

Review Copy

Personal Identity and Ethics
A Brief Introduction

David Shoemaker



BROADVIEW GUIDES to PHILOSOPHY

© 2009 David Shoemaker

All rights reserved. The use of any part of this publication reproduced, transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, or stored in a retrieval system, without prior written consent of the publisher—or in the case of photocopying, a licence from Access Copyright (Canadian Copyright Licensing Agency), One Yonge Street, Suite 1900, Toronto, Ontario M5E 1E5—is an infringement of the copyright law.

Library and Archives Canada Cataloguing in Publication

Shoemaker, David, 1964-
Personal identity and ethics : a brief introduction / David Shoemaker.

Includes bibliographical references and index.
ISBN 978-1-55111-882-6

1. Self (Philosophy). 2. Identity (Philosophical concept). 3. Ethics.
I. Title.

BD450.S45 2008 126 C2008-904149-6

Broadview Press is an independent, international publishing house, incorporated in 1985. Broadview believes in shared ownership, both with its employees and with the general public; since the year 2000 Broadview shares have traded publicly on the Toronto Venture Exchange under the symbol BDP.

We welcome comments and suggestions regarding any aspect of our publications—please feel free to contact us at the addresses below or at broadview@broadviewpress.com.

North America

PO Box 1243, Peterborough, Ontario, Canada K9J 7H5
2215 Kenmore Ave., Buffalo, NY, USA 14207
Tel: (705) 743-8990; Fax: (705) 743-8353
email: customerservice@broadviewpress.com

UK, Ireland, and continental Europe

NBN International, Estover Road, Plymouth, UK PL6 7PY
Tel: 44 (0) 1752 202300; Fax: 44 (0) 1752 202330
email: enquiries@nbninternational.com

Australia and New Zealand

UNIREPS, University of New South Wales
Sydney, NSW, Australia 2052
Tel: 61 2 9664 0999; Fax: 61 2 9664 5420
email: info.press@unsw.edu.au

www.broadviewpress.com

This book is printed on paper containing 100% post-consumer fibre.

PRINTED IN CANADA



Introduction

Motivation

Consider the following six cases.

Case 1: Gretchen has been in a terrible motorcycle accident, which has caused such internal damage that she will die in just a few days. A lifelong atheist and advocate of the view that “once you’re dead, you’re dead,” Gretchen now begins to wonder if she’d been wrong all along. She would, after all, very much like to be able to anticipate some kind of continued survival after the death of her body, but such anticipation will be rational only if it’s possible for her to survive the death of her body. But is it? Her body and brain will in fact cease to exist in a few days (she has requested cremation after she dies), so how could it even be remotely possible that she, Gretchen, will survive if her body and brain won’t?

Case 2: Carlos and Tanya are having a discussion about abortion. Carlos says, “Abortion is wrong. It would be wrong to kill me, wouldn’t it? Well, that fetus from which I developed was also me, so it would have been wrong to kill it as well.” “I disagree,” says Tanya. “While you certainly developed from a fetus, that fetus wasn’t you, just like an acorn isn’t an oak tree. What you are is a person, after all, a being with the capacity for not only consciousness but also self-consciousness, whereas an early-stage fetus, say, has neither capacity. Thus, insofar as you aren’t the same sort of things, it could be wrong to kill you without being wrong to kill the fetus from which you grew.”

Case 3: When Meredith is 55 she is diagnosed with early-stage Alzheimer’s. She knows exactly what the disease does to one’s mind, since she watched her mother

die from it. She thus signs an advance directive instructing doctors that they are not to use any extraordinary means to keep her alive if she gets seriously ill after becoming demented. But once Meredith actually gets to that demented state and contracts pneumonia, it turns out that she is perfectly content and, when asked, expresses a preference to stay alive. Whose wishes are to be honored here, the early-stage patient or the late-stage patient?

Case 4: Howard and Annie are top-notch reproductive scientists. Annie has made a breakthrough in her research that would enable her to clone an adult human being, that is, to take one of his cells, coax it back into an undifferentiated state, combine it with an egg whose nucleus has been removed, and then implant the resulting zygote in a woman's uterus, from which a human being with the same genetic structure as the original cell donor will eventually be born. Howard is horrified by the prospect of cloning, however, and objects by saying that cloning someone would be wrong insofar as it would rob the clone of his own unique identity; instead, he'd just be a copy of someone else. Annie scoffs, noting that human identity has nothing to do with genetic or physical structure; instead, it's entirely about psychology, and given that the clone would grow up in a very different environment from the original, he'd certainly develop psychologically in very different ways from the original, leaving them both with two distinct identities.

Case 5: Sitting around with his family after Thanksgiving dinner, a slightly tipsy Phil laughingly brings up an old family story about how when he was ten years old he caused his younger sister Jen to fall out of a tree, breaking her arm. "I still blame you for that," Jen says, suddenly quite serious, "I get angry at you whenever I think of it." "Oh, c'mon, Jen," replies Phil, "it was thirty years ago and I'm nothing like that ten-year-old anymore. Surely you can't still be mad at me for what that stupid little ten-year-old did!" "Oh, I am," mutters Jen, "because no matter how much time has passed or how much you may have changed, you are still the one who pushed me out of that tree."

Case 6: Darren and Samantha have two young sons, Brad and Albert. Brad is perhaps the cutest baby of all time, and so his parents allow him to model infant clothes and, as he gets older, toddler clothes. Brad has very sensitive skin, though, and the

Review Copy

clothes he models aren't made from very good fabrics, so he gets a skin irritation every time he does a modeling job. Nevertheless, he gets paid a lot of money, all of which his parents save in a special account. Albert, on the other hand, while not terribly cute, is obviously a budding genius. Darren decides one day that, when the boys get older, they should take all the money earned by Brad's modeling career and give it to Albert so he can afford to go to the best schools: it will make the most good come out of that money. Samantha vociferously disagrees, saying that Brad was the one who sacrificed for that money, and so he's the one who should get it as compensation. Giving it to Albert would simply be unfair, she thinks.

All of these cases are in some way about the relation between personal identity and ethics. **Ethics**, very generally, is about the way(s) in which we ought to live our lives. This includes the actions we ought or ought not perform, the attitudes and concerns we ought or ought not have, and the character traits we ought or ought not develop. Many people take the term "morality" to be interchangeable with the term "ethics," and that will be fine for our purposes.¹ In addition, for many the term "ethical" is interchangeable with the term "practical," and this is also a substitute you will see throughout the book. In particular, one might construe this book as being about the relation between personal identity and our **practical concerns**, those practices and patterns of caring that are central to the living of our lives, both with respect to our treatment of ourselves—call these our **self-regarding practical concerns**—and our treatment of others—call these our **other-regarding practical concerns**. The relevant self-regarding practical concerns here include our anticipation of the future, the special sort of caring we have only for ourselves, and our concern to survive into an afterlife. The relevant other-regarding practical concerns here include our attitudes toward specific moral issues like abortion, cloning, advance directives, and so forth, as well as our more abstract commitments to various practices of moral responsibility, compensation, and fair distribution of resources.

1 Others think ethics is better construed as the *study* of morality, where morality consists of the principles and standards that guide action and character. This is a narrower understanding of "ethics" than we will employ.

Now the cases described above are about the various interesting and fruitful ways in which considerations of personal identity may be particularly relevant for these practical/ethical concerns, be they self-regarding or other-regarding. To see this, consider the cases in more detail.

Case 1 is about the relation between identity and the rationality of anticipating survival in some sort of afterlife. It seems, after all, that for Gretchen rationally to anticipate surviving her death, it has to be possible for there to exist someone in heaven (or hell!), say, who *is* Gretchen, and who is not just a replica of her. But this way of putting the matter implies that, in general, it is rational for me to anticipate only *my own* future experiences, which on its face seems quite true: surely I can't anticipate *someone else's* experiences! So while I may worry about what it will be like for you to undergo some experience, or even have a genuine understanding of what it will be like for you, I cannot look forward to *having that experience* in the way you can. And closely aligned with anticipation of experiences is a special kind of concern, a concern you have, it would seem, only for yourself; indeed, it's what we typically call "self-concern." Now to appreciate the "specialness" of this type of concern, suppose I tell you that I've managed to create a full blown replica of you in the next room, an actual human being who's exactly like you in every respect, and I further tell you that tomorrow morning I'm going to torture him for several hours. Surely you'll be concerned for him, perhaps even fearful on his behalf. But notice the difference in the attitude you'll have were I to tell you instead that I'm going to torture *you* tomorrow morning. Now the type of concern you'll have is of a different sort, and its difference seemingly consists in your being able to anticipate the experiences of the torture victim in the second version of the story, but not the first. The basic lesson, then, seems to be that both rational anticipation and this kind of special concern depend in some crucial way on personal identity, on the anticipator and anticipatee, the concerned and the concerned-for, being one and the same person.

Cases 2–4 are all about the relation between personal identity and various issues in the field called **applied ethics**, wherein one attempts to apply abstract moral theories and principles to concrete, real-life cases. *Case 2*

focuses on the key issue of personhood and our essential nature. What am I, precisely? Was I ever actually a fetus, or did I—a person—come into existence at some point after that fetus, that is, did I develop *from* a fetus? If the former, then given my obviously significant moral status now (I have a right to life, surely), wouldn't such moral status have to be shared as well by all the parts of my past, lending moral protection to fetuses generally? If the latter, then wouldn't my moral status as a person simply be fundamentally different from that of fetuses, so that what it's wrong to do to me would not apply (or would at least carry less weight) with respect to fetuses? The import of these questions for our general topic is that the correct theory of personal identity itself seems to depend on what the actual nature is of those individuals whose identity is being tracked; that is, we can't put together the proper theory of our identity (and its relation to ethics) unless we first understand what we are. And this is a source of great controversy.

Case 3 raises different sorts of issues. Normally, it is assumed that the authority of an advance directive to determine the care of some patient stems from its being the directive of the patient herself, that is, the *signer* of the directive and the *subject* of the directive are one and the same person. If this is true, then in Meredith's case she is expressing two contrary preferences, one at the time of the signing, and one at the time of the illness. Which preference thus has authority here? Many people think that, in the case of advance directives, the preferences of the earlier self are authoritative, but this reaction seems to conflict with the way in which we treat past preferences in most other contexts. Suppose that you and I have been friends since we were 20, and when I was 21 I told you, "If I'm not a professional poet by the time I'm 40, I want you to take out an ad in the *New York Times* proclaiming what a failure I am." But now that I am forty, and not a poet, I'm likely to say something like, "Do you remember that crazy demand I made of you at 21? Please just forget about it—I'm so embarrassed by my pretentiousness back then." In this case, we are likely to think the earlier preference is no longer binding; the preferences I have *now* are what counts. So why not think this in Meredith's case? Why not think the preference of her demented self, given that it is expressed *now*,

is binding? Of course, one (serious) difference is that Meredith's later self is indeed demented, so it may be that we need to decide what would be best for the patient in a way we don't do when it comes to the poet promise, say. But *who* is the precise subject here? Is it simply the unit comprised of the Later Meredith (LM) and the months she has left? Or is it the unit comprised of Meredith's entire life, in which case these last few months are a kind of tragic coda, which will, if allowed to occur, undermine many of the goods of Meredith's life as a whole? How we answer this question will go a long way towards determining how Meredith is to be treated here.

On the other hand, it may be that LM and Earlier Meredith (EM) are actually different persons, or that LM *isn't even a person at all*. Consider the former possibility: even if LM and EM are different persons, that doesn't yet answer the question of which one's preferences are authoritative. For while it's true that I typically have no say in the medical care someone *else* is to receive, this isn't always the case. After all, I *do* have a say in the medical care my child, or another close relative, is to receive. And so it may be the case that, even in the absence of identity between EM and LM, EM's directive is authoritative. What, though, if LM isn't a person at all? Wouldn't this then imply that the preferences of EM—a person—are authoritative? Maybe not. After all, even if LM isn't a person, that doesn't yet mean that EM and LM aren't still *identical*, as intimated in our brief discussion of *Case 2*: if what we are essentially *isn't* in fact persons, then the fact that some stage in my life isn't a person may be irrelevant to determining which preferences are authoritative.

Case 4, on cloning, is really about the relation between personal identity and *uniqueness*. Uniqueness is something many of us value, and we think that what makes us unique *just is* what determines our personal identity. On this view, however, cloning me would rob both the clone and me of uniqueness, and thus rob us each of a personal identity. But is genetic identity what matters for *personal* identity? If so, then wouldn't all identical twins lack both uniqueness and a personal identity? Perhaps, though, the intentional creation of a clone is a relevant difference here, or perhaps identical twins *do* lack an important kind of uniqueness. At any

rate, more needs to be said to identify the essential features, if any, of a person's (unique?) identity.

Case 5 is about moral responsibility, something that seems to depend on a principle about personal identity: one person can be morally responsible only for his own actions, and so he cannot be morally responsible for the actions of someone else. The dispute between Phil and Jen, however, isn't over this principle—a principle they both accept; rather, it's over whether or not Phil is in fact the same person as the little boy who pushed his sister out of the tree. Jen insists that of course he is; who else could that boy have been? Indeed, doesn't Phil now say things like "Remember when *I* pushed you out of that tree"? Phil, on the other hand, while perhaps agreeing that, in one sense he was that little boy, disagrees that that's the sense of identity that *matters* with respect to things like moral responsibility. Instead, Phil is suggesting that the type of identity that matters is in some crucial way merely psychological. In other words, when a person changes a great deal psychologically, his identity (in this different sense) changes, and so he may no longer be responsible for what he did prior to the changes. And many people would agree with respect to legal responsibility as well: when a hardened criminal is genuinely converted to Christianity while in prison, there are many who would advocate absolving him (the "new" him) of his (the "old" his) crimes. So which sense of "identity" is appropriate for cases of moral and legal responsibility? Furthermore, is *any* sense of "identity" appropriate? After all, perhaps what Phil is suggesting is that, while he is in fact the same person as that little boy, that fact about identity is just *irrelevant*, given the abundant psychological changes that have taken place since then. What matters for attributions of responsibility, then, might instead be about the psychological relations—or perhaps even some *other* kind of relations—between the blamee and the original agent.

Case 6 is about compensation. Here, most of us would agree that compensation is something that, once again, presupposes personal identity: I can truly be compensated only for my own past sacrifices, and I cannot be compensated for burdens I have undergone by benefits being given to someone else. This is, after all, the principle Samantha clearly seems to accept in maintaining that benefiting smart little Albert wouldn't count

as compensation to cute little Brad for the sacrifices Brad made as a child. But not only is she making a conceptual point, she is also maintaining that it's *wrong* to benefit Albert for the burden undergone by Brad, that, more generally, uncompensated sacrifices are just unfair. Darren, on the other hand, disagrees, maintaining that it's not immoral to benefit Albert with the money made from Brad's burden, given that this distribution will actually make the world a better place (perhaps Darren has noticed that Brad has become a bit of a slacker these days, and so would just squander the money). This is a true ethical dispute, but it's a dispute that nevertheless may depend on the truth about personal identity, in several different ways. First, if it turns out that 18-year-old-Brad is actually a *different person* than two-year-old-Brad—given the vast psychological and physical differences between them, say—then Samantha's worries about compensation are moot: giving the money to Albert would not be unfair, given that two-year-old-Brad no longer exists. Second, even if the two "Brads" are indeed the same person, that fact may simply not be very *morally* important, given the vast psychological and physical differences between them. In other words, the moral importance of personal identity may vary in proportion to how *strong* the identity relation actually is—perhaps, after all, identity comes in degrees and so perhaps the tighter the various relevant psychological or physical strands are, the more morally important Samantha's principle of compensation becomes. Finally, there is clearly a sense in which the two "Brads" are the same person, and if this is the sense of identity that matters morally, and if Samantha's principle about compensation is indeed an important moral factor, then it is difficult to see how Darren's view is very tenable.

These are just a few of the many interesting and important issues that arise at the intersection of personal identity and ethics, but they should be enough to motivate our project. In general, people are led to investigate the nature of personal identity precisely because of its relation to our practical concerns, but, as we will see, what they find might surprise them, perhaps even pushing them to reconsider the nature and significance of the practical concerns with which they began the investigation. In other words, what would seem to be a straightforward relation between personal

identity and ethics may turn out to be anything but. In what follows, then, we will explore in detail this tangled and multifaceted relation. In doing so, we will try to sort out and understand the various contenders for the “true” theory of personal identity, along with the implications each theory would have for our practical concerns.

Before embarking on that project, however, we need to take the time to discuss some crucial concepts and distinctions, which we will do in the remainder of this Introduction. The book from there on out is divided into two parts, each focused on a different domain of ethics. In Part A, “Personal Identity and Self-Regarding Ethics,” we will focus on identity and the practical concerns we have regarding what we ought to do and care about with respect to *ourselves*. In Part B, “Personal Identity and Other-Regarding Ethics,” we will focus on identity and the practical concerns we have regarding what we ought to do and care about with respect to *other people*.

In Chapter One, then, we will begin our exploration of the relation between personal identity and self-regarding concerns, motivated by the issues raised in *Case 1* about the rationality of anticipation and the possibility of immortality. In doing so, we will lay out four rather crude theories of personal identity, theories focused individually on souls, memories, bodies, and brains, and we will discuss crippling problems with all four. In Chapter Two, then, we will attempt to develop two much more sophisticated theories of identity—the Psychological Criterion and the Biological Criterion—still with an eye towards their relation to self-regarding concerns (in particular to anticipation and self-concern), and we will discuss various strengths and weaknesses of each theory. In Chapter Three, spurred by a kind of standoff we’ll be left with in Chapter Two, we will investigate two much more radical approaches to the issue, namely, a view called “narrative identity,” and a view that personal identity is in fact not what matters for anticipation and self-concern. While there will be problems with these two theories as well, they will join the Psychological Criterion and the Biological Criterion as legitimate contenders for the “right” theory to anchor the relation between identity and ethics throughout the remainder of the book.

We will then turn to our other-regarding concerns in Part B. We will begin, in Chapter Four, with some important identity-related moral conundrums at the beginning of life: abortion, stem cell research, and cloning. We will then talk in Chapter Five about genetic intervention (both prenatal and postnatal genetic therapy), and any obligations we might have regarding bringing entire populations of people into existence. In Chapter Six, we will turn to a discussion of a couple of important moral issues at the end of life, including advance directives and causing the “deaths” of multiple personalities during therapeutic treatment of that psychological disorder. In Chapter Seven, we will explore the issue of moral responsibility, which has been the main focus of many people interested in the topic of personal identity over the years (and was the subject of our *Case 5*). In Chapter Eight, we will turn to a discussion of the relation between personal identity and ethical *theory*, which will include not only a discussion of the compensation issue of *Case 6*, but also a discussion of which theories of identity bolster the plausibility of which theories of ethics.

In the brief concluding chapter of the book, we will switch gears rather dramatically in order to discuss outright an important abstract question that hangs implicitly over the entire enterprise, namely, what is the right or best *method* for investigating the relation between personal identity and ethics? There are three assumptions about method made by most writers on this relation. First, they assume that the *motivation* for an investigation into the nature of personal identity comes from our practical concerns, that it's because we want answers about various practical questions that we are led to explore identity. This is not the only motivation available, however, for someone might be interested in personal identity solely insofar as it's an interesting metaphysical issue *independently* of our identity-related practical concerns. It's quite possible, though, that these different motivations may yield different results, about both the nature of identity and its implications for our practical concerns. Now it should be obvious that we are adopting the more popular assumption here, given that we have started off explicitly with cases of practical concerns to motivate the project. But it will be worth keeping in mind the alternative route into the

project, and in the last chapter we will discuss whether or not we should be worried by possible conflicting methods here.

The second methodological assumption of most writers is that personal identity is prior to ethics,¹ so that our practical concerns ought to answer to, and so be revised in light of, the correct theory of identity. This will also be our default method. We will thus be applying various plausible theories of identity to our ethical issues, assuming that if one particular theory is true, it will have one set of implications for our practical concerns, but if another theory of identity is true, it will have a *different* set of implications for our practical concerns. Nevertheless, this method may be questioned. Instead, one might believe that ethics is actually prior to personal identity, so that our theories of identity must answer to, and so be revised in light of, our practical concerns. This conflict will arise in Chapter Eight, on identity and ethical theory (given that it's a challenge presented by some ethical theorists), and we will address it both there and in the concluding chapter.

The final assumption of most writers on this topic is that what we are seeking is *the* relation between identity and ethics, that we are seeking one theory of personal identity that can stand in the same relation to all of our practical concerns, both self-regarding and other-regarding. So, one might think, if there is a relation between identity and ethics, it's going to be between all of our practical concerns and a psychological-based theory of identity, say, or between all of those concerns and a biological-based theory of identity. And this is the default assumption we will work with as well. But as we proceed, we will come to question this assumption, given that one theory of identity will do really well at relating to some of our practical concerns but do poorly with some others, themselves which will seem more closely related to a different theory of identity altogether. We might well wonder, then, whether or not different theories of identity are relevant to different practical concerns, whether there just is no single criterion of identity that bears the appropriate relation to all our practical concerns.

1 This means that decisions must be made about identity first; only then can ethical issues which depend on these decisions be considered.

Review Copy

These are the three difficult methodological questions we will take on explicitly in the final chapter. Indeed, the book as a whole gets more difficult as it progresses, primarily because of the increasingly abstract nature of the topics. My aim, though, will be to make those more difficult aspects of the book as clear and accessible as possible and to prepare the reader throughout the journey for the challenges to come.

Concepts and Distinctions

Before we launch into the metaphysical investigations of Chapter One, we need to get clear on a few important concepts and distinctions, some of which are specific to our project, and others of which we'll run across occasionally that have general philosophical importance. What we are looking for, in working out the first part of the relation between identity and ethics, is a criterion of personal identity. Unfortunately, the terms "criterion" and "identity" are actually ambiguous. Consider the latter term first. Suppose that, when handed a photograph by your mother of your five-year-old self, you sigh and say, "I'm afraid, mom, that I am no longer the same person as that little child you loved so much." What are you saying, exactly? Now there's a sense in which you *aren't* the same person insofar as you are different (likely psychologically) from that child. But there is clearly another sense in which you *are* the same person as that child, given your clear implication that *you yourself* have changed, where change is compatible with identity.

This ambiguity in the phrase "same person" stems from an ambiguity in the more general term "identity," so it's with reference to that term that we need to make our first essential distinction:

X and Y are **qualitatively identical** if and only if they have exactly similar qualities.

X and Y are **quantitatively identical** if and only if whatever is true of X is true of Y and vice versa.

The best way to explain this distinction is by considering examples. If you bought two brand new copies of this book (go ahead, try it!), they would likely be qualitatively identical, that is, all of their qualities—shape, size, pages, words, cost, and so forth—would be exactly similar to one another. But the books nevertheless could not be *quantitatively* identical to each other. Why not? Well, for one thing, they occupy different points in space-time, so whatever is true of one (namely, its particular location in space-time) could not be true of the other. To make this point even more explicit, you could simultaneously hold one book right side up and the other upside down.

Each copy of the book, then, will be quantitatively identical *only with itself*. Quantitative identity is also known as **numerical identity**, and in both cases the relevant quantity, or number, is *one*, as in *one and the same thing*. So if “X” refers to the copy of the book held steady in your right hand, say, at *this* precise moment in time, and “Y” refers to the copy of the book still in your right hand at this *later* moment in time, then X is quantitatively identical to Y insofar as it is one and the same thing as Y. Notice further that quantitative identity, unlike qualitative identity, is compatible with qualitative changes. So if you were to take your copy of the book and then fold up the corner of the title page, it would still be the same book—quantitatively—as the book you bought, even though one of its original qualities (an unfolded title page) had changed.

One of the main questions we will be concerned with, then, is about *quantitative*, or numerical, identity: what makes me the same person as the child in that photo, even though we’re not qualitatively identical to one another (we in fact share very few, if any, qualities)?

This leads to the second distinction we need to make. Again, what we’ll be looking for is a criterion of personal identity across time, but “criterion” has two senses as well. To draw the appropriate distinction, we will use some standard philosophical terminology that is, unfortunately, rather unwieldy:

X is a **metaphysical criterion** of Y just in case X provides an explication¹ of what Y consists in, an explication of Y’s nature. (*Metaphysics* is the

1 An *explication* is a detailed, formal explanation, with attention to theoretical issues and implications.

philosophical study of the principles of existence—of beings in general, or of particular kinds of things.)

X is an **epistemological criterion** of Y just in case X provides a way of identifying Y. (*Epistemology* is the philosophical study of knowledge: what it is and how you get it.)

To understand this distinction, suppose I tell you that what makes you the same person as that child in the photo is that you both have the same *soul*. You then respond, “Well maybe, but how could we ever *know* that I have the same soul as that child? After all, the soul is supposedly nonphysical, but because we can have evidence only for the existence of physical objects, we could never have any evidence whatsoever for the existence of souls, so we could never determine whether or not I have a soul, whether or not that child had a soul, and whether or not my soul and that child’s soul are the *same* soul!” While this would be an important claim to consider, it would not be directly responsive to my initial assertion, and that’s because we are each discussing a different kind of criterion of personal identity. On the one hand, I have offered a *metaphysical* criterion: the nature of personal identity, I say, consists in persistence of the same soul. On the other hand, you have denied the existence of a particular *epistemological* criterion of identity: we have no empirical means of identifying souls, you’re saying. But then what we’re doing is just talking past one another. For even if we could never know whether or not your soul now is the same as the soul had by that child in the photograph (if you even actually have a soul at all), it could still be true that personal identity consists in sameness of soul, that what would *make you identical* to that child is the identity of your (shared) soul. And alternatively, what provides the means for our identification of the bearer of identity may be in place (or not) regardless of what actual metaphysical criterion of identity is true.

So what sense of “criterion” are we going to be interested in here? Clearly it’s the metaphysical sense of the term: we want to know what the *nature* of identity consists in, regardless of whether or not we could ever know if that nature obtains in any individual case. Nevertheless, as

we shall see, some people have insisted that there should still be a close relation between the nature of personal identity and how we can come to identify when it obtains. If, after all, I could never actually know when some (metaphysical) criterion of personal identity obtained, could it really be the right criterion? We make what seem to be justified judgments of identity all the time (for example, every time we recognize our friends). Could these really be unjustified, though? If our ordinary judgments of identity don't (or can't) track what some theory tells us is the true nature of identity, that theory may lose some points. We will say more about this problem in the first chapter.

While we will primarily be interested in a metaphysical criterion of personal identity across time, there is a somewhat related criterion that will crop up occasionally as well, and it will be good for us to say something in detail about it right away. This is a **criterion for membership in a kind**. A "kind" is just a grouping of items that all have something in common. So staplers, lamps, tigers, and pearls are all *kinds*. A criterion for membership in a kind, then, is going to tell us what makes X a member of a kind Y, all of whose members share some identifying feature.

Now there are two relevant sorts of question we might ask, both of which seem related to identity. On the one hand, we might see some object and ask, "What kind of thing is that?" Or we may ask a more specific version of this question: "I know that's an animal (i.e., it's a member of the kind "animal"), but what kind of *animal* is it?" Or even more specifically, "I know that's an animal and that it's a squirrel, but what kind of *squirrel* is it?" On the other hand, we may ask "Is that the same thing/animal/squirrel we saw here the other day?" The first sort of question is about the kind-identity of some object, whereas the second is a question about the numerical identity across time of some object. Each of these questions relies on a different criterion to answer it: a criterion for kind membership will tell us what it takes for some object to have an identity as a member of a particular kind, and a criterion of numerical identity across time will tell us what it takes for some object to be one and the same thing at different times. On its face, then, it would seem that these are quite distinct sorts of criteria.

Consider an example. Suppose a rich person built an elaborate building in 1850 that served as his family homestead until he and his family all died off, at which point it was bequeathed to a neighbor who in 1900 turned it into a church. Fifty years later it was sold to a historical society, and they transformed the building into a museum. Fifty years later in 2000 they sold it to a private group that turned the building into a nightclub, which it remains to this day. Suppose you visited that building in 1901, remembering it having been the old homestead of the rich guy's family, and you ask, "What makes this building a church?" You're asking a question about kind-membership, about what makes this object (the building) a member of the kind "church." And fifty years later you could well ask what now makes the building a museum, and fifty years later (my, how you've grown old), you could ask what makes it a nightclub. In each case, the answer will have something to do with the intentions of the owners and attendees, as well as the functions served by the various parts of the building with respect to the intended end.

The question of kind-membership identity will call for a different sort of answer than the question of numerical identity across time. Suppose that in 1949 you returned to the site and asked, "Is this the same church as the one I saw in 1901?" What makes a building the same church across time may be quite different from what makes it a church to begin with. To be the same church may involve preservation of the same basic physical structure, or at least some kind of continuity of form of that structure, across time, whereas being a church itself may simply be a product of the intentions of some owner(s) in determining the function of some physical space. And so it may go with respect to identity and persons: the question of what makes me the same person across time asks for something different than the question of what makes me a person to begin with.

Nevertheless, the questions asked (and the criteria produced) are related, and some authors we explore will attempt to exploit their relation in the following way. There may be good reason to think that what makes for membership in a kind will actually be an essential ingredient in a criterion for identity across time of its members. So to be a member of the kind *church*, a building will have to have several parts that function in

“church-like” ways per the intention of the owner. So it will have to have a gathering area, and pew-like seats, and a stage, and perhaps a pulpit-like area, and so forth. But if these are the properties that *make it* a church, it might be thought that those properties will be necessary to its ongoing preservation, so that it cannot lose them without also losing its identity. This suggests that kind membership provides a crucial condition for identity across time: what makes X a member of kind Y is also what must be preserved in order for X to be the *same* Y across time. So it may be with us: if what makes me a member of the kind “person” are certain psychological capacities, then those capacities must be preserved in order for me to be the *same* person across time.

Nevertheless, as we will see in Chapter Two, there is a powerful argument available that shows that this truth applies only to kinds whose members are *essentially* members, things that simply wouldn’t exist at all if they weren’t members of that particular kind. So while the building is a member of the kind *church* once the owner cleared out the congregation space, built the pulpit, and so forth, the building itself could still exist—and did, in our example—even after it ceased to be a church: the building is not *essentially* a church. In other words, the building exists before, during, and after the existence of the church, and so the criterion for membership in the kind *church* does not in fact give us the identity conditions for the *building*. Those identity conditions will instead come from the criterion for membership in the kind *building*, which will likely have something to do with a thing’s having an intentionally-shaped physical structure.

How, then, does all of this relate to the question of *our* identity? Some authors think that being a person is like being a church, where what I am is not essentially a person but something much more fundamental, something that exists before, during, and after its incarnation as a person. What is this thing? A particular sort of animal, a biological organism. But if this is what I am essentially, then my identity conditions must have their source in that *biological* nature, which means that what preserves my identity across time may have nothing whatsoever to do with what makes me a person, with my psychology. This also suggests that *my* identity, the identity of the individual that I am, may not necessarily be a *personal*

identity after all. That is, for these authors, the issue of “personal identity” has been mislabeled: instead of being about whether or not a person at one time is the same person as a person at another time, it is really about whether or not something that is a person at one time is identical to *something*—person or not—at another time. Realizing this distinction, they say, enables us to resolve all sorts of problems.

More on this argument in Chapter Two. What matters for now is just that, while there is an important distinction between one’s numerical identity across time and one’s identity as a member of a kind, these two sorts of identity are nevertheless related in very interesting and subtle ways, and it will be important for us to keep this in mind as we go.

One final distinction before we conclude. There is a crucial philosophical distinction that will crop up repeatedly as we go, a distinction between *necessary* and *sufficient* conditions. Here is the gloss on each:

X is a **necessary condition** of Y just in case there could be no Y without X.

X is a **sufficient condition** of Y just in case X in and of itself guarantees Y.

X is a **necessary and sufficient condition** of Y just in case if there’s X there’s Y, and if there’s Y there’s X.

The easiest way to explain this distinction is by means of example. Start with necessary conditions. It is a necessary condition of a person’s being the President of the United States that he or she be born in the United States. One can’t be President of the U.S. without being born in the U.S. Nevertheless, this isn’t a sufficient condition of being President, for there are millions and millions of U.S.-born citizens who aren’t President: being born in the U.S. isn’t enough to guarantee that one becomes President; far from it. When X is a necessary condition for Y, we say, “Y, only if X.”

Now consider sufficient conditions. Suppose you are six feet tall. Being 6’2” is a sufficient condition for being taller than you, then. And so is being 6’5” or being 6’1/2”. Are any of these *necessary* conditions for being taller than you? No. It’s not as if I wouldn’t count as taller than you if I were 6’8”,

say. When X is a sufficient condition for Y, we say, “If X, then Y.”

Finally, consider necessary and sufficient conditions. Together these provide a guarantee of the presence of some thing that also couldn’t be present without the conditions in question. So the necessary and sufficient conditions of someone’s being a lesbian, say, are that (a) that person is a woman, and (b) that person has a predominant sexual interest in women. Note that neither of these conditions alone is sufficient (although they’re each necessary): just being a woman isn’t enough to guarantee you’re a lesbian, and neither does just having a sexual interest in women (you could be a heterosexual male, for instance). But whenever someone meets these conditions she will be a lesbian, and wherever there’s a lesbian she will meet these two conditions. When X is a necessary and sufficient condition for Y, we say, “Y if and only if X.”

Given all of these important distinctions, then, the first question we will pursue may be put as follows: *What makes person X at some time (t_1) quantitatively identical to person Y at a later time (t_2)?* This is a question about the nature of the quantitative identity relation, not the membership in a class relation. It also seems to be demanding both necessary and sufficient conditions, although both conditions may not turn out to be essential for what we need in some cases. And of course this general formulation itself may not survive throughout our entire inquiry—there are many who think it’s just the wrong formulation of the identity question, as we will see in both Chapters Two and Three—but for now it gives us enough, namely, a clear entrée into the puzzle of personal identity. And that is the topic to which we can now, finally, turn.