

## WHAT WE BELIEVE OURSELVES TO BE

I enter the Teletransporter. I have been to Mars before, but only by the old method, a space-ship journey taking several weeks. This machine will send me at the speed of light. I merely have to press the green button. Like others, I am nervous. Will it work? I remind myself what I have been told to expect. When I press the button, I shall lose consciousness, and then wake up at what seems a moment later. In fact I shall have been unconscious for about an hour. The Scanner here on Earth will destroy my brain and body, while recording the exact states of all of my cells. It will then transmit this information by radio. Travelling at the speed of light, the message will take three minutes to reach the Replicator on Mars. This will then create, out of new matter, a brain and body exactly like mine. It will be in this body that I shall wake up.

Though I believe that this is what will happen, I still hesitate. But then I remember seeing my wife grin when, at breakfast today, I revealed my nervousness. As she reminded me, she has been often teletransported, and there is nothing wrong with *her*. I press the button. As predicted, I lose and seem at once to regain consciousness, but in a different cubicle. Examining my new body, I find no change at all. Even the cut on my upper lip, from this morning's shave, is still there.

Several years pass, during which I am often Teletransported. I am now back in the cubicle, ready for another trip to Mars. But this time, when I press the green button, I do not lose consciousness. There is a whirring sound, then silence. I leave the cubicle, and say to the attendant: 'It's not working. What did I do wrong?'

'It's working', he replies, handing me a printed card. This reads: 'The New Scanner records your blueprint without destroying your brain and body. We hope that you will welcome the opportunities which this technical advance offers.'

The attendant tells me that I am one of the first people to use the New Scanner. He adds that, if I stay for an hour, I can use the Intercom to see and talk to myself on Mars.

'Wait a minute', I reply, 'If I'm here I can't *also* be on Mars'.

Someone politely coughs, a white-coated man who asks to speak to

me in private. We go to his office, where he tells me to sit down, and pauses. Then he says: 'I'm afraid that we're having problems with the New Scanner. It records your blueprint just as accurately, as you will see when you talk to yourself on Mars. But it seems to be damaging the cardiac systems which it scans. Judging from the results so far, though you will be quite healthy on Mars, here on Earth you must expect cardiac failure within the next few days.'

The attendant later calls me to the Intercom. On the screen I see myself just as I do in the mirror every morning. But there are two differences. On the screen I am not left-right reversed. And, while I stand here speechless, I can see and hear myself, in the studio on Mars, starting to speak.

What can we learn from this imaginary story? Some believe that we can learn little. This would have been Wittgenstein's view.<sup>1</sup> And 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 words have some logical force beyond what our past needs have invested them with.'<sup>2</sup>

This criticism might be justified if, when considering such imagined cases, we had no reactions. But these cases arouse in most of us strong beliefs. And these are beliefs, not about our words, but about ourselves. By considering these cases, we discover what we believe to be involved in our own continued existence, or what it is that makes us now and ourselves next year the same people. We discover our beliefs about the nature of personal identity over time. Though our beliefs are revealed most clearly when we consider imaginary cases, these beliefs also cover actual cases, and our own lives. In Part Three of this book I shall argue that some of these beliefs are false, then suggest how and why this matters.

#### 75. SIMPLE TELETRANSPORTATION AND THE BRANCH-LINE CASE

At the beginning of my story, the Scanner destroys my brain and body. My blueprint is beamed to Mars, where another machine makes an organic *Replica* of me. My Replica thinks that he is me, and he seems to remember living my life up to the moment when I pressed the green button. In every other way, both physically and psychologically, my Replica is just like me. If he returned to Earth, everyone would think that he was me.

Simple Teletransportation, as just described, is a common feature in science fiction. And it is believed, by some readers of this fiction, merely to be the fastest way of travelling. They believe that my Replica *would be me*. Other science fiction readers, and some of the characters in this fiction, take a different view. They believe that, when I press the green button, I die. My

Replica is *someone else*, who has been made to be exactly like me.

This second view seems to be supported by the end of my story. The New Scanner does not destroy my brain and body. Besides gathering the information, it merely damages my heart. While I am in the cubicle, with the green button pressed, nothing seems to happen. I walk out, and learn that in a few days I shall die. I later talk, by two-way television, to my Replica on Mars. Let us continue the story. Since my Replica knows that I am about to die, he tries to console me with the same thoughts with which I recently tried to console a dying friend. It is sad to learn, on the receiving end, how unconsoling these thoughts are. My Replica then assures me that he will take up my life where I leave off. He loves my wife, and together they will care for my children. And he will finish the book that I am writing. Besides having all of my drafts, he has all of my intentions. I must admit that he can finish my book as well as I could. All these facts console me a little. Dying when I know that I shall have a Replica is not quite as bad as, simply, dying. Even so, I shall soon lose consciousness, forever.

In Simple Teletransportation, I do not co-exist with my Replica. This makes it easier to believe that this *is* a way of travelling—that my Replica *is* me. At the end of my story, my life and that of my Replica overlap. Call this the *Branch-Line Case*. In this case, I cannot hope to travel on the *Main Line*, waking up on Mars with forty years of life ahead. I shall remain on the Branch-Line, on Earth, which ends a few days later. Since I can talk to my Replica, it seems clear that he is *not* me. Though he is exactly like me, he is one person, and I am another. When I pinch myself, he feels nothing. When I have my heart attack, he will again feel nothing. And when I am dead he will live for another forty years.

If we believe that my Replica is not me, it is natural to assume that my prospect, on the Branch Line, is almost as bad as ordinary death. I shall deny this assumption. As I shall argue later, I ought to regard having a Replica as being about as good as ordinary survival. I can best defend this claim, and the view that supports it, after briefly discussing part of the past debate about personal identity.

#### 76. QUALITATIVE AND NUMERICAL IDENTITY

There are two kinds of sameness, or identity. I and my Replica are *qualitatively identical*, or exactly alike. But we may not be *numerically identical*, or one and the same person. Similarly, two white billiard balls are not numerically but may be qualitatively identical. If I paint one of these balls red, it will not now be qualitatively identical to itself yesterday. But the red ball that I see now and the white ball that I painted red are numerically identical. They are one and the same ball.

We might say, of someone, 'After his accident, he is no longer the same person'. This is a claim about both kinds of identity. We claim that *he*, the same person, is *not* now the same person. This is not a contradiction. We

merely mean that this person's character has changed. This numerically identical person is now qualitatively different.

When we are concerned about our future, it is our numerical identity that we are concerned about. I may believe that, after my marriage, I shall not be the same person. But this does not make marriage death. However much I change, I shall still be alive if there will be some person living who is numerically identical with me.

The philosophical debate is about the nature both of persons and of personal identity over time. It will help to distinguish these questions:

- (1) What is the nature of a person?
- (2) What is it that makes a person at two different times one and the same person?
- (3) What is necessarily involved in the continued existence of each person over time?

The answer to (2) can take this form: 'X today is one and the same person as Y at some past time *if and only if*. . .' This answer states the *necessary and sufficient conditions* for personal identity over time. And the answer to (2) provides the answer to (3). Each person's continued existence has the *same* necessary and sufficient conditions.

In answering (2) and (3) we shall also partly answer (1). The necessary features of our continued existence depend upon our nature. And the simplest answer to (1) is that, to be a person, a being must be self-conscious, aware of its identity and its continued existence over time.

We can also ask

- (4) What is in fact involved in the continued existence of each person over time?

Since our continued existence has features that are not necessary, the answer to (3) is only part of the answer to (4). Being happy, for our example, is not necessary to our continued existence, but it may be part of what someone's continued existence in fact involves.

Though question (2) is about numerical rather than qualitative identity, this does not imply that qualitative changes do not matter. On one view, certain kinds of qualitative change destroy numerical identity. If certain things happened to me, the truth may not be that I become a very different person. The truth may be that I cease to exist, and the resulting person is someone else.

#### 77. THE PHYSICAL CRITERION OF PERSONAL IDENTITY

Many writers use the ambiguous phrase 'the criterion of identity over time'. Some mean by this 'our way of telling whether some present object is

identical with some past object'. But I shall mean *what this identity necessarily involves, or consists in*.

In the case of most physical objects, on what I call the *standard view*, the criterion of identity over time is the spatio-temporal physical continuity of this object. This is something that we all understand, even if we fail to understand the description I shall now give. In the simplest case of physical continuity, like that of the Pyramids, an apparently static object continues to exist. In another simple case, like that of the Moon, an object moves in a regular way. Many objects move in less regular ways, but they still trace physically continuous spatio-temporal paths. Suppose that the billiard ball that I painted red is the same as the white ball with which last year I made a winning shot. On the standard view, this is true only if this ball traced such a continuous path. It must be true (1) that there is a line through space and time, starting where the white ball rested before I made my winning shot, and ending where the red ball now is, (2) that at every point on this line there was a billiard ball, and (3) that the existence of a ball at each point on this line was in part caused by the existence of a ball at the immediately preceding point.<sup>3</sup>

Some kinds of thing continue to exist even though their physical continuity involves great changes. A Camberwell Beauty is first an egg, then a caterpillar, then a chrysalis, then a butterfly. These are four stages in the physically continuous existence of a single organism. Other kinds of thing cannot survive such great changes. Suppose that an artist paints a self-portrait and then, by repainting, turns this into a portrait of his father. Even though these portraits are more similar than a caterpillar and a butterfly, they are not stages in the continued existence of a single painting. The self-portrait is a painting that the artist destroyed. In a general discussion of identity, we would need to explain why the requirement of physical continuity differs in such ways for different kinds of thing. But we can ignore this here.

Can there be gaps in the continued existence of a physical object? Suppose that I have the same gold watch that I was given as a boy even though, for a month, it lay disassembled on a watch-repairer's shelf. On one view, in the spatio-temporal path traced by this watch there was not at every point a watch, so my watch does not have a history of full physical continuity. But during the month when my watch was disassembled, and did not exist, all of its parts had histories of full continuity. On another view, even when it was disassembled, my watch existed.

Another complication again concerns the relation between a complex thing and the various parts of which it is composed. It is true of some of these things, though not true of all, that their continued existence need not involve the continued existence of their components. Suppose that a wooden ship is repaired from time to time while it is floating in harbour, and that after fifty years it contains none of the bits of wood out of which it was first built. It is still one and the same ship, because, as a ship, it has

displayed throughout these fifty years full physical continuity. This is so despite the fact that it is now composed of quite different bits of wood. These bits of wood might be qualitatively identical to the original bits, but they are not one and the same bits. Something similar is partly true of a human body. With the exception of some brain cells, the cells in our bodies are replaced with new cells several times in our lives.

I have now described the physical continuity which, on the standard view, makes a physical object one and the same after many days or years. This enables me to state one of the rival views about personal identity. On this view, what makes me the same person over time is that I have the same brain and body. The criterion of my identity over time—or what this identity involves—is the physical continuity, over time, of my brain and body. I shall continue to exist if and only if this particular brain and body continue both to exist and to be the brain and body of a living person.

This is the simplest version of this view. There is a better version. This is

*The Physical Criterion:* (1) What is necessary is not the continued existence of the whole body, but the continued existence of *enough* of the brain to be the brain of a living person. *X* today is one and the same person as *Y* at some past time if and only if (2) enough of *Y*'s brain continued to exist, and is now *X*'s brain, and (3) there does not exist a different person who also has enough of *Y*'s brain. (4) Personal identity over time just consists in the holding of facts like (2) and (3).

(1) is clearly needed in certain actual cases. Some people continue to exist even though they lose much of their bodies, perhaps including their hearts and lungs if they are on Heart-Lung Machines. The need for (3) will be clear later.

Those who believe in the Physical Criterion would reject Teletransportation. They would believe this to be a way, not of travelling, but of dying. They would also reject, as inconceivable, reincarnation. They believe that someone cannot have a life after death, unless he lives this life in a resurrection of the very same, physically continuous body. Some of the Christians who believe this insist that they be buried. They believe that if, like Greek and Trojan heroes, they were burnt on funeral pyres, and their ashes scattered, not even God could bring them to life again. God could create only a Replica, someone else who was exactly like them. Other Christians believe that God could resurrect *them* if He reassembled their bodies out of the bits of matter that, when they were last alive, made up their bodies. This view is like the first of the views about the reassembly of my gold watch.<sup>4</sup>

#### 78. THE PSYCHOLOGICAL CRITERION

Some people believe in a kind of psychological continuity that resembles physical continuity. This involves the continued existence of a purely mental

entity, or thing—a soul, or spiritual substance. I shall return to this view. But I shall first explain another kind of psychological continuity. This is less like physical continuity, since it does not consist in the continued existence of some entity. But this other kind of psychological continuity involves only facts with which we are familiar.

What has been most discussed is the continuity of memory. This is because it is memory that makes most of us aware of our own continued existence over time. The exceptions are the people who are suffering from amnesia. Most amnesiacs lose only two sets of memories. They lose all of their memories of having particular past experiences—or, for short, their *experience memories*. They also lose some of their memories about facts, those that are about their own past lives. But they remember other facts, and they remember how to do different things, such as how to speak, or swim.

Locke suggested that experience-memory provides the criterion of personal identity.<sup>5</sup> Though this is not, on its own, a plausible view, I believe that it can be part of such a view. I shall therefore try to answer Locke's critics.

Locke claimed that someone cannot have committed some crime unless he now remembers doing so. We can understand a reluctance to punish people for crimes that they cannot remember. But, taken as a view about what is involved in a person's continued existence, Locke's claim is clearly false. If it was true, it would not be possible for someone to forget any of the things that he once did, or any of the experiences that he once had. But this is possible. I cannot now remember putting on my shirt this morning.

There are several ways to extend the experience-memory criterion so as to cover such cases. I shall appeal to the concept of an overlapping chain of experience-memories. Let us say that, between X today and Y twenty years ago, there are *direct memory connections* if X can now remember having some of the experiences that Y had twenty years ago. On Locke's view, this makes X and Y one and the same person. Even if there are *no* such direct memory connections, there may be *continuity of memory* between X now and Y twenty years ago. This would be so if between X now and Y at that time there has been an overlapping chain of direct memories. In the case of most people who are over twenty three, there would be such an overlapping chain. In each day within the last twenty years, most of these people remembered some of their experiences on the previous day. On the revised version of Locke's view, some present person X is the same as some past person Y if there is between them continuity of memory.

This revision meets one objection to Locke's view. We should also revise the view so that it appeals to other facts. Besides direct memories, there are several other kinds of direct psychological connection. One such connection is that which holds between an intention and the later act in which this intention is carried out. Other such direct connections are those which hold when a belief, or a desire, or any other psychological feature, continues to

be had.

I can now define two general relations:

*Psychological connectedness* is the holding of particular direct psychological connections.

*Psychological continuity* is the holding of overlapping chains of *strong* connectedness.

Of these two general relations, connectedness is more important both in theory and in practice. Connectedness can hold to any degree. Between X today and Y yesterday there might be several thousand direct psychological connections, or only a single connection. If there was only a single connection, X and Y would not be, on the revised Lockean View, the same person. For X and Y to be the same person, there must be over every day *enough* direct psychological connections. Since connectedness is a matter of degree, we cannot plausibly define precisely what counts as enough. But we can claim that there is enough connectedness if the number of connections, over any day, is *at least half* the number of direct connections that hold, over every day, in the lives of nearly every actual person.<sup>6</sup> When there are enough direct connections, there is what I call *strong* connectedness.

This relation cannot be the criterion of personal identity. A relation *F* is *transitive* if it is true that, if X is F-related to Y, and Y is F-related to Z, X and Z *must* be F-related. Personal identity is a transitive relation. If Bertie was one and the same person as the philosopher Russell, and Russell was one and the same person as the author of *Why I Am Not a Christian*, this author and Bertie must be one and the same person.

Strong connectedness is *not* a transitive relation. I am now strongly connected to myself yesterday, when I was strongly connected to myself two days ago, when I was strongly connected to myself three days ago, and so on. It does not follow that I am now strongly connected to myself twenty years ago. And this is not true. Between me now and myself twenty years ago there are many fewer than the number of direct psychological connections that hold over any day in the lives of nearly all adults. For example, while these adults have many memories of experiences that they had in the previous day, I have few memories of experiences that I had twenty years ago.

By 'the criterion of personal identity over time' I mean what this identity *necessarily involves or consists in*. Because identity is a transitive relation, the criterion of identity must be a transitive relation. Since strong connectedness is not transitive, it cannot be the criterion of identity. And I have just described a case in which this is shown. I am the same person as myself twenty years ago, though I am not now strongly connected to myself then.

Though a defender of Locke's view cannot appeal to psychological

transitive. He can appeal to

*The Psychological Criterion:* (1) There is *psychological continuity* if and only if there are overlapping chains of strong connectedness. X today is one and the same person as Y at some past time if and only if (2) X is psychologically continuous with Y, (3) this continuity has the right kind of cause, and (4) there does not exist a different person who is also psychologically continuous with Y. (5) Personal identity over time just consists in the holding of facts like (2) to (4).

As with the Physical Criterion, the need for (4) will be clear later.

There are three versions of the Psychological Criterion. These differ over the question of what is the *right* kind of cause. On the *Narrow* version, this must be the *normal* cause. On the *Wide* version, this could be *any reliable* cause. On the *Widest* version, the cause could be *any* cause.

The Narrow Psychological Criterion uses words in their ordinary sense. Thus I remember having an experience only if

- (1) I seem to remember having an experience,
- (2) I did have this experience,

and

- (3) my apparent memory is causally dependent, in the normal way, on this past experience.

That we need condition (3) can be suggested with an example. Suppose that I am knocked unconscious in a climbing accident. After I recover, my fellow-climber tells me what he shouted just before I fell. In some later year, when my memories are less clear, I might seem to remember the experience of hearing my companion shout just before I fell. And it might be true that I did have just such an experience. But though conditions (1) and (2) are met, we should not believe that I am remembering that past experience. It is a well-established fact that people can never remember their last few experiences before they were knocked unconscious. We should therefore claim that my apparent memory of hearing my companion shout is not a real memory of that past experience. This apparent memory is not causally dependent in the right way on that past experience. I have this apparent memory only because my companion later told me what he shouted.<sup>7</sup>

Similar remarks apply to the other kinds of continuity, such as continuity of character. On the Narrow Psychological Criterion, even if someone's character radically changes, there is continuity of character if these changes have one of several normal causes. Some changes of character are deliberately brought about; others are the natural consequence of growing older; others are the natural response to certain kinds of experience. But there would not be continuity of character if radical and unwanted changes

Though it is memory that makes us aware of our own continued existence over time, the various other continuities have great importance. We may believe that they have enough importance to provide personal identity even in the absence of memory. We shall then claim, what Locke denied, that a person continues to exist even if he suffers from complete amnesia. I would rather suffer amnesia than have surgery that would give me a quite different and obnoxious character.

Besides the Narrow version, I described the two Wide versions of the Psychological Criterion. These versions extend the senses of several words. On the ordinary sense of 'memory', a memory must have its normal cause. The two Wide Psychological Criteria appeal to a wider sense of 'memory', which allows either any reliable cause, or any cause. Similar claims apply to the other kinds of direct psychological connection. To simplify my discussion of these three Criteria, I shall use 'psychological continuity' in its widest sense, that allows this continuity to have *any* cause.

If we appeal to the Narrow Version, which insists on the normal cause, the Psychological Criterion coincides in most cases with the Physical Criterion. The normal causes of memory involve the continued existence of the brain. And some or all of our psychological features depend upon states or events in our brains. The continued existence of a person's brain is at least part of the normal cause of psychological continuity. On the Physical Criterion, a person continues to exist if and only if (a) there continues to exist *enough* of this person's brain so that it remains the brain of a living person, and (b) no different person ever has enough of this person's brain. (a) and (b) are claimed to be the necessary and sufficient conditions for this person's identity, or continued existence, over time. On the Narrow Psychological Criterion, (a) is necessary, but not sufficient. A person continues to exist if and only if (c) there is psychological continuity, (d) this continuity has its normal cause, and (e) there does not exist a different person who is also psychologically continuous with this person. (a) is required as part of the normal cause of psychological continuity.

I shall argue that the two Wide Psychological Criteria are both better than the Narrow Criterion. A partial analogy may suggest why. Some people go blind because of damage to their eyes. Scientists are now developing artificial eyes. These involve a glass or plastic lens, and a micro-computer which sends through the optic nerve electrical patterns like those that are sent through this nerve by a natural eye. When such artificial eyes are more advanced, they might give to someone who has gone blind visual experiences just like those that he used to have. What he seems to see would correspond to what is in fact before him. And his visual experiences would be causally dependent, in this new but reliable way, on the light-waves coming from the objects that are before him.

Would this person be *seeing* these objects? If we insist that seeing must involve the normal cause, we would answer No. But even if this person cannot see, what he has is *just as good as* seeing, both as a way of knowing

what is within sight, and as a source of visual pleasure. If we accept the Psychological Criterion, we could make a similar claim. If psychological continuity does not have its normal cause, some may claim that it is *not* true psychological continuity. We can claim that, even if this is so, this kind of continuity is *just as good as* ordinary continuity.

Reconsider the start of my imagined story, where my brain and body are destroyed. The Scanner and the Replicator produce a person who has a new but exactly similar brain and body, and who is psychologically continuous with me as I was when I pressed the green button. The cause of this continuity is, though unusual, reliable. On both the Physical Criterion and the Narrow Psychological Criterion, my Replica would *not* be me. On the two Wide Criteria, he *would* be me.

#### 79. THE OTHER VIEWS

I am asking what is the criterion of personal identity over time—what this identity involves, or consists in. I first described the spatio-temporal physical continuity that, on the standard view, is the criterion of identity of physical objects. I then described two views about personal identity, the Physical and Psychological Criteria.

There is a natural but false assumption about these views. Many people believe in what is called *Materialism*, or *Physicalism*. This is the view that that there are no purely mental objects, states, or events. On one version of Physicalism, every mental event is just a physical event in some particular brain and nervous system. There are other versions. Those who are not Physicalists are either *Dualists* or *Idealists*. Dualists believe that mental events are *not* physical events. This can be so even if all mental events are causally dependent on physical events in a brain. Idealists believe that all states and events are, when understood correctly, purely mental. Given these distinctions, we may assume that Physicalists must accept the Physical Criterion of personal identity.

This is not so. Physicalists could accept the Psychological Criterion. And they could accept the version that allows any reliable cause, or any cause. They could thus believe that, in Simple Teletransportation, my Replica would be me. They would here be rejecting the Physical Criterion.<sup>8</sup>

These criteria are not the only views about personal identity. I shall now describe the other views that are either sufficiently plausible, or have enough supporters, to be worth considering. This description will be too abstract to be fully understood, before the details are filled out in later chapters. But it is worth giving this description, both for later reference, and to provide a rough idea of what lies ahead. If much of this summary seems, on a first reading, either obscure or trivial, *do not worry*.

I start with a new distinction. On the Physical Criterion, personal identity

over time just involves the physically continuous existence of enough of a brain so that it remains the brain of a living person. On the Psychological Criterion, personal identity over time just involves the various kinds of psychological continuity, with the right kind of cause. These views are both *Reductionist*. They are Reductionist because they claim

- (1) that the fact of a person's identity over time just consists in the holding of certain more particular facts,
- and
- (2) that these facts can be described without either presupposing the identity of this person, or explicitly claiming that the experiences in this person's life are had by this person, or even explicitly claiming that this person exists. These facts can be described in an *impersonal* way.

It may seem that (2) could not be true. When we describe the psychological continuity that unifies some person's mental life, we must mention this person, and many other people, in describing the *content* of many thoughts, desires, intentions, and other mental states. But mentioning this person in this way does not involve either asserting that these mental states are had by this person, or asserting that this person exists. A similar claim applies to the Physical Criterion. These claims need further arguments, which I shall later give.

Our view is *Non-Reductionist* if we reject either or both of the two Reductionist claims.

Many Non-Reductionists hold what I call the view that *we are separately existing entities*. On this view, personal identity over time does not just consist in physical and/or psychological continuity. It is a separate, further fact. A person is a separately existing entity, distinct from his brain and body, and his experiences. On the best-known version of this view, a person is a *purely mental* entity: a Cartesian Pure Ego, or spiritual substance. But we might believe that a person is a separately existing *physical* entity, of a kind that is not yet recognised in the theories of contemporary physics.

There is another Non-Reductionist View. This view denies that we are separately existing entities, distinct from our brains and bodies, and our experiences. But this view claims that, though we are not separately existing entities, personal identity *is* a further fact, which does not just consist in physical and/or psychological continuity. I call this the *Further Fact View*.

The Physical and Psychological Criteria are versions of the Reductionist View. And there are two versions of each Criterion. As I have said, what is *necessarily* involved in a person's continued existence is less than what is *in fact* involved. Believers in the Reductionist Criteria disagree when considering imaginary cases. But they would agree about what is in fact

involved in the existence of actual people. They would start to disagree only if, for example, people began to be Teletransported.

On the Reductionist View, each person's existence just involves the existence of a brain and body, the doing of certain deeds, the thinking of certain thoughts, the occurrence of certain experiences, and so on. It will help to extend the ordinary sense of the word 'event'. I shall use 'event' to cover even such *boring* events as the continued existence of a belief, or a desire. This use makes the Reductionist View simpler to describe. And it avoids what a Reductionist believes to be the misleading implications of the words 'mental state'. While a state must be a state *of* some entity, this is not true of an event. Given this extended use of the word 'event', all Reductionists would accept

- (3) A person's existence just consists in the existence of a brain and body, and the occurrence of a series of interrelated physical and mental events.

Some Reductionists claim

- (4) A person *just is* a particular brain and body, and such a series of interrelated events.

(4) uses the *is of composition*, as in the claim that a statue is a piece of bronze. This is not the *is of identity*. A statue and a piece of bronze are not one and the same thing. This is shown by the fact that, if we melt the statue, we destroy the statue but do not destroy the piece of bronze. Such a statue is composed of a piece of bronze. In the same sense, (4) claims that a person is composed of a particular brain and body, and a series of interrelated physical and mental events.

Other Reductionists claim

- (5) A person is an entity that is *distinct* from a brain and body, and such a series of events.

On this version of the Reductionist View, a person is not merely a composite object, with these various components. A person is an entity that *has* a brain and body, and *has* particular thoughts, desires, and so on. But, though (5) is true, a person is not a *separately existing* entity. Though (5) is true, (3) is also true.

This version of Reductionism may seem self-contradictory. (3) and (5) may seem to be inconsistent. It may help to consider Hume's analogy: 'I cannot compare the soul more properly to anything than to a republic, or commonwealth.'<sup>9</sup> Most of us are Reductionists about nations. We would accept the following claims: Nations exist. Ruritania does not exist, but France does. Though nations exist, a nation is not an entity that exists separately, apart from its citizens and its territory. We would accept

- (6) A nation's existence just involves the existence of its citizens,

living together in certain ways, on its territory.

Some claim

- (7) A nation just *is* these citizens and this territory.

Others claim

- (8) A nation is an entity that is distinct from its citizens and its territory.

We may believe that (8) and (6) are not inconsistent. If we believe this, we may accept that there is no inconsistency between the corresponding claims (3) and (5). We may thus agree that the version of Reductionism expressed in (3) and (5) is a consistent view. If this version is consistent, as I believe, it is the better version. It uses our actual concept of a person. In most of what follows, we can ignore the difference between these two versions. But at one point this difference may have great importance.<sup>10</sup>

Besides claiming (1) and (2), Reductionists would also claim

- (9) Though persons exist, we could give a *complete* description of reality *without* claiming that persons exist.

I call this the view *that a complete description could be impersonal*.

This view may also seem to be self-contradictory. If persons exist, and a description of what exists fails to mention persons, how can this description be complete?

A Reductionist could give the following reply. Suppose that an object has two names. This is true of the planet that is called both *Venus* and the *Evening Star*. In our description of what exists, we could claim that *Venus* exists. Our description could then be complete even though we do not claim that the *Evening Star* exists. We need not make this claim because, using its other name, we have already claimed that this object exists.

A similar claim applies when some fact can be described in two ways. Some Reductionists accept (4), the claim that a person just is a particular brain and body, and a series of interrelated physical and mental events. If this is what a person *is*, the same fact can be described in two ways. We can claim either

- (10) that there exists a particular brain and body, and a particular series of interrelated physical and mental events.

or

- (11) that a particular person exists,

If (10) and (11) are two ways of describing the *same* fact, a complete description need not make *both* claims. In a complete description we could claim (10), and fail to claim (11). Though this person exists, a complete

description need not claim that he exists. The fact that he exists has already been reported in claim (10).

Other Reductionists accept (5), the claim that a person is distinct from his brain and body, and his acts, thoughts, and other physical and mental events. On this version of Reductionism, claim (10) does not describe the very same fact that claim (11) describes. But claim (10) *implies* or *entails* claim (11). More weakly, given our understanding of the concept of a person, if we know that (10) is true, we shall know that (11) is true. These Reductionists can say that, if our description of reality either states or implies or would enable us to know the existence of everything that exists, our description is complete. This claim is not as clearly true as the claim that a complete description need not give two descriptions of the same fact. But this claim seems plausible. If it is justified, and the Reductionist View is true, these Reductionists can completely describe reality without claiming that persons exist.<sup>11</sup>

My claims about Reductionism draw distinctions that, in this abstract form, are hard to grasp. Fortunately, there are ways of making it clear whether we accept or reject a Reductionist View about some kind of thing. If we accept a Reductionist View, there may be cases where we believe the identity of such a thing to be, in a quite unpuzzling way, *indeterminate*. We would *not* believe this if we reject the Reductionist View about this kind of thing.

Consider, for example, clubs. Suppose that a certain club exists for several years, holding regular meetings. The meetings then cease. Some years later, some of the members of this club form a club with the same name, and the same rules. We ask: 'Have these people reconvened the *very same* club? Or have they merely started up *another* club, which is exactly similar?' There might be an answer to this question. The original club might have had a rule explaining how, after such a period of non-existence, it could be reconvened. Or it might have had a rule preventing this. But suppose that there is no such rule, and no legal facts, supporting either answer to our question. And suppose that the people involved, if they asked our question, would not give it an answer. There would then be no answer to our question. The claim 'This is the same club' would be *neither true nor false*.

Though there is no answer to our question, there may be nothing that we do not know. This is because the existence of a club is not separate from the existence of its members, behaving in certain ways. The continued existence of a club just involves its members having meetings, that are conducted according to the club's rules. If we know all the facts about how people held meetings, and about the club's rules, we know everything there is to know. This is why we would not be puzzled when we cannot answer the question, 'Is this the very same club?' We would not be puzzled because, without answering this question, we can know everything about what happened. When this is true of some question, I call this question *empty*.

When we ask an empty question, there is only one fact or outcome that we are considering. Different answers to our question are merely different descriptions of this fact or outcome. This is why, without answering this empty question, we can know everything that there is to know. In my example we can ask, 'Is this the very same club, or is it merely another club, that is exactly similar?' But these are not here two different possibilities, one of which must be true.

When an empty question has no answer, we can decide to *give* it an answer. We could decide to call the later club the same as the original club. Or we could decide to call it another club, that is exactly similar. This is not a decision between different views about what really happened. Before making our decision, we already knew what happened. We are merely choosing one of two different descriptions of the very same course of events.

If we are Reductionists about personal identity, we should make similar claims. Suppose that we accept the Psychological Criterion. We can describe cases where the psychological connectedness between me now and some future person will hold only to a reduced degree. If I imagine myself in such a case, I can always ask, 'Am I about to die? Will the resulting person be me?' On this version of the Reductionist View, in some cases there would be no answer to my question. My question would be *empty*. The claim that I am about to die would be neither true nor false. If I know the facts about both physical continuity and psychological connectedness, I know everything there is to know. I know everything, even though I do not know whether I am about to die, or shall go on living for many years.

When it is applied to ourselves, this Reductionist claim is hard to believe. In such imagined cases, something unusual is about to happen. But most of us are inclined to believe that, in any conceivable case, the question 'Am I about to die?' must have an answer. And we are inclined to believe that this answer must be either, and quite simply, Yes or No. Any future person must be either me, or someone else. These beliefs I call the view that *our identity must be determinate*.

I shall next describe two explanatory claims. The first answers a new question. What unites the different experiences that are had by a single person at the same time? While I type this sentence, I am aware of the movements of my fingers, and can see the sunlight on my desk, and can hear the wind ruffling some leaves. What unites these different experiences? Some claim: the fact that they are all *my* experiences. These are the experiences that are being had, at this time, by a particular person, or *subject of experiences*. A similar question covers my whole life. What unites the different experiences that, together, constitute this life? Some give the same answer. What unites all of these experiences is, simply, that they are all mine. These answers I call the view that *psychological unity is explained by ownership*.

The views described so far are about the nature of personal identity. I shall end with a pair of views that are about, not the nature of this identity, but its importance. Consider an ordinary case where, even on any version of the Reductionist View, there are two possible outcomes. In one of the outcomes, I am about to die. In the other outcome I shall live for another forty years. If these forty years would be worth living, the second outcome would be better for me. And the difference between these outcomes would be judged to be important on most theories about rationality, and most moral theories. It would have rational and moral significance whether I am about to die, or shall live for another forty years. What is judged to be important here is whether, during these forty years, there will be someone living who will *be me*. This is a question about personal identity. On one view, this is always what is important. I call this the view that *personal identity is what matters*. This is the natural view.

The rival view is that *personal identity is not what matters*. I claim

*What matters is Relation R*: psychological connectedness and/or continuity with the right kind of cause.

Since it is more controversial, I add, as a separate claim

The right kind of cause could be any cause.

It is in imaginary cases that we can best decide whether what matters is Relation R or personal identity. One example may be the Branch-Line Case, where my life briefly overlaps with that of my Replica. Suppose that we believe that I and my Replica are two different people. I am about to die, but my Replica will live for another forty years. If personal identity is what matters, I should regard my prospect here as being nearly as bad as ordinary death. But if what matters is Relation R, with any cause, I should regard this way of dying as being about as good as ordinary survival.

The disagreement between these views is not confined to imaginary cases. The two views also disagree about all of the actual lives that are lived. The disagreement is here less sharp, because, on both views, all or nearly all these lives contain the relation that matters. On all of the plausible views about the nature of personal identity, personal identity nearly always coincides with psychological continuity, and roughly coincides with psychological connectedness. But, as I shall argue later, it makes a great difference which of these we believe to be what matters. If we cease to believe that our identity is what matters, this may affect some of our emotions, such as our attitude to ageing and to death. And, as I shall argue, we may be led to change our views about both rationality and morality.

I have now given a first description of several different views. Stated in this abstract way, this description cannot be wholly clear. But what is now obscure may, when I discuss these views, become clear.

How are these views related to each other? I shall claim. what some deny.

that many of these views stand or fall together. If this is so, it will be easier to decide what the truth is. When we see how these views are related, we shall find, I believe, that we have only two alternatives. It is worth stating in advance some of the ways in which, as I shall argue, these views are related.

If we do not believe that we are separately existing entities, can we believe that personal identity is what matters? Some writers think we can. I shall argue that we cannot.

If we do not believe that we are separately existing entities, can we claim that personal identity does not just consist in physical and psychological continuity, but is a further fact? Some writers think we can. I shall argue that we cannot.

If we believe that our identity must be determinate, must we believe that we are separately existing entities? Having the first belief does not imply having the second. We might believe both that we are not separately existing entities, and that, to any question about personal identity, there must always be an answer, which must be either Yes or No. There are some writers who accept this view. But I shall argue that this view is indefensible. Only if we are separately existing entities can it be true that our identity must be determinate.

It would be possible to claim that we are separately existing entities, but deny that our identity must be determinate. But there are few people who would combine these claims.

Suppose next that we believe that psychological unity is explained by ownership. We believe that the unity of a person's consciousness at any time is explained by the fact that this person's different experiences are all being had by this person. And we believe that the unity of a person's whole life is explained by the fact that all of the experiences in this life are had by this person. These are the explanations given by those who claim that we are separately existing entities. Can we give these explanations if we reject that claim? Some writers suggest that we can. But I shall argue that we cannot.

I shall also argue for the following conclusions:

- (1) We are not separately existing entities, apart from our brains and bodies, and various interrelated physical and mental events. Our existence just involves the existence of our brains and bodies, and the doing of our deeds, and the thinking of our thoughts, and the occurrence of certain other physical and mental events. Our identity over time just involves (a) Relation R—psychological connectedness and/or psychological continuity, either with the normal cause or with any cause, provided (b) that there is no different person who is R-related to us as we once were.
- (2) It is not true that our identity is always determinate. I can always ask, 'Am I about to die?' But it is not true that, in every case, this

question must have an answer, which must be either Yes or No. In some cases this would be an empty question.

- (3) There are two unities to be explained: the unity of consciousness at any time, and the unity of a whole life. These two unities cannot be explained by claiming that different experiences are had by the same person. These unities must be explained by describing the relations between these many experiences, and their relations to this person's brain. And we can refer to these experiences, and fully describe the relations between them, without claiming that these experiences are had by a person.
- (4) Personal identity is not what matters. What fundamentally matters is Relation R, with any cause. This relation is what matters even when, as in a case where one person is R-related to two other people, Relation R does not provide personal identity. Two other relations may have some slight importance: physical continuity, and physical similarity. (In the case of a few people, who are very beautiful, physical similarity may have great importance.)

Here is a brief sketch of how I shall argue for my conclusions. I shall first try to answer some objections to my claim that we could describe our lives in an *impersonal* way. I shall then try to show that, even if we are not aware of this, we are naturally inclined to believe that our identity must always be determinate. We are inclined to believe, strongly, that this *must* be so. I shall next argue that this natural belief cannot be true unless we are separately existing entities. I shall then argue for conclusion (1), that we are not such entities. And I shall argue that, because (1) is true, so are my other three conclusions.

Most of us would accept some of the claims that I shall be denying. I shall thus be arguing that most of us have a false view about ourselves, and about our actual lives. If we come to see that this view is false, this may make a difference to our lives.