identity restricted to a kind whose instances have periods when they are ants, stars, trees, gases, rivers, or any other type of thing without the capacity for reflective mental life.” (p.457). However, it is not at all clear that this is something that an avowed relativist like Johnston is entitled simply to stipulate. Johnston claims that there are no absolute facts about personal identity: that the relation of personal identity is just that relation around which subjects organize their person-directed practices.
8.1 The Contacti

A good way to begin to explore these possibilities is by way of a case, first presented by Hirsch, of an imaginary community, the Contacti, who appear to possess what, from our point of view, looks like a completely bizarre concept of survival. There are two questions to ask about the Contacti. First, is this apparently bizarre concept a possible concept of survival i.e. is it the sort of concept around which a community could successfully organize enough of their person-directed practices? Second, is this bizarre concept a possible concept for us—is it the sort of concept around which we could organize our person-directed practices and concerns?

The Contacti are a community who have a concept of personal identity according to which when two people come into contact the two people, as we might describe it, ‘swap bodies’ for the period of contact. If two people, A and B, are touching, the Contacti regard the sentence ‘the person A who used to be associated with body-A is now associated with body-B, and the person B who used to be associated with body-B is now to be associated with body-A’ as true. When contact is released, the association of people with bodies reverts back to the way it was before contact. So, when contact is released, the Contacti say ‘The person A who was in the A-body before contact and who became associated with body-B during contact has now become associated with body-A again’ and ‘the person B who was associated with body-B before contact and who became associated with body-A during contact is now in body-B again’. (To keep things simple let’s imagine that the Contacti never touch more than one person at any one time.) The Contacti do not believe that touching actually results in the transmission of characteristics from one person to the other. Nor, in fact, does it. Upon touching a person, although you become that person, and that person you, for the period of touching, touching does not actually result in the transmission of characteristics. The body that you touch does not inherit your memories, character traits, beliefs, desires, physical abilities, appearance, expertise, intentions, and so on. Nor do the Contacti believe that people are non-material entities which mysteriously swap brains or bodies upon contact. (Nor, in fact, are there any such non-material entities in the Contacti’s world.) It is just that the Contacti have a concept of personal identity which, unlike
ours, tracks personal identity with touching when two people are in exclusive contact, otherwise (when people are not touching anyone) just as we do. In general, if A is a person who throughout their life touches two other people B and C, then A's life history includes all of A's life during the period when A was not touching anyone, plus the part of B's life history for the period when A was touching B, and the part of C's life history for the period when A was touching C. As we might represent it four-dimensionally:

But, as Hirsch sets up the example of the Contacti, although the Contacti apply the language of personal identity in this strange way, they do not actually organize their person-directed concerns and practices around the concept of exclusive contact. For their person-directed practices are not at all as we would expect them to be were they actually organizing them around the relation of exclusive contact. Indeed, as Hirsch describes the case, their person-directed practices and concerns are just the same as our own. For example, they are not at all reticent to touch a person who is in extreme pain, as we would expect them to be were they organizing their attitudes of self-concern around the Contacti concept. For their Contacti concept would tell them that were they to touch a person in pain, they would become the person in pain. Nor do they expect to remember directly the experiences had by a person they touched during a period of exclusive touching. Nor do they possess bizarre expectations about the relations between actions and desert. For example, if one of the Contacti, upon deciding to commit murder, grabs another Contacti person just before pulling the trigger, the Contacti charge the same person as we would for committing murder. Namely, the person (physical body) who pulled the trigger. Not the person who, by their Contacti concept, is the same person as the person who committed the crime.
So, as Hirsch sets up the example, though the Contacti talk as if the Contacti concept was their concept of personal identity, in fact it is completely inert as regards all their person-directed practices and concerns.

As Mark Johnston quite rightly points out, then, as it stands the Contacti concept, though indeed an alternative concept, is not in fact their concept of survival. Johnston writes,

In general, the concept of being the same Cp as used by the Contacti does not seem to be their concept of personal identity just because they use that concept neither to guide their future-oriented and retrospective concerns, nor to shape their expectations about the relationship between earlier action and later desert of praise and blame, nor to focus their patterns of anticipation of experiences and memories of those experiences.4

But the pressing question for a pluralist in considering whether or not the Contacti relation of touching could count as an acceptable alternative relation of survival is not whether or not a community does in fact organize their person-directed practices and concerns around the relation of exclusive contact, but whether or not a community could successfully organize their person-directed practices around the Contacti relation of touching. And here both Hirsch and Johnston are united in maintaining that the Contacti concept is not the sort of concept around which a community could successfully organize their person-directed practice and concerns. As Hirsch writes,

...such a language must lead to total havoc. For one of the essential functions of our concept of the self is to enable us to retain our separate identities in the course of complex social interactions, including of course, physical contact. In Contacti this would be impossible. People who spoke that language would feel impelled to touch each other, or not to touch each other, in ways that are completely irrational or even socially harmful. For example, if someone were in pain, no one would want to touch him, including the doctors. On the other hand, everyone would be anxious to touch the rich and successful (which is already something of a problem). The whole idea is evidently insanely unworkable...If we tried to develop the Contacti fantasy in the emotive direction, then we imagine people whose lives are in countless ways bizarre and grotesque, from our ordinary point of view. I am not in fact even confident that we can make the fantasy fully intelligible in this direction...5

And Johnston concurs,

...if things were otherwise and the Contacti had the responses we would expect if we took their use of 'person' at face value, then the example of the Contacti either would be one of a tribe with our concept of personal identity and strange collateral beliefs about the effect of touching or would be difficult to develop coherently.6

Bizarre and grotesque, from our point of view, a community who genuinely organized their person-directed practices and concerns around the Contacti concept might well seem. But, for the pluralist, this is no good reason to rule out such a community as having an acceptable personal identity concept. Only if such a concept were not the sort of concept around which an informed community could successfully organize their person-directed practices and concerns would the pluralist have a principled reason to rule such a concept out of court. And, contrary to both Johnston and Hirsch, I think that the Contacti concept, developed in the emotive direction as it must be if we are to imagine it functioning as their concept of personal identity, is a potentially socially workable concept. It is not, however, given the way we are presently, albeit contingently, psychologically constituted, a possible concept for us. This is fortunate if, like Wiggins, you are worried that once we abandon constraints served up by nature, there is no principled way to render a pluralism suitably constrained—no principled way, that is, to rule out bizarre concepts as acceptable concepts for us. Perhaps unfortunate, in that it turns out that pluralism does not make for the radical possibility of organizing our person-directed practices and concerns in terms of excitingly weird and wonderful concepts.

I want to have two attempts at showing how the Contacti concept might be socially workable. The first is to show how it can be made socially workable, indeed scarcely at all bizarre, but at the cost of it seeming not to be the Contacti's concept of personal identity. This I take to be further grist to the pluralist's central claim that personal identity is whatever relation it is that organizes our person-directed practices; that a relation that does not play this role for a community is not their personal identity relation. The second is to make it socially workable as a concept of personal identity and, in the process,

---

to illustrate why (fortunately or unfortunately) it does not seem to be a psychologically possible concept for us.

One preliminary. On either of these two stories (or, indeed any story) the Contacti relation of exclusive contact could not be the only concept which a community employ in organizing their person-directed practices and concerns. At least two other concepts—a concept of bodily continuity or bodily identity and a concept of informational continuity (roughly, psychological continuity) or informational identity—must also do some organizing work. As, indeed, both of these concepts do for us. For countless of our everyday person-directed practices depend upon our having a concept or concepts which enable us to track people’s physical and/or psychological characteristics. We identify and reidentify people with specific interests in mind. When we go to see a particular surgeon, we want the person with the surgeon’s physical and psychological expertise. When we go to see a teacher, we want the person with the teacher’s expertise. When we want to play tennis, we want to play with the person who has the tennis-playing abilities. When we arrange for someone to help us move house, we want someone with the requisite physical strength. When we want to laugh over holiday photos with the friend who we went on holiday with, we want the person with our friend’s holiday memories. And so on and so forth. Successful organization of these manifold sorts of everyday person-directed practices depends on our employing concepts which enable us to track individual’s physical and/or psychological characteristics.

Indeed this failure to track personal identity with continuity of individual’s physical and/or mental characteristics is precisely what strikes us as so bizarre about Contacti concept in the first place. For, if touching did result in the transmission of a person’s characteristics, then the case of the Contacti* community would be just like the case of a community of Teletransporters (except, of course, touching rather than the teletransporter mechanism would be the cause of the transmission of a person’s characteristics). And nothing strikes us as particularly bizarre or grotesque about a community who have a concept of survival according to which a person survives teletransportation.
That the Contacti concept fails to track personal identity with individual’s physical and/or psychological characteristics is a *prima facie* reason for thinking that the Contacti concept could not be socially workable from the start. For, when two people are in exclusive contact, the Contacti concept does not track personal identity with continuity of these characteristics so central to, and seemingly necessary for, successful social interaction. Imagine, for example, that you, as a member of a Contacti community, are very sick and need to find a particular medical specialist. When you go to find the specialist you discover that she is touching someone. Now, remember that touching does not *actually* change anything in the brains and bodies of the people touching. Upon touching a person, though you become that person, and that person you, for the period of touching, touching does not result in the transmission of characteristics. The body that you touch does not inherit your memories, character traits, beliefs, desires, physical abilities, appearance, expertise, intentions, and so on. You want the person with the medical specialist’s skills, training, memories, beliefs, and expertise. And these do not swap bodies with touching. So in order to get the person you want, you will actually need to employ another concept—a concept of bodily identity—since, although you know the doctor has swapped bodies with the other person upon touching, you want the person with the doctor’s expertise, and (since bodily and informational continuity go together in this community) they will be in the body which was the doctor’s before touching. Similarly, if you want an engineer to build a bridge and go to ask the engineer only to discover that he’s touching someone who knows nothing about bridges, you will have to direct your request to the person who has the engineer’s beliefs, memories, and expertise, not to the person who is, by your Contacti concept, identical with the engineer. For you want the person with the engineer’s memories, beliefs, and expertise, and these do not swap bodies with touching. So too, if you want the strongest person in the Contacti community to manhandle a two tonne weight, and you go to find them, only to discover that the strongest person is touching a weakling, a concept of bodily identity will have to override your concept of touching as your concept of survival. For you want the person who has the strongest body, not the person who is, by your Contacti concept, identical with them. And so on.
For purely pragmatic reasons, in all of these cases and countless more like them, in order to get the person with the things that you want, you need also to employ a concept of bodily and/or informational identity; concepts which, in all of these sorts of cases, look like they would have to trump the Contacti concept as the concept in accordance with which people are identified and reidentified in circumstances where two Contacti people are touching. For in order to function successfully in countless everyday social situations it seems as if the Contacti would have to track personal identity with bodily and/or informational identity, not touching, when two people are in exclusive contact.

All of this might seem to show that it is no accident that personal identity theorists have focussed almost exclusively on relations of physical and/or psychological continuity and connectedness as candidate relations for personal identity. For such continuities are indeed necessary for personal identity if personal identity is to play the social role definitive of the concept. The pluralist’s claim would then merely be that these continuities are not sufficient for personal identity: that they are necessary for personal identity, but they are only necessary for personal identity because a relation that did not track these would not be a relation which could play the social role definitive of the concept.

The question posed by the Contacti is exactly the question of whether these two other concepts (in terms of one or other of which personal identity has typically been defined) are necessary for personal identity, or whether they might rather be independent resources on which we contingently draw in our application of the concept of personal identity, and personal identity can be something apart from them. Physical continuity theorists, for example, view psychological continuity as only contingently associated with personal identity; for physical continuity theorists, psychological continuity is just such an independent resource. Psychological continuity theorists, in contrast, view physical continuity as an independent resource, only contingently associated with personal identity. The question is, and contrary to initial appearances, might both physical and psychological continuity be thus only contingently associated with personal identity? This is, of course, what non-reductionists such as Cartesians claim (and, as we shall see in section 8.5, it gets them into some deep trouble). But survival might be
distinct from physical and/or psychological continuity without it being such a mysterious 'further fact'. It might just consist in some more particular relation among person-stages other than physical continuity or psychological continuity—it might, for example, consist in the relation of exclusive contact.

To the first story, then: the story of the Contacti*.

8.2 The Contacti*

One way of making the Contacti relation socially workable in the face of the sort of apparent counter-examples mentioned by Hirsch (in the preceding quotation) is to imagine a community who, in addition to employing the Contacti concept of 'personal identity', also employ some other concept which operates so as to offset or counteract the sorts of potentially socially problematic consequences of the Contacti concept to which Hirsch points. So we can imagine a community, the Contacti*, who, in addition to organizing their person-directed practices and concerns around the Contacti concept of touching, also organize their person-directed practices and concerns around an altruistic concept to the effect that the lives of others are of equal worth or value as their own. An altruistic concept, to put it in familiar Utilitarian terms, which yields a moral imperative of the form 'always act to maximise utility'.

So, for example, we can imagine that although the Contacti* are not at all reticent to touch a Contacti* person who is in excruciating pain, this is not because they do not organize their person-directed concerns and expectations around the Contacti concept of touching. On the contrary. It is because they also organize their person-directed concerns and expectations around an altruistic concept which impels them, even in the knowledge that upon touching the person in agony it will be them who feel excruciating pain, to touch those who are in pain in order to help, comfort, or cure them. For helping, comforting, or curing those in pain, even at the cost of incurring pain oneself (literally), will maximise utility. So it is not that the Contacti* do not expect to feel excruciating pain when they go to touch another Contacti* in excruciating pain in order to help them. They do. After all, upon touching the person in pain, it will be them who is in pain. It is simply that they also organize their person-directed concerns and expectations around an altruistic
concept which impels them, against their own better interests, to help those Contacti* in pain.

Similarly, we can imagine that the Contacti concept of personal identity, in tandem with their altruistic concept, governs their expectations about the relations between earlier action and desert of praise or blame. Except, of course, in the Contacti* community we are imagining, there would be no cause for blame. For in the Contacti community we are imagining, there would be no murder, grabbing of innocent hostages to avoid punishment for a crime or other such morally deviant behaviour. For in the Contacti* community we are imagining the Contacti concept of personal identity goes hand in hand with, and is locked in place by, a mutually reinforcing altruistic concept of valuing the lives of other Contacti* as much as one's own. If the Contacti* modified their altruistic concept, they would have also to modify their concept of personal identity. Now, if we define a community in terms of shared practices, and the practices that are definitive of our Contacti* community are guided by the Contacti concept of personal identity and an altruistic concept to the effect that the lives of other Contacti* are equally as important as one's own, then selfishly-motivated killing or taking of innocent hostages would never occur. A morally deviant person who takes a non-utility maximising course of murdering another Contacti* would have departed from the Contacti* community. Of course, there might be killing in the Contacti* community, if such killing were to increase utility. But such killing would not be blameworthy. On the contrary, since it is a means to increase utility, it would be praiseworthy. In fact, because all in the Contacti* community, by definition, are guided by this altruistic concept, and assuming their own death would increase utility, rather than waiting for some other Contacti* to kill them they would simply kill themselves.

As regards Hirsch's last imagined difficulty for the Contacti concept developed in the emotive direction, we can suppose that the Contacti* do not flock to touch the rich and successful, though they might wish themselves to be rich and successful, because touching the rich and successful would not make them rich and successful. (That is to say, I think that Hirsch's example here is misdescribed). For we are trying to imagine a Contacti community who organize their person-directed practices, including their legal and social entitlements, around their Contacti concept. And touching a person who is
rich and successful will not make you rich and successful because touching will merely make it true that you are now to be identified with the body of the person which used to be identified as the body of the person who was rich and successful; and things such as wealth and fame go with people, they do not stay with bodies. No problem here, then, about social intractability.

I could go on developing the story of the Contacti*, but this is far enough to make the necessary point. The question to be asked is whether the Contacti*, as I have presented them so far, successfully organize (or extrapolating out in the same vein, could successfully organize) enough of their person-directed practices around the relation of exclusive contact for it to count as their concept of personal identity? For, as we noted in Chapter 5, a relation does not have to satisfy all of the survival role in order to count as a relation of personal identity for a community, it just has to satisfy enough of it. Does the relation of exclusive contact play enough of the survival role in the Contacti* to count as the relation of survival? Do we, intuitively, think that the Contacti concept counts as their concept of personal identity?

I think it does not (and telling the story further—describing case by case how the altruistic concept operates to offset the otherwise socially problematic consequences of organizing around the relation of exclusive contact—would make little if any difference to this conclusion). Imagine that we are watching the Contacti* and wondering what their concept of personal identity is. We see that they are not at all reluctant to touch the sick; we see that they have no apparent practices of punishment; we see that they do not run around grabbing hold of the rich and successful or grabbing persons before they rob banks; and so on. In short, although they appear to have less crime than us and appear to value their own continued existence less than us, they don’t seem to behave in any way that would suggest that they have a remarkably different concept of personal identity from us. In fact, what differences there are seems better explained by their employing a slightly different moral concept: by their being impartial utility maximisers.

Indeed, we might not merely question whether or not the Contacti concept counts as the Contacti*’s concept of survival, we might question whether or not they have any concept of survival at all. For they are staunch universal utilitarians. Their person-directed practices are organized around
maximising utility; and doing what maximises utility may not require that they employ a survival concept at all. As we have already seen, the Contacti* have no need to employ a concept of survival in the context of blame and punishment, for since they all always act to maximise utility there is no crime or otherwise blameworthy activity to punish. And it may be that they similarly have no need to employ a concept of survival in other contexts. Their practices of love, attributions of ongoing rights and obligations, and so on, may all be governed by the maximising imperative without any need to employ a concept in accordance with which people are identified and reidentified as the same people over time. On the other hand, it might turn out that they do need to employ a concept of survival in aid of maximising utility. If it were to turn out that having an ongoing loving relationship with a single, same partner will maximize utility better than would engaging in a series of relationships with random many, for example. This may be true of the Contacti*, or it may not. But either way, if they have a concept of survival at all, it is not the Contacti concept. For that concept is entirely motivationally inert for them.

In retrospect, this should not be surprising. For the thought behind positing the altruistic concept in the first place was to offset what Hirsch took to be the socially unworkable upshots of the Contacti genuinely organizing their attitudes of self-concern and other such psychological motivations around the relation of exclusive contact. It was, in effect, to try to make them psychologically and practically as much like us as possible. For if we organize our person-directed practices successfully, and the Contacti are just like us, then we must admit that the Contacti too organize their person-directed practices successfully. But as much as adding the altruistic concept made the Contacti* community paradigmatically socially workable, it simultaneously makes it much less plausible to think that the Contacti concept is their concept of personal identity. For the more like us the Contacti, the less plausible it is to think that they are organizing around the relation of exclusive contact. For a community who genuinely organized around that relation would be very different from us. They would be reluctant to touch those in pain. And they would, from our point of view (although not, of course, from theirs), punish innocent people for crimes. But why should we suppose that the way in which we organize our person-directed practices, is
the only way it is possible for a community successfully or coherently to organize them?

So let's forget about offsetting the Contacti concept of personal identity to make it paradigmatically socially workable i.e. to give them person-directed practices just like our own. For this looks like it was gerrymandered in favour of our concept(s) right from the start. Let's start afresh and imagine what a community, call them the Contacti**, who really organized their person-directed practices and concerns around the Contacti concept would look like. Could such a community coherently organize enough of their person-directed practices and concerns around the Contacti relation of exclusive contact?

8.3 The Contacti**

Amongst the platitudes about personal identity in the cluster that I think we weight very heavily are the platitudes about self-interest, memory and responsibility. Indeed, Johnston defines personal identity solely in terms of these practices, taking them alone to be jointly necessary and sufficient for personal identity. If Johnston is right about the heavy weight we assign to these platitudes (and I think he is), then were the Contacti** able successfully to organize these of their person-directed practices around the Contacti relation of exclusive contact we might well be prepared to say that the Contacti**, although undeniably bizarre, nonetheless employ the Contacti concept as their concept of personal identity. Let me describe how the Contacti** might successfully organize their practices of punishment and self-concern around the relation of exclusive contact by way of just one short story.

Imagine that a Contacti ** bank robber wants to rob a bank. Being self-interested, he would prefer that he himself not be blamed and punished for it. So he decides, rather than risk himself being later caught and punished for the crime, the thing to do is to maintain exclusive contact with an innocent bystander throughout the course of the robbery so that, after contact is released, the innocent bystander, not him, will be punished for the crime. Being a well-enculturated Contacti**, he believes that there is nothing unjust about this (not that he would be particularly worried even if there were!), for the person who is the same person as the person who committed the crime is
the person who deserves to be blamed and punished for it. And the innocent bystander, not him, will be the person who, if everything goes according to plan, commits the crime. Likewise, the innocent bystander will regard it as perfectly just that she gets blamed and punished for the crime. For, upon being grabbed by the robber for the duration of the robbery, her life history will include the stage of the robber for the period in which they were in exclusive contact. She (or a stage of her) robbed the bank, and so she deserves to be blamed and punished for it.

All does go to plan. The robber goes into the bank, gun in hand, grabs an innocent bystander and demands the money from the teller. Having mastery of the Contacti concept of personal identity, he then starts stuffing the money into the pockets of the body who used to be the innocent bystander, but who, now, of course, is him. After all, being self-interested, he doesn’t want to give the money to the innocent bystander (as he would be doing were he to stuff the money into the pockets of who we would call the robber)! Of course, later, once contact is released, the robber will once again be associated with the robber-body. Stuffing the money into the pockets of the robber-body would be in the robber’s long-term interests i.e. in his or her interests after contact is released. But, now, and for the duration of contact, the innocent bystander is in the robber-body. And to now put the money in the pockets of the robber-body would be to give the money to the innocent bystander. Not something that the (short-term) self-interested robber wants to do. Similarly, the innocent bystander, we might imagine, disapproving of robbery, desperately tries to take the money out of the robber’s pockets again i.e. out of the pockets of who we would call the innocent bystander. And perhaps, wanting to be able to return the money to its rightful owner, desperately starts trying to stuff the money into his own pockets i.e. into the pockets of the person who we would call the robber.

The point to glean is that in the Contacti** community, acting self-interestedly requires acting in the interests of the body touched for the period of exclusive contact. For, for the period of exclusive contact, those are your interests. Such a pattern of self-concern is certainly very strange from our point of view. Not least because, to return to the robbery, when a Contacti** robber acts self-interestedly (e.g. stuffs the money into the pockets of the person who is now them but who, post-contact, will be the innocent
bystander) they bring it about that, once they release contact and run out of the bank, the innocent bystander ends up with the money. Psychologically bizarre Contacti** robbers undoubtably are, but I don’t see any reason to suppose that acting in this way is actually incoherent. (I will return to this point again shortly.)

What of the sorts of practices—practices which require our tracking individual’s informational and bodily characteristics—that, *prima facie*, looked particularly problematic for the Contacti concept? Suppose that a Contacti** person, Susan, comes to see a particular medical specialist, Fran. (Let’s say that the person who has the medical specialist’s expertise, memories etc is wearing a white coat.) Upon entering the doctor’s room, Susan discovers that Fran is touching another body, which, prior to contact, was Caroline’s body. Unperturbed, and in accordance with her Contacti concept, Susan turns to the person in the white coat and says ‘Caroline, you must temporarily have the doctor’s memories and expertise, could you give me a check-up?’ Of course, as we have already noted, in order to identify the person with the doctor’s knowledge, Susan will need also to employ a concept of bodily identity (or, in a case where bodily and informational identity come apart, a concept of informational identity as well). For as we noted, contact does not actually result in the transfer of characteristics from body to body. However, this does not mean that the Contacti cannot employ their Contacti concept, which does not track these things, as their concept of personal identity in these cases. As already discussed, like Susan, they believe that Caroline, not Fran, is the person who temporarily has the doctor’s expertise.

Of course, where patients are presently experiencing great pain, Contacti** doctors may well be reluctant to come into contact with the patient. For, for the period of touching, they will be the person in great pain. But perhaps Contacti** doctors get androids to deliver large doses of morphine to the patient before contact. Or perhaps they feel that the exorbitant fees they charge for their services are sufficient to compensate for short-term pain. (In that, they may be psychologically not so different from some doctors round here!) In any case, where patients are not actually in pain prior to contact, Contacti** doctors would most likely have no qualms whatsoever about touching sick patients—would it be so bad to have cancer for ten minutes?
What of memory? Return to the above case of Caroline and the doctor, Fran. Let’s assume, for simplicity’s sake, that Caroline is the first person with whom Fran has come into contact. Before contact with Caroline, Fran views the memories she has as her own: as memories as of things she herself has experienced. When contact occurs, just as other Contacti* (such as Susan) view the person with the doctor’s memories, expertise etc as Caroline, so Fran thinks to herself ‘I am now having Caroline’s memories’. In retrospect (i.e. after contact is released), Fran will think to herself, ‘I am now having my own memories again, but the memories I have from the period in which I was in contact with Caroline are of things that Caroline experienced.’ For my life history includes a stage of Caroline for the period in which we were in exclusive contact.

If the concept of memory that the Contacti* employ seems familiar, it is because it is the psychological continuity theorist’s, q-memory (except, of course, unlike Parfit’s example of Jane and Paul in Venice, Contacti* people can have q-memories without there being underlying neurophysiological change). On the Contacti concept of personal identity, a person may remember things that she herself has not experienced (as, for example, when doctor Fran expects to remember Caroline’s experiences for the period in which they were in exclusive contact). On the assumption that there is nothing incoherent about the concept of q-memory (and psychological continuity theorists had better hope that there’s not!), there is nothing incoherent about the Contacti* organizing their expectations about the relationship between memories and remembered-experiences in terms of their Contacti concept.

Having seen how the Contacti* seem to manage successfully to organize the aforementioned, heavily weighted sorts person-directed practices around the Contacti concept of exclusive contact, we can see how the story will go for the rest. The question is, is the Contacti concept their concept of personal identity: that is to say, do they organize enough of their person-directed practices around the Contacti concept in order for it to count as their concept of personal identity? And the answer seems to me, intuitively, ‘yes’. When we look at the Contacti* we see a community who systematically act in the interests of the body touched for the period of exclusive contact; who punish the person grabbed by a robber during a
robbery and who, moreover, regard it as perfectly just that the person grabbed should be so blamed and punished; who expect to remember the experiences had by the person touched during a period of exclusive contact; who identify and reidentify people on the basis of exclusive contact; and so on. True, they employ concepts of bodily and/or informational identity in aid of the application of their concept of personal identity. But relations of bodily and/or informational identity are not the relations around which they organize their person-directed practices and concerns for periods in which they are in exclusive contact. It is the relation of exclusive contact, not the relations of bodily and/or informational identity that are psychologically motivational for the Contacti**. Whereas the Contacti* were psychologically motivated by the altruistic concept, not the Contacti concept of touching, the Contacti** are genuinely motivated by, and act in accordance with, the Contacti relation of touching. They act in the interests of the body they touch for the period of exclusive contact. If they have to choose who gets tortured during a period of contact they will choose that their old body gets the torture (in our terms, that their body gets the torture). They do not regard it as in any way unjust that they get punished for crimes which were planned by somebody then and now distinct from them; crimes, moreover, from which they didn’t even get any loot. And so on. Unlike Johnston, who defines survival summary-style in terms of these three sorts of person-directed, I define the survival more widely in terms of the whole cluster of platitudes about survival. To describe bit by bit how the Contacti could similarly successfully organize all of these in terms of the relation of exclusive contact would be to write an entire novel about this strange community, the Contacti**. But if Johnston is right in thinking that the person-directed practices in terms of which he defines personal identity are jointly necessary and sufficient for personal identity, or at least very heavily weighted in the cluster (and I think that he is), then it seems that the Contacti concept is the Contacti**’s concept of personal identity. And this seems the intuitively right thing to say. The Contacti concept of exclusive contact is then a possible concept of survival, for it is the sort of relation around which an informed community could successfully organize enough of their person-directed practices and concerns.
However, fortunately or unfortunately, the Contacti concept of exclusive touching does not seem to be a possible concept for beings with the sort of psychology that we have. For example, acting self-interestedly in the Contacti** community requires, first, predicting what would be in the interests of the body touched and, second, being disposed to act in accordance with those interests for the period of contact, even where those interests are at odds with what were your interests prior to contact. For example, if an honest policeman is grabbed by a bank-robber during a hold-up, the policeman must be disposed to change his desires to act in the evil interests of the ex robber-body and takeover the hold-up. Likewise, in such a situation, the robber must be disposed to change his preferences to act in the interests of the body which used to be associated with the honest policeman prior to contact. The robber must desperately try to stop the robbery; the honest policeman must desperately try to carry it out.

As psychological characteristics are not actually transferred upon contact, when you are in contact with another, figuring out what is in your self-interest requires an enormous effort of empathy. It requires, literally, that you put yourself in the shoes of the body touched for the period of exclusive contact. For in order to act in your own interests for the period of contact, you need to work out what that body's interests are. And since you do not have privileged psychological access to those interests (because you are not actually psychologically continuous with the person you touch) the only way you can work out what is in your interests is by predicting what the interests of the body you are touching are—what are the intentions, aspirations, projects, values, goals etc. of the body you are touching? What would that body want you to do? And that, for the Contacti**, is exactly equivalent to asking what is it in *my* interests to do? For the Contacti**, self-interest, like memory, is not strictly self-interest, but q-interest. However, not only do you have to predict what the interests of the body touched are in order to work out what is in your self-interest, you must also be disposed to change what were your interests prior to contact in such a way as to do what is in the interests of the body touched, even in circumstances (such as the robber touching the policeman) where those interests are diametrically opposed to what were your interests prior to contact.
If incoherence threatens anywhere, for the Contacti**, it is here. But I do not think that q-interest of this sort is actually incoherent. It is rather that, for beings with our psychology, the Contacti concept of exclusive contact would coexist in an extremely unstable alliance with concepts of bodily and informational continuity. Given our sort of psychology, I think, it would not be long before we would begin to wear metal cages or engage in other contact-preventing measures to avoid being grabbed by those too psychologically dissimilar from us or of whose interests and values we disapprove; or taking bullets for people who, in retrospect, are complete strangers (if we were still around to retrospect, that is); or handing over our life savings to pay for what, in retrospect, is someone else’s debt; or regarding it as perfectly just that we should be punished for a crime committed by a person who, in retrospect, is someone else. It would not be long, I think, before beings with our psychology would begin to act in our long-term interests i.e. in the interests of the person who is physically and psychologically continuous with us post-contact, rather than in the interests of the body touched for the period of contact. For, for example, when people with our psychology rob a bank they want to end up with the money. We would take great precautions to ensure that we come into contact with only those friends and family in whose interests we would want to act anyway. We would cease to bother addressing the doctor as ‘Caroline’; and, when we remember some experience, worrying about whether we could truly say that it was of something that we ourselves had experienced, or whether it was rather of something that someone else had experienced while we were touching them. And so on. In a relatively short time, I think, concepts of bodily and informational continuity would win out to replace the Contacti concept as the concept around which we organize our person-directed practices and concerns. The relation of exclusive contact just does not seem to be a relation around which beings with our psychological dispositions, interests and motivations could stably care for very long.

** 8.4 Two possible lessons**

What conclusions should we draw from the case of the Contacti (and the Contacti* and the Contacti**!)? I think there are two possible lessons, either of which I find perfectly congenial (although, since I think that the story of the Contacti** describes a coherent community wherein the relation of
exclusive contact functions successfully as their concept of personal identity, I think that the second is most likely the right one).

The first lesson. If you think, despite my story-telling efforts, that the story of Contacti** does not (or could not), in the end, describe a coherent community, then there is a good a priori reason for thinking that only relations which underpin the transmission of individuals’ physical and/or psychological characteristics are candidate relations for survival. For, as noted at the outset, where the Contacti relation of exclusive contact will break down, if it breaks down at all, will be in its failure to track survival with bodily and/or informational continuity. What we will then think is that it is a priori true that survival is a disjunction of the relations of physical or psychological continuity; and pluralism will simply tell us which of these two disjuncts (or some slight variation thereon) is the survival relation for a given community. On this view, pluralism will simply tell us which way a community precisifies the concept of survival in indeterminate cases; either precisification of which is a perfectly acceptable survival concept. (I think that this is the way that Mark Johnston thinks of pluralism). This would be a less exciting and radical view, but a pluralist view, nonetheless.

If, on the other hand, we think that the Contacti** does describe a coherent community wherein the Contacti concept functions successfully as their concept of personal identity (as I tend to), then we will think that pluralism has more scope than this. It does not merely decide which of physical or psychological continuity—the only acceptable candidate survival relations—constitute survival for a community. It is not a priori true that survival is either physical and/or psychological continuity. Survival for any community is rather whatever relation among person-stages it is that that community organizes their person-directed practices and concerns around. And pluralism allows for much greater object-level scope than the comparatively narrow range of candidates which have been canvassed in the personal identity to date. Whether we ourselves are in a position to avail ourselves of some of the weirder and more wonderful possible concepts depends only on contingent (although probably well-engrained) facts about our psychological constitution.
8.5 Is the Cartesian View possible?

Our discussion of the Contacti concept serves more than merely to illustrate the possible object-level scope of pluralism. It also serves to make vivid a worry, mentioned only very briefly in passing in chapter 1, about the intelligibility of Cartesian views of personal identity, or indeed, of any view that similarly leaves relations of physical and/or psychological continuity and connectedness only contingently associated with survival.

Recall that the Cartesian view is a non-reductionist view: survival consists in some further fact (for Cartesians, the continued existence of the same immaterial entity or soul), over and above, and not reducible to, facts about physical and/or psychological connections and continuities among person stages. On the Cartesian view, facts about physical and psychological continuity are mere fallible evidence for personal identity (i.e. the continued existence of the same soul), they are in no way constitutive of, or necessary for, personal identity. As Richard Swinburne writes,

...although apparent memory and brain continuity are, as they obviously are, evidence of personal identity, they are fallible evidence and personal identity is something distinct from them. Just as the presence of blood stains and fingerprints matching those of a given man are evidence of his earlier presence at the scene of the crime, and the discovery of Roman-looking coins and buildings is evidence that the Romans lived in some region, so the similarity of P2’s apparent memory to that of P1 and his having much of the same brain matter, is evidence that P2 is the same person as P1. Yet blood stains and fingerprints are one thing and a man’s earlier presence at the scene of the crime another. His presence at the scene of the crime is not analysable in terms of the later presence of bloodstains and fingerprints. The latter is evidence of the former, because you seldom get bloodstains and fingerprints at a place matching those of a given man, unless he has been there leaving them around. But it might happen. So, the suggestion is, personal identity is distinct from, although evidenced by, similarity of memory and continuity of brain.7 (emphasis added).

This feature of the Cartesian view—that it allows that the soul may exist, continue to exist, and even cease to exist quite independently of the physical and psychological continuities that are our evidence for it—has lead some to

---

maintain that the view is unintelligible. As Kant explains the worry, the Cartesian view is perfectly compatible with the possibility of there being a whole series of substances of which the first transmits its state together with its consciousness to the second, the second its own state with that of the preceding substance to the third, and this in turn the state of all the preceding substances together with its own consciousness and with their consciousness to another. The last substance would then be conscious of all the states of the previously changed substances, as being its own states, because they would have been transferred to it together with the consciousness of them.8

Assuming Kant’s sceptical hypothesis is coherent, what it shows is that on the Cartesian view personal identity is unknowable. For nothing in our first-person experience or our third-person practices of identifying and reidentifying people on the basis of observed physical and/or psychological continuities can rule out the possibility that what we naturally and habitually take to be cases of the persistence of the same person over time might in fact be a succession of different people or souls, each communicating its consciousness to the next. For all that we know through introspection and observation, souls might constantly be flitting round from body to body, psychology to psychology, even as individual’s mental and physical characteristics remain entirely unchanged. In making personal identity thus unknowable, it is widely argued, the Cartesian view violates a central condition on an account of personal identity: that it be able to account for the kind of knowledge that we have of our personal identity, including the special kind of knowledge that we seem to have of our own continued existence over time.

However, the possibility described by Kant points to a related but broader and potentially far more serious difficulty for the Cartesian view; a difficulty, as we saw in chapter 1, of which Locke was acutely aware. The difficulty is that, in a world where souls systematically fail to underpin relations of physical and/or psychological continuity and connectedness, sameness of soul, just like the Contacti relation of exclusive contact, may not be an acceptable candidate relation for survival. For, in a world where souls are so badly behaved, sameness of soul may not be the sort of relation around

which a community could successfully organize their person-directed practices and concerns. This worry, though more pointed on the account of survival I have argued for, is nonetheless a very serious worry on any account of personal identity. For, as we saw in chapter 2, in addition to the requirement of knowability, a second central and standard condition on an acceptable account of personal identity is that the account be able to make intelligible the special sort of importance that personal identity has for us.

Note that Kant's example depends on the assumption that there are souls in the actual world, and that they are well-behaved. For otherwise why should we even begin to entertain the worry that badly behaved souls might not count as souls i.e. candidate relations for survival? What is it about an immaterial entity that makes it count as a soul or candidate relation for survival? It is the supposition that that immaterial entity plays the survival role in the actual world. Only then does the worry about whether or not an immaterial entity which did not behave as it is supposed to behave in the actual world would still count as a soul (i.e. a candidate relation for survival) get off the ground.

Objectors to the Cartesian view have divided roughly into two camps. The first maintains that, insofar as the Cartesian view allows for the possibility that souls might be badly-behaved, the view is unintelligible. Others maintain that the Cartesian view is not unintelligible (personal identity might have consisted in sameness of soul), it is just as a matter of fact false. For there are no immaterial entities in this world. The pluralist agrees that in a world, such as ours, where there are no immaterial egos, survival will not consist in the continued existence of an immaterial entity. For that will not be what we are organizing around. But what will the pluralist say about worlds where there are such things? In particular, is the badly behaved immaterial entity an acceptable candidate relation for survival?

---

9I owe this point to David Braddon-Mitchell.
10This is, of course, the pluralist answer. The practice-independent theorist would give a different answer. For example, survival consists in these immaterial entities because God said that survival consists in these immaterial entities. I hope that by now these answers will seem unsatisfactory.
What the pluralist will say here depends on which of the two lesson is the lesson to be drawn from the case of the Contacti. Those who believe that the first lesson is the right one will have it that only relations of physical and/or psychological continuity are relations around which a community could successfully organize their person-directed practices and concerns. So that any relation other than these (such as the relation of exclusive contact, or sameness of soul in a world where souls do not underpin transmission of individual’s physical and/or psychological characteristics) is not an acceptable candidate relation for survival. If the first lesson is the one to be drawn from the case of the Contacti, then the (badly behaved) immaterial entity is ruled out \textit{a priori} as an acceptable candidate relation for personal identity.

If the second lesson is the right one to be drawn, however, we will take a different view. For if the Contacti relation which fails to track personal identity with physical and/or psychological continuity is nonetheless the sort of relation around which a community could successfully organize their person-directed practices and concerns, then an immaterial entity, even when badly behaved, is similarly an acceptable candidate relation for survival. But what pluralism of this second genre gives to the Cartesian view with one hand, it takes away with the other. For just as the Contacti relation seems an impossible survival concept for beings psychologically constituted like us, so (in a world where immaterial entities similarly fail to track physical and/or psychological continuities among person-stages) sameness of immaterial thing is not a psychologically possible survival concept for us. Either way, then, badly behaved immaterial entities seem not to be candidate survival relations for us.

What, however, of a world where immaterial entities do not come apart from the physical and/or psychological continuities among person-stages that are evidence for them? Is a world in which there are immaterial entities, and where subjects organize their person-directed practices around those immaterial entities, a world in which survival consists in the continued existence of those immaterial entities? The worry here is that such a community will be like the Contacti as Hirsch originally described them, or like my description of the Contacti*. Both communities talk as if the Contacti concept was their survival concept, but in fact it is completely redundant as
regards their person-directed practices. They say they organize around the Contact relation of exclusive contact, but in fact they organize around something else (in the former case, relations of physical and/or psychological continuity, in the latter, an altruistic concept). An exactly analogous worry arises for the case of a community who organize around a well-behaved immaterial entity: they say they organize around an immaterial entity whose continued existence corresponds exactly to the relations of physical and/or psychological continuity among person-stages which are evidence for it, but in fact they just organize around the relations of physical and/or psychological continuity. For, for all intents and purposes, their survival-oriented behaviour will be indistinguishable from that of a community who organizes straightforwardly around relations of physical and/or psychological continuity. What reason might we have, then, for thinking that the soul is doing any organizing work? That is to say, what reason could we have for thinking that sameness of immaterial entity, rather than relations of physical and/or psychological continuity, is their survival relation?

Here is a reason. If they are genuinely organizing around an immaterial entity, although their actual survival-oriented practices will be practically indistinguishable from a community who organize around relations of physical and/or psychological continuity, their counter-factual dispositions will be different. Were the immaterial entity to come apart from relations of physical and/or psychological continuity, they would track (or be disposed to track) survival with that immaterial entity rather than with physical and/or psychological continuity. For a community to genuinely organize around an immaterial entity will be for them to have these distinctive counter-factual dispositions. For it to be true that Swinburne, for example, genuinely organizes his person-directed practices around an immaterial entity, rather than simply around the relations which are evidence for it, it must be the case that (unlike Parfit, Shoemaker, Williams, Unger and the like) he is disposed to track survival with an immaterial entity, rather than with the relations which are typically evidence for it, in counter-factual circumstances where they come apart. That is how pluralism admits the well-behaved immaterial entity as a candidate survival relation for us.
PLURALISM IN THE ACTUAL WORLD

Pluralism has important and timely implications for the ongoing personal identity debate in this world; a debate which, having been dominated to date by competing practice-independent accounts, has taken the form of an increasingly desperate search for the single, independent answer to the question of survival.

As we have seen, contemporary conceptual analysis in the case of survival, as elsewhere, has gone by way of what we might call the method of cases. Theorists formulate putative accounts of the necessary and sufficient conditions for survival over time; accounts which are evaluated in view of how well they accord with our intuitive responses (taken to reflect our prior grasp of the concept of survival) gleaned from actual and hypothetical puzzle cases. The right account of survival, it is assumed, will be the one which best predicts and explains our intuitive reactions to these puzzle cases.

Notoriously, however, the intuitive deliverances of the method of cases have proved incredibly problematic for practice-independent accounts, which assume that there is a single, independent answer to the question of personal identity. The problem has been that our intuitive reactions to puzzle cases do not seem to settle on any particular one of the numerous, competing candidate properties that the different practice-independent accounts claim to take the honour of instantiating the single, practice-independent truth of the matter. And given the assumption that the right account of survival is the one which can predict and explain our intuitive reactions to puzzle cases, the fact that our intuitive responses appear not to converge on any one of these competing accounts, would seem to suggest, given practice-independence, that all these accounts are far from adequate by their own lights.

Somewhat surprisingly though, practice-independent theorists have responded to this problem not by reassessing their assumption that the question of survival is itself the kind of question which admits of a single, independent answer in the first place, but by devising more and more
ingenious puzzle cases designed to prompt intuitive reactions which will accord with their preferred account, and by engaging in more and increasingly elaborate attempts to explain away those recalcitrant intuitions which would seem to favour some other competing account. The feeling seems to be that if we could just find quite the right way of presenting these puzzle cases then people’s intuitive responses will converge and we will then have hit upon the single, independent answer to the question of survival.

Of course, even if we did find a way of presenting these puzzle cases so our intuitive reactions in fact converged and/or a principled explanation as to why they had not so converged in the past we would not necessarily have found the independent answer to the question of personal identity or survival. What we would then have found is that in this world, as a matter of fact, survival is realized by a single relation (it might be a single disjunctive relation). We would not have found that personal identity or survival is necessarily instantiated by this particular relation. That is to say, the fact that different subjects in this world converge in their opinions about what relation survival in fact consists in would not necessarily show that survival for all communities in all worlds is whatever relation it is that is survival for us in the actual world. This is arguably true, for example, in the case of colour. In the case of colour, as a matter of fact, normal observers in normal conditions converge in judging that fire engines are red, grass is green and the sky is blue. But many doubt that the mere fact that there is this convergence shows that ‘red’ rigidly names the (probably very disjunctive) property that it names in the actual world. For colour seems intimately connected to subject’s colour experience in a way in which squareness, for example, is clearly not to subject’s experience of square. So we want to leave it open that things might still be red for normal observers in normal conditions, even if it were the case that a different property caused things to look red. And this is not simply because ‘red’ happens to pick out a particularly messy disjunction in the actual world, it is because we think that what fixes the reference of ‘red’ in any world depends on what property (or properties) cause things to look to red to observers in that world; and different properties may play that role for different subjects in different worlds. All this is simply to say that pluralism might still be true—survival might still be multiply realizable—even if survival were not actually (as I think it in fact is) multiply realized. Pluralism
might still be the right analysis of survival, even if it were the case that everyone in the actual world happened to organize their person-directed practices around the same relation. That would just tell us that we all agree about what survival is in the actual world. It will not, in itself, tell us what survival is in all worlds. To answer that, we need to investigate the meta-issue.

But the problem in the case of personal identity, unlike the case of colour, has been that, after more than a century of attempting to get the presentation of puzzle cases quite right, peoples' intuitive responses in the actual world seem no closer to converging around the same relation. The problem is not so much that individual subjects have internally inconsistent intuitive reactions to puzzle cases (although that may sometimes be the case), the problem is that different subjects have different opinions about what relation survival consists in. The problem, that is to say, is not the fact of intra-personal divergence, it is the fact that there is substantial inter-personal divergence. And, what has been particularly troubling for practice-independent theorists, is that different subjects seem no closer to converging even when those subjects have fully informed, considered and consistent beliefs about the matter. For, philosophers like Parfit, Perry, Shoemaker, and Lewis, amongst others, on the one hand, and Williams, Wiggins, Nagel, and Unger, amongst others, on the other, have an informed, considered, consistent and precise grasp of their person-directed beliefs and concerns and the property around which they are organized. They are in conditions of increasing information and critical reflection if anyone is. But their informed, considered and consistent opinions as to the property or properties which instantiate personal identity or survival do not converge. The former insist that the latter are mistaken, for physical continuity is no necessary condition for personal identity, whilst the latter insist that the former are mistaken, for physical continuity is necessary for personal identity. And they continue to diverge even when they know everything about the connections between person-stages which they themselves claim to be the only things they need to know in order to decide whether or not a person survives. Yet the search for the single, independent truth of the matter has raged on by way of more and more ingenious puzzle cases designed to show once and for all how and why
members of opposing camps are mistaken, and more and more elaborate attempts to explain why it is that they had been led astray in the first place.

But if the concept of personal identity is essentially connected to concepts of subjects' person-directed practices and concerns in the specified conditions then we have an alternative explanation as to why these two groups of philosophers so vehemently disagree about the properties which instantiate personal identity. It is not, I will argue, that one or other of the two groups are fundamentally mistaken about survival, or mean something significantly different by survival, it is just that different relations realize the survival role for them. They, like the Teletransporters and the Somataphiles, just care, and blamelessly so, about different properties in survival; properties which, in virtue of the common role they play in their person-directed concerns, they take to constitute the concept of survival. Rather than viewing one or other of their disparate intuitive responses to puzzle cases as mistaken reactions to be explained away, we should see these responses (informed and consistent as they are) as revealing what property is it that, as a matter of fact, constitutes survival for them; as revealing what they care about in survival. That is to say, we should hold pluralism not merely to be possibly true, but actually true: true in the actual world.

In so saying, I am rejecting some of the more standard sorts of explanations for this lack of object-level convergence. So I'd better say a little more about why. Perhaps the best place to begin is with a famous set of puzzle cases first presented by Williams¹.

Case 1
Suppose that you and another person, Brown, are reasonably physically similar. Now, imagine that your brain is physically altered so that it comes to contain the information and have the dispositions of Brown's. Brown's brain is similarly altered with the result that it contains the information and dispositions yours has now. So the person who wakes up in your (present) body after the operation will seem to remember doing all the things that Brown now remembers doing and will seem to have all of Brown's beliefs, desires, intentions, character traits, and so on. They will think that they are

¹Williams (1970).
Brown, and so will everybody else. Similarly, the person who wakes up in Brown's body will seem to have your (present) psychology: they will think that they are you, they will seem to remember all the things that you now remember doing, have all your character traits, beliefs, desires, intentions, and so on. Which, then, of the two resulting persons is you?

To make the question more vivid, suppose that you are told the following news: one or other of the people (the person in your body or the person in Brown's) will, after the alteration, be tortured for twenty-four hours. There is no way that you can prevent this. But you are offered the following choice: you can choose which body gets the future torture. Which body would you choose?

The widespread intuitive response to this case is (or is supposed to be) that the person who inherits your distinctive dispositional psychology—memories, intentions, aspirations, and other character traits—is you; so that, in choosing which person gets the future torture, you choose the person in your body. In this case, the possession of a particular body seems less crucial for personal identity than the possession of a distinctive sort of psychological make-up. The dominant reaction to this case, then, is to track personal identity over time with psychological connectedness and continuity where psychological and physical continuity so come apart. What philosophers take this case to show, then, is that psychological, but not physical, continuity is necessary for personal identity. This case, then, seems to support a psychological continuity account of personal identity. For that is the commitment which explains our intuitive response. But now consider the following case.

Case 2
Suppose that you are given the following news. You are to be tortured in twenty-four hours. There is no way you can avoid the impending torture, but you are offered the following consolations. Firstly, before the torture is administered, you will suffer complete amnesia regarding your present experiences. Secondly, you will be given a set of delusive memories about your past. Thirdly, you will be given a set of behavioural dispositions which are completely at odds with your present character and constitute a basic change in your present personality. The question to ask yourself is would
you take the consolations. Would they make you any less deeply anxious about the impending torture?

And the dominant response here is (or is supposed to be) a definitive ‘no’. As Bernard Williams reiterates, none of these changes or so-called ‘consolations’ seem to relieve any of the special apprehension and anxiety you feel toward the impending torture. These changes would just be adding insult to injury. As if it weren’t bad enough that you were going to be tortured, you would have your psychology fiddled with as well. In this case then, we seem to track personal identity with bodily, not psychological, continuity where the two come apart. What this case seems to show, then, is that physical, not psychological, continuity is necessary for survival. For that survival requires physical continuity, but not psychological continuity, seems to explain our intuitive response to this case.

But what has incited so much interest in these two Williams cases is not merely that we have disparate intuitive responses to the cases (in one case we seem to track survival with psychological continuity, whilst in the other we seem to track it with physical continuity, where the two come apart). What is particularly interesting about the Williams cases is that we appear to have blatantly inconsistent or contradictory intuitive reactions to the very same case. For Case 2 is just a redescription of Case 1. The only difference between the two being that in Case 1 we are told both sides of the story—we are told about what happens to both you and Brown—whilst in Case 2 we are told only about what happens to you. So Case 2 and Case 1 are just different presentations of the very same case. And, taken together, our intuitive reactions to these two presentations of what amounts to the very same case seem to show that neither physical nor psychological continuity is a necessary condition for personal identity or survival, though each may be sufficient. But this seems a very odd conclusion.

Not surprisingly, much has been made of our apparently inconsistent intuitive responses to these two cases. Williams himself takes what he assumes is our tendency to revise our response to the first case in light of the second to auger strongly in favour of a physical continuity account of personal identity. Whilst psychological continuity theorists have engaged in elaborate attempts to explain away our intuitive reaction to the second case
either as based on misleading features in the presentation of the cases, or by amending or supplementing existing psychological accounts so as to be able to explain them (or both). And a few philosophers have even taken our inconsistent intuitive reactions to these, and related, cases as dealing a telling blow to theorizing about survival by way of the method of cases altogether.

Robert Nozick advocates one of the best known and widely endorsed attempts to explain away our intuitive reaction to the second case—the case which appears to support a physical continuity account. According to Nozick, our intuitive reactions to these two puzzle cases can be explained by recourse to what he calls the 'closest-continuer theory'. Briefly and roughly, according to the closest continuer theory, we track the identity of a person over time with the closest (and close enough) continuer of that person. The closest continuer theory is itself neutral between physical and psychological continuity theories, for in itself it says nothing about what relations or weighted set of relations determine closeness. But, according to psychological continuity theorists, the closest continuer theory, in tandem with a psychological continuity construal of closeness, can explain away our intuitive reactions to the second presentation of the case. For the closest continuer theory enables psychological continuity theorists to point to what they claim is an important difference between the first and second presentations of the puzzle case. The crucial difference is that in Case 1 we are presented with a better continuer (a psychological continuer) of the original person than the person (a mere physical continuer) with whom we are presented in Case 2. This is the significance of leaving out what happens to Brown in Case 2. What our intuitive reactions show here, then, is that although a physical continuer is close enough or sufficient for personal identity (Case 2), a psychological continuer is better (Case 1). And so, despite initial appearances, our intuitive reactions to these cases support a psychological continuity account of survival. Unsurprisingly, many psychological continuity theorists have been quick to follow Nozick’s line of response.

Unfortunately, however, according to many physical continuity theorists\(^2\), the closest continuer theory cannot save the psychological

\(^2\)And even a few psychological continuity theorists!
continuity account as the right account of personal identity (i.e. the account upon which our intuitive responses converge). For, they claim, there are many other puzzle cases where the closest continuer theory not only appears unable to explain our intuitions, but actually yields extremely counterintuitive conclusions. Here is one such case they cite in support. Imagine again that your brain is physically altered so that it comes to possess the memories, intentions, character traits and other psychological dispositions of Brown’s. You wake up, thinking you are and have always been Brown. But, next to you, there is a state-of-the-art machine in the process of physically altering the brain of another person, White, so that White too will come to possess the psychology of Brown and, since the machine is state-of-the-art, will make White a slightly better psychological continuer of Brown than you. On the closest continuer view, whether or not you are Brown all depends upon whether or not you turn the machine off. If you do, you will be Brown (for you will be the best continuer of Brown), if you don’t, White will be Brown (for White will be a better continuer of Brown than you). Further against the closest continuer theory as an explanation of people’s intuitive reactions to Cases 1 and 2 we might note that, even where what happens to the other person (Brown) is included in the description of the case, many people’s intuitive reactions still seem to favour a physical continuity account in Case 2. With these details included the case reads:

**Case 2**

Suppose that you are given the following news. You are to be tortured in twenty-four hours. There is no way you can avoid the impending torture, but you are offered the following consolations. Firstly, before the torture is administered, you will suffer complete amnesia regarding your present experiences. Secondly, you will be given a set of delusive memories about your past. Thirdly, you will be given a set of behavioural dispositions which are completely at odds with your present character and constitute a basic change in your present personality. As a final consolation, the same kinds of psychological changes will be effected in another subject, who, after suffering amnesia, will seem to remember the things you remember now and will exhibit your personality traits. The question to ask yourself is would you take the consolations. Would they make you any less deeply anxious about the impending torture?
And here again, despite the mention of the purportedly 'closer' continuer, the answer is often 'no'. (There is another reason for including the details about what happens to the other person in this case. It blocks what seems to me a more straightforward response than the closest-continuer theory that psychological continuity theorists might make to the second case as originally presented—namely, that the reason you are filled with dread in Case 2 is that the psychological changes prior to the torture will kill you (you will die), not that you will survive the psychological changes and then be tortured. Being told that your psychology will be realized in another brain and body, blocks this quick response. For then 'you', as glossed by psychological continuity theorists, will not die.)

Unger is another physical continuity theorist who is concerned to explain away the apparent intuitive appeal of psychological continuity accounts in cases such as Case 1, and many others like it. This is of particularly pressing concern for Unger because, recall, Unger is quite explicit in asserting both that the right account of survival is the one which accords with our 'deepest beliefs' about our survival conditions revealed in our intuitive responses to puzzle cases, and that we all share a 'deep conception' of our survival conditions, namely, a version of the physical continuity account: "Briefly and roughly, one of us will survive from an earlier to a later time when whatever physically realizes that person's core psychology continuously (enough) realizes that core of basic psychological capacities from the earlier time to the later time."3 This seems a very bold claim, particularly in light of the well-documented apparent appeal of psychological continuity accounts; that is, if our intuitive responses to many puzzle cases are anything to go by. Unger's response is that, in fact, our intuitive responses to many puzzle cases (unsurprisingly, those which support a psychological continuity account) are not much to go by. Our intuitive reactions to these puzzle cases fail truly to be revelatory of our 'deepest beliefs' because we have been misled by pernicious details in the description or presentation of the case, or by our prior intellectual commitments. Those puzzle cases which appear to support a psychological continuity account are too sparse in background details (leading us to make unwarranted or

---

sometimes even incoherent assumptions), or are misleadingly and question-beggingly described (as, for example, when the case of teletransportation is described as *teletransportation*; or when the person emerging from the process is called ‘you’); or reflect not the untutored intuitions of common folk, but the preconceived and vested intellectual views of philosophers.  

Unger may be right to caution us about the influence that these sorts of presentational details may have on people’s intuitive responses. However, we need some principled, independent grounds for discounting some of our reactions as based on misleading features in the presentation of puzzle cases, whilst elevating others as truly revelatory of our deepest beliefs. That is to say, we need some principled way of distinguishing between distorting and non-distorting features of puzzle cases. For psychological continuity theorists may (and indeed, often do) explain those of our intuitive reactions which seem to support a physical continuity account in exactly the same way—by pointing to (different) misleading features in the presentation of the case or to the prior intellectual commitments of respondents. Moreover, by Unger’s criteria, if either of the two cases just presented are misleadingly described, it is Case 2—the case which seems to support of physical continuity account—wherein the person throughout is described as ‘you’, not Case 1. This is not to say that Case 2 could not be redescribed ‘properly’ in a way that may similarly support a physical continuity account, but it is to express some scepticism about Unger’s claim that our intuitive reactions to puzzle cases, once properly described, will unanimously support a version of the physical continuity account. For Case 1 doesn’t seem to exhibit any particularly misleading descriptive features; and Unger’s account, which denies that distinctive psychology is necessary for survival, doesn’t

4 Whereas Unger regards such details as irrelevant and pernicious, Stephen White (1989) thinks that they are crucially relevant. White’s claim, I think, is that what explains our intuitions about personal identity (and about what matters in personal identity) are not so much the first-order facts about physical continuity or psychological continuity or the like, but second-order facts, facts about properties (or relations) having the property of playing a certain role in the person-directed practices of the people described in the puzzle cases. White, it seems, thinks that our intuitive reactions to puzzle cases can be explained by positing the second-order property as the realizer of the survival role for us.

5 Unger (1990), esp. pp.7-15.

6 Presumably the psychological literature is the place to look for this.
seem able to explain why we have the intuition in Case 1 that the persons swap bodies. Nor does it seem able to explain why so many regard purely-informational teleporting (often, but not always, misleadingly called ‘teletransportation’) as survival-preserving. In short, whilst some have intuitive reactions which would seem to support Unger’s account, others do not. And it seems that not all of these contrary intuitions can be explained away as rooted in pernicious case descriptions or prior commitments and so not revelatory of our shared ‘deepest beliefs’, all the more so in the absence of any principled grounds for so doing.

So far, the evidence for lack of object-level convergence in the actual world has come from people’s (divergent) intuitive responses to possible cases. However, some deny that people’s intuitive reactions to possible cases are a reliable guide to their actual survival concept at all. Some, such as Wittgenstein and Quine, maintain that possible cases cannot tell us anything interesting about our actual concepts, but at most about how we predict that we would extend them were certain counter-factual situations actually to obtain.

Wittgenstein writes,

It is as if our concepts involve a scaffolding of facts...If you imagine certain facts otherwise...then you can no longer imagine the application of certain concepts.7

And, in a similar vein, Quine writes

The method of science fiction has its uses in philosophy but... I wonder whether the limits of the method are properly heeded. To seek what is ‘logically required’ for sameness of person under unprecedented circumstances is to suggest that our words have some logical force beyond what our past needs have invested them with.8

Insofar as Quine and Wittgenstein are right, and possible cases tell us, not about our prior concept, but about how we guess we would extend it, possible cases are not much use in conceptual analysis. For they are

---

7Wittgenstein (1967), Proposition 350.
8Quine (1972), p.490.
supposed to tell us about our prior concept, not prompt us to develop a new one, *a posteriori*.

Kathleen Wilkes expresses a slightly different reservation. According to Wilkes, possible cases present us with the following insoluble dilemma: either the background to such cases is the world as we know it (in which case many possible cases are impossible because they violate fundamental laws), or the background is different from our world, in which case we cannot draw any philosophical conclusions from such fantasy cases because, in a world indeterminately different from our own, we do not know what we would want to say about anything. The first horn of Wilkes' 'dilemma' rests on the assumption that the world (physical and social) comes as a big, inseparable package deal—we cannot vary the social facts, without varying the rest of the package. Suppose we grant Wilkes this horn—given all the physical facts about the world (the initial conditions and the physical laws), the world (including the social facts) could not be any different from the way they in fact are. For possible cases in the case of survival typically do vary some of the background in any case. (They are cases in which teletransportation, brain-scrubbing and fission occur). Now there is certainly a sense in which such cases do leave some things 'indeterminate' (exactly what the laws of nature are in a world where fission regularly occurs, for example). But these cases do not leave *everything* indeterminate and mysterious—they tell us, for example, everything we seem to need to know about what fission, brain-scrubbing and teletransportation involve. As a result, we *do* know what we want to say about these cases, even if some of us want to say different things.

More recently, Johnston too has claimed that we should dispense with possible cases in theorizing about survival. According to Johnston, what our apparently inconsistent intuitive reactions to Cases 1 and 2 reveal is that the concept of personal identity operative in our responses to these cases is a conception of ourselves as 'bare loci of mental life', according to which neither physical nor psychological continuity is necessary for personal identity or survival, though we may typically regard them as evidence for it. And, Johnston argues, we should be very suspicious of any method of theorizing about personal identity which delivers up and supports a demonstrably false view about our nature.
Chapter 9  

Others, however, grant the validity of the use of possible cases in the project of conceptual analysis, but downplay the significance of the apparent disagreement in the case of survival. Our intuitive reactions to puzzle cases are fallible predictions about how we would respond if certain hypothetical circumstances were actually to obtain; and we (or some of us) might guess wrong. For example, we may guess that we wouldn't get into the Teletransporter were we standing in front of one, but were we actually standing there we would take a different view. This line of response maintains that, contrary to our differing predictions, we would all converge in our opinions about whether or not teletransportation is survival-preserving, were Teletransportation actually available. Some of us just aren't very good predictors.

But dispensing with the use of possible cases, or treating our intuitive responses as (very fallible) predictions about how we think we would respond were possible cases actually to obtain, will not make the disagreement go away. For people disagree about actual cases. Is the born-again Christian the same person as the sinner who committed the crime? Will I be the same person as the senile old woman (still a person) that I shall become? Is Fineas Gage (the unlucky railway foreman who had a steel pole blown thorough his left frontal lobe while tamping dynamite and underwent radical character transformation from a friendly, capable and efficient man to a man who was prone to violent, unprovoked fits of aggression; so changed was he that his friends and family insisted that he was no longer the same person) after the accident the same person as the person before the accident? About all of these actual cases, and quite a few more like them, well-informed people with consistent beliefs about the matter just go different ways. They diverge along exactly the same lines that philosophers such as Parfit, Perry, Shoemaker, Noonan, Unger, Williams, Nagel and others diverge about survival in the possible cases.

Once we concede, as I think we should, that people who are neither in ignorance or error simply do not converge in their opinions about exactly what relation 'survival' refers to in possible and even actual cases, what should we say about survival in the actual world? There are, I think, three possible views.
The first is that what such disagreement shows is that the concept of survival is just vague around the edges. There is a set of core, more or less common, cases about which we all agree: cases where physical and psychological continuity go together to a reasonably great degree. But, when it comes to cases outside this core—cases where physical and psychological continuity come apart—the concept of survival starts to go vague; and the explanation for intuitive divergence around these indeterminate edges is that different people think it ought to be precisified in different ways. Unsurprisingly, this is an explanation for the intuitive divergence in the case of survival which has received very little attention in the personal identity debate. For if you think that the concept of survival is vague around the edges, then you think that there is not much more to be said about it than that.

However, the claim that the explanation for the intuitive divergence about non-core cases is that the concept is vague around the edges has an unhappy consequence. The consequence is that, if this were right, we should have to say that very few people have mastery of the concept of survival, including many apparently very proficient users of the concept. For, on the assumption that the concept of survival is vague in this way, what people should be saying is that either way of precisifying it in indeterminate cases is an equally acceptable way of going. If it is indeterminate whether or not the born-again Christian is the same person as the earlier sinner, then what people should say is that the precisification according to which the born-again is the same person as the sinner and the precisification according to which the born-again is not the same person as the sinner, are equally acceptable ways of going. Similarly for possible cases, such as the case of teleportation: the precisification according to which a person survives teleportation and the precisification according to which a person determinately does not survive teleportation are equally acceptable ways of going. Insofar as people are not ecumenical in this way—and, needless to say, many are not—they fail to have mastery of the concept of survival. On the plausible assumption that most of us have reasonable mastery of our own concepts, it cannot be that the concept is vague.

If it seems unlikely that the explanation for the disagreement is that the concept of survival is vague, what might the explanation be? According to
pluralism, we should say this: the disputing parties mean the same thing by 'survival' (they mean 'whatever plays the survival role for a community'), it is just that that role is multiply realized: different relations realize the survival role for different people. However, there is one, last competing explanation. The competing view, the rigidifying view, would have it that the two groups mean something different by survival—they have different survival concepts, survival and survival*. If survival rigidly picks out physical continuity (as Unger, for example, claims it does) then people such as Parfit, Perry, Lewis and Shoemaker, who use the word 'survival' to refer to psychological continuity, mean something different by survival: they mean survival*.

Note that, whichever of these two remaining ways we choose to go (rigidifying or pluralist), will signal an end to personal identity debate as we know it—namely, as a shoot-out between physical and psychological theories, one or other of which must be mistaken. For whether we take pluralism or the rigidifying view to be the right moral to draw from the debate, the disagreement between physical and psychological continuity theorists will be a 'no-fault' disagreement. If pluralism is right, advocates of the different views just blamelessly organize their person-directed practices and concerns around different relations; if the rigidifying view is right, then they just mean different things by 'survival'.

The dispute between pluralism and the rigidifying view is an empirical dispute about what people mean by 'survival'—do they mean 'the relation that plays the survival role for a community' or do they mean 'the relation that actually plays the survival role for me', a posteriori 'physical continuity' (as Unger claims) or (as psychological continuity theorists typically claim) 'psychological continuity'? Prima facie, the dogmatic adversarialism which has characterized the personal identity debate might seem to suggest that people are rigidifiers, not pluralists. For pluralism requires that people think that, although they would not survive some process which by their lights they do not survive, other people, with different lights, might. And although few in the debate have actually claimed that

---

9Unger's meta-theory (that survival is (rigidly) whatever relation we most deeply believe it to be) may, in these circumstances, lead Unger to maintain that survival rigidly picks out the disjunction of physical or psychological continuity. So that the two groups do not, after all, mean different things. Unger would then be a pluralist of the more constrained variety.
advocates of opposing views simply mean something different by 'survival', they have commonly claimed that advocates of opposing views are mistaken—a person's survival (any person's survival, not just their own) requires the obtaining of the relation that they care about, and those who think otherwise are just wrong. However, as I mentioned in chapter 1, the difficulty in resolving this question decisively one way or the other here and now is that the standard sorts of puzzle cases employed in the personal identity debate are not of a sort which will help us answer this question. But the Williams' cases are an exception. For they do test people's intuitions about the meta-issue; and, what's more, provide empirical data to suggest that people actually are pluralists about survival. For, insofar as people allow that the person in Case 2, despite loss of psychological continuity, might nonetheless throughout properly be called 'you', they allow that the issue of survival can be settled a different way from the way in which it is settled in Case 1 where 'you' instead refers to psychological continuity. And this can only be explained if people's pre-theoretic semantic intuitions are pluralist, not rigidifying.

Here is an analogy. Many philosophers maintain that water is rigidly H₂O. According to these philosophers, what people mean when they say 'water' is H₂O (not 'the stuff that falls from the sky, runs from taps, is drinkable etc'). We might test this claim that what people mean by 'water' is H₂O by presenting people with a thought experiment which begins 'Imagine a world in which water is not H₂O...'. Now, if people are rigidifiers about water they ought not know what we mean by 'water' in such a case. What is the stuff that is not H₂O; how is that 'water'? Insofar as people do not have this reaction they do not straightforwardly mean H₂O by 'water'. Insofar as they can make sense of the thought that there might be something properly called 'water' in a world devoid of H₂O they must mean something more than merely H₂O by 'water'. They must mean something like 'the stuff that plays the water role'.

Just so in the Williams' case. By describing the person whose psychology is wiped as 'you', Case 2 is asking people to imagine a world in which the issue of survival is settled in a different way from Case 1; where 'survival' (or 'you') refers not to psychological continuity among person-stages but to physical continuity. And insofar as people understand what
'you' means in Case 2, they are not rigidifiers. For if they thought that 'survival' straightforwardly meant psychological continuity (as their reactions to Case 1 might suggest), they ought not find coherent the description of Case 2 wherein 'you' refers instead to physical continuity. Insofar as they can understand what it means to call a person who is not psychologically continuous 'you', they cannot mean by 'you' simply psychological continuity. By the term 'you' they must mean something like 'whatever relation plays the survival role'.

The rigidifier may respond to this as they have responded to similar cases elsewhere (in particular, in the case of 'water'). Namely, by granting that it is epistemically possible that 'survival' might be used to refer to physical rather than psychological continuity, or vice versa. But, they will say, although epistemically possible, it is not logically or metaphysically or conceptually possible. It is logically necessary that 'survival' refers to psychological continuity, or whatever relation it is that the rigidifier claims that 'survival' rigidly designates. Now, this might be a valid move where we have some independent reason or argument for thinking that 'survival' is a rigid designator. For example, that 'survival' names a natural kind. But, in the case of survival, our intuitions are the only argument and data there is. And, prima facie, these support pluralism.

Insofar as people are willing to allow that 'survival' might properly (or not improperly) be used to refer to a different relation, then, they are pluralists, not rigidifiers, about 'survival'. Now, it may be that some people (such as Unger) think that they and others are not willing to allow this; that survival for any community is what is survival for us and that another community who cared about something else would not survive (although, I think, our intuitive responses to the Williams cases provide prima facie evidence against this). However, the pluralist might, if she wanted, concede that Unger is right about this. Perhaps some people do think, wrongly, that they use 'survival' in this way. Certainly, the Somataphiles as I originally presented them used 'survival' in the way in which Unger, as a rigidifier, must claim that we all do. The Somataphiles, recall, thought that no one, not even the Teletransporters, survived teletransportation, because teletransportation did not preserve the relation that they, the Somataphiles, cared about. But, as I went onto argue not in so many words, insofar as the
Somataphiles thought this they thought it because they uncritically and mistakenly assumed practice-independence: they believed that ‘survival’ latched onto a natural kind, or some other such independently privileged thing. Perhaps, in similarly having such false meta-level beliefs, some of us are like the Somataphiles. The question then is: would we still be like the Somataphiles once we ceased to have such false meta-level semantic and/or metaphysical beliefs? Would we still persist in our belief that ‘survival’ was a rigid designator after informed deliberation?

Imagine that we come to have substantial interaction with a community in which fission is a regular occurrence, and who organize their person-directed practices around the assumption that only one of the fission products is a survivor of the original person (the other being a wholly new and different person). In this community, rights, legal entitlements, ongoing obligations, personal relationships etc. are inherited by the fission product, call him or her, Righty, who emerges first (milliseconds before the other, Lefty). What would we think of this community’s view of survival? Initially, I think, we would try to convince them that they are wrong about survival: that is to say, that they are mistaken to treat Righty as the same person as the original, and Lefty as a new and different person; that this is arbitrary discrimination against another perfectly good successor, and so on. In just the same way as, initially, the Somataphiles tried to convince the Teletransporters that they were wrong about survival, and vice versa. However, suppose that our protests were to no avail—they continue to regard Righty as the only legitimate successor of the pre-fission person. In our ongoing interactions with them would we conclude, after sufficient and informed deliberation, that they’re not talking about survival or would we conclude that they are talking about survival, it is just that what relation is survival for them is not the same as what is survival for us.

Pluralism, as an analysis of how we use the word ‘survival’, stakes much of its claim on the latter: that, if not immediately, then certainly after a little deliberation, we would come to believe that, without having changed the subject, they simply and blamelessly settle the issue of survival a different way. And the Williams cases provide prima facie support for thinking that we are at least willing to entertain that this might be the right thing to say about such a community. Insofar as we may be reluctant to allow that this
community are still talking about ‘survival’ it will be because we are parochially attached to the particular relation that we care about. There is nothing wrong with parochial attachment of this sort, except that, in the case of survival, it conflicts with another central intuition: the intuition that survival, whatever it is, must be principled. Surely survival for every community shouldn’t depend on something as utterly psychologically contingent as what we happen to care about.10

I think that pluralism is right as a theory about how we use the word ‘survival’. And my hope is that, once others reflect upon the meta-issue, they will come to realize that they are pluralists too. But I concede that the evidence from empirical psychology is far from entirely in as yet. (At least now, I hope, having seen the need for such empirical investigation into the meta-issue, we can begin it in earnest.) However, suppose that further investigations in empirical psychology deliver the conclusion that many of us are rigidifiers, not pluralists, about survival; that we refuse to allow that ‘survival’ properly refers to any relation other than the one or ones to which we use it rigidly to refer. Would the bell then toll for pluralism?

Well, it would signal the end to pluralism as a piece of descriptive conceptual analysis: as an analysis of how we all use the word ‘survival’. But pluralism may yet retain a vital place in analysing survival. For pluralism will then be the theorist’s view. The theorist may not, and typically does not, care whether their categories coincide with the categories that ordinary folk use. For the way ordinary people use words may not carve up the world in any particularly theoretically useful or illuminating way. Anthropologists investigating the cultural practices of different communities, for example, typically do not care greatly about whether the categories that they use to

10 Someone who is unconvinced by this argument will have at least the following take-home argument for a watered-down pluralism—that both role and actual realizer are essential for it being determinate whether or not a person survives. In the absence of either—where something plays the survival role for a community but that relation is not what actually realizes the survival role for us, or where there is the relation that actually plays the survival role for us but it does not play the survival role for a community—it is simply indeterminate whether or not people in that community survive. On this view, as against the rigidifier, that something plays the survival role for a community is essential for it that relation counting as survival, but (with the rigidifier, as against the pluralist) this view has it that actual realizer is also required for it to be determinately true that a community are talking about survival. As earlier noted, something along exactly these lines seems to be David Wiggins’ view.
investigate communities are endorsed by the communities that they are investigating. When they are looking with interest into different conceptions of the good life for example, they would not abandon the whole project simply because they discover in the course of these investigations that the members of the communities investigated themselves regard each other as meaning different things by ‘the good life’ (although they may record this as an interesting and relevant piece of sociological data in itself). For they are interested in the commonalities between the two communities; commonalities (such as the common motivational role that different properties play) which would be denied and obscured were they to adopt the view taken by the communities as their theoretical kind. Just so, I think, if it turns out that ordinary people use the word ‘survival’ to refer only to the relation that they themselves care about, then it will turn out that the folk use the word ‘survival’ in a theoretically uninteresting way. The theorist who wants to understand and explain interesting commonalities between communities who organize their person-directed practices and concerns around different relations (such as the Teletransporters and the Somataphiles) needs a different category to carve up the world in an sociologically or anthropologically interesting way; and that category will be pluralism.
If pluralism is right, then we should revise our beliefs not merely about what to say of other communities in other worlds, but also about what to say about survival in the actual world. Moreover, we ought (the 'ought' of instrumental rationality) to become minimalists about survival. However, the implications of pluralism are not confined to the case of personal identity or personal survival alone. From pluralism about survival, I think, we can draw some more general lessons for issues which are of substantial contemporary concern in moral and political philosophy. In particular, we can draw some lessons about the relationship between individual's attitudes of identification (with nation, culture, gender, race and class) and the question of personal identity, with which philosophers wanting to explain such phenomena as nationalism, gendered identity and ethnicity, amongst others, have been increasingly concerned. I want very briefly, in closing, to begin to suggest what some of these morals might be. To explore these implications in greater detail is a further, future project.

Many social and political philosophers, most notably (but not exclusively) Marxists, communitarians, post-modernists, feminists and post-structuralists, have been anxious to insist that 'self-identity' or 'personal identity' is socially constructed. Indeed, this has become almost a truism in most contemporary moral and political philosophy. Insofar as these moral and political philosophers have engaged at all with the personal identity debate in analytic philosophy, it has typically been with an eye to pointing out the utter irrelevance of this debate for the issues—national identity, ethnicity, gendered identity, racial identity and the like—with which they are concerned. The personal identity debate in analytic philosophy is, to their mind, not only inanely insular and parochial, but completely misguided.

These objections to analytic discussions of personal identity find their clearest (and most sympathetic) expression with Ross Poole, in the context of a discussion of nationalism. Poole, unlike many others, does not despair of the possibility of analytic accounts of personal identity contributing something of interest and importance to an understanding of nationalism. But he does think that analytic accounts of personal identity

---

1See, for example, Jonathan Ree (1992).
will not do as they stand. Poole’s objection is that, since Locke, analytic accounts of personal identity have got the direction of explanation between social practices attitudes and personal identity the wrong way round. He writes,

Whilst most subsequent philosophy has rejected Locke’s account, it has accepted his way of posing the problem...The Lockean tradition assumes that personal identity is a physical or psychological relation which underlies and helps explain certain social practices, and has sought—unsuccessfully—to discern what that relationship is.²

This sort of approach, Poole thinks, is of little use in moral and political philosophy because "understanding personal identity is not a matter of discerning continuing essence, but of discerning the different social practices which construct and sustain it".³

The sort of account of personal identity that Poole and others think we need for an understanding of social and political phenomena such as nationalism is an account which Poole himself traces to Hegel, Nietzsche and Marx. It is an account which "reverses the direction of explanation: it is the social practices which underlie and explain personal identity"⁴, not personal identity which underlies and explains the social practices. In short, I think, the account Poole and others want is a practice-dependent account of personal identity of just the sort I have advocated.

At this point, analytic philosophers might well reply that it is no surprise that their accounts of personal identity do not serve to shed much light on the connection between personal identity and social attitudes of identification. For what Poole and others want when they want an account of personal identity which explains the social practices which construct people’s identities is an account of qualitative personal identity; an account of how and why people form the various identifications with nation, gender, race, class or culture, that they do. And this is a quite different and separate question from the question of what makes a person one and the same numerical individual persisting through time. When Poole and others criticize analytic philosophy for failing to deliver the sort

²Poole (1992), p.15. For the reasons discussed in Chapter 1, I think that Locke is an unfortunate choice of example here. For, if there is one philosopher in the analytic tradition who advocates just the sort of account of personal identity Poole wants, it is Locke.
³Ibid., p.16.
⁴loc.cit.
of account of personal identity that can shed light on the question of how and why people form the identifications they do, they have made an elementary confusion: they have simply failed to realize that the question of qualitative personal identity is different from the question of numerical personal identity; and that it is the latter, not the former, question that analytic philosophers have sought to answer in discussions of personal identity. They have failed to realize that just because the words ‘personal identity’ are the same, it does not follow that they are being used in the same sense or for the same question. Most analytic philosophers, I think, even practice-independent ones, would be more than happy to grant that a person’s qualitative identity—the sort of person that a person is; the beliefs, desires and behavioural dispositions that they have—is determined, to a significant degree, by social practices. But, they will say, although the two sorts of personal identity masquerade under the same name, the question of qualitative personal identity is a quite different, and independent, question from that of numerical personal identity; and numerical identity, unlike qualitative identity, is determined quite independently of social practices.

Between these two different philosophical traditions and questions, however, there is some middle ground; but it takes practice-dependence about (numerical) personal identity to see it. For, once we are practice-dependent about (numerical) personal identity, these two questions (the question of numerical personal identity and the question of qualitative personal identity), though certainly different questions, are not so independent as philosophers of personal identity in the analytic tradition have typically thought. For on a practice-dependent account, the concept of (numerical) personal identity is a cluster concept; and amongst the cluster may well figure platitudes about the importance for numerical personal identity of continuing identification with a nation, a gender, a class, a race or a culture. \( X \) identifies as \( Y \) just if \( Y \) is a member of the cluster, a weighted most of which is sufficient for numerical identity. The question of numerical identity is then much more closely associated with the question of qualitative identity than previously thought. For a person’s numerical identity is defined, \textit{inter alia}, in terms of social attitudes and practices which were previously thought the unique and distinct province of qualitative personal identity. And now, moreover, we have an explanation as to why the two different questions are both questions about ‘personal identity’ nonetheless.
Exactly how closely entwined the two sorts of personal identity are is an interesting question. As I noted at the outset, too much qualitative change can destroy numerical identity. Might it be, then, that ceasing to identify with a nation, culture, race, class or gender might destroy numerical identity? Might it mean, not merely that a later person is a very different sort of person to an earlier person, but that, as a result of such refiguration, the earlier person literally ceases to exist? Temporal-phase pluralism certainly happily accommodates this possibility. To pick just one topical example, it might be that, in ceasing to identify as a lesbian (by, *inter alia*, becoming attracted to men) a person may cease to exist, constructing a new and different person, a person who may (or may not) regard earlier lesbian-identifying person-stages as parts of them. This would give a literal gloss to the claim that some lesbian-feminists and queer-theorists have made, that ceasing to identify as a lesbian means, not merely that the lesbian ceases to exist, but that the self itself ceases to exist.

I must say that it seems to me unlikely that, where the platitude about continuing to identify as a lesbian figures amongst the personal identity cluster at all, it figures sufficiently heavily weighted so that to lose it alone would constitute a person’s literal death (although, if sufficient other of the platitudes were gone, ceasing to identify as a lesbian might be the final straw that makes for a person’s death). Whatever the plausibility of these claims, however, it is an advantage of temporal-phase pluralism, I think, that, alone amongst accounts of personal identity in analytic philosophy, it can give rigorous and intelligible expression to these popular claims. These further, interesting applications of pluralism I leave for another time.
BIBLIOGRAPHY


