

Self: Personal Identity

Eric T. Olson

University of Sheffield

In W. Banks, ed., *Encyclopedia of Consciousness* (Elsevier Academic Press), vol. 2, 301-312.

Introduction

The problems of personal identity

- The persistence question

- The population question

- The personal-ontology question

- The question of what matters in identity

The persistence question

Proposed answers to the persistence question

The psychological-continuity view

Fission

The beginning and end of life

The persistence question and personal ontology

The too-many-minds problem

Conclusion

keywords

commissurotomy, consciousness, dissociative identity disorder, David Hume, John Locke, memory, metaphysics, organism, persistence, psychological continuity

glossary

cerebrum The largest and uppermost part of the brain, responsible for higher cognitive functions.

dissociative-identity disorder A rare condition in which a human being acts as if he or she were inhabited alternately by a number of people, with different personalities, memories, and cognitive abilities.

hemispherectomy The surgical removal of a cerebral hemisphere.

persistence conditions The conditions logically necessary and sufficient for a thing to continue existing, as opposed to ceasing to exist.

personal ontology The basic metaphysical nature of human people, such as whether we are material or immaterial.

temporal part A part of an object that has the same spatial extent as that object has during a period of time, and the same temporary properties, but a shorter temporal extent: for example, the first half of a football match.

synopsis

Personal identity deals with the many philosophical questions about ourselves that arise by virtue of our being people. The most frequently discussed is what it takes for a person to persist through time. Many philosophers say that we persist by virtue of psychological continuity. Others say that our persistence is determined by brute physical facts, and psychology is irrelevant. In choosing among these answers we must consider not only what they imply about who is who in particular cases, both real and imaginary, but also their implications about our metaphysical nature in general.

Introduction

Personal identity deals with philosophical questions about ourselves that arise by virtue of our being people (or persons). Some of these questions are familiar ones that occur to all of us: What am I? When did I begin? What will happen to me when I die? Others are more abstruse.

Many philosophers, following Locke and Hume, give consciousness a central role in answering questions of personal identity. They say that many of these questions are nothing more than questions about the unity and continuity of consciousness, or at least that facts about consciousness and related psychological matters suffice to settle the facts about personal identity. Others disagree, saying that consciousness is irrelevant to most questions of personal identity.

(A terminological note: this article is about the self in the sense that my self is simply myself--me, the author--and not about other senses of the word self.)

The Problems of Personal Identity

There is no single problem of personal identity, but rather a wide range of loosely connected questions. Here are some of the most central and important

ones.

The persistence question

One question is what it takes for a person (or a human person) to persist or survive from one time to another. What sorts of adventures could you possibly survive, in the broadest sense of the word possible? What sort of event would necessarily bring your existence to an end? What determines which past or future being is you? What, in other words, does our persistence through time consist in?

Suppose you point to a boy or girl in an old class photograph and say, "That's me." What makes you that one, rather than one of the others? What is it about the way that person then relates to you that makes her you? For that matter, what makes it the case that anyone at all who existed back then is you? Is a certain boy or girl you by virtue of psychological connections to you? Is it, for instance, because you are now able to remember things that happened to her back then? Or should we look instead to brute physical or biological relations? Might that person be you because she is the same biological organism as you are, or because you and she have the same body? Or does the answer lie somewhere else entirely? This is the question of personal identity over time, or the persistence question. It applies to our survival into the future as well as our existence in the past.

Historically this question arises when people wonder whether we might continue to exist after we die (in Plato's Phaedo, for instance). Whether this could happen depends on whether biological death necessarily brings one's existence to an end. Imagine that after your death there really will be someone, in the next world or in this one, who is in some ways like you. How would that being have to relate to you for it actually to be you, rather than someone else? What would the Higher Powers have to do to keep you in existence after your death? Or is there anything they could do? The answer to these questions depends on the answer to the persistence question.

The population question

We can think of the persistence question as asking which of the characters introduced at the beginning of a story have survived to become the characters at the end of it. But we can also ask how many characters are on the stage at any one time. What determines how many of us there are now? If there are some six billion people on the earth at present, what facts--biological, psychological, or what have you--make this the case? The question is not what causes there to be a certain number of people at a given time, but rather what there being that number of people consists in. It is like asking what sort of configuration of pieces amounts to black's winning a game of chess, rather than like asking what sorts of moves might lead to its winning. This is the population question. (It is sometimes called the question of synchronic identity, as opposed to the diachronic identity of the persistence question.)

One possible answer is that the number of people at any given time is simply the number of human organisms there are then, which is determined by brute biology. (We may have to discount human organisms in an immature or radically defective state--human embryos and human vegetables, for instance--as many deny that they count as genuine people.) Another answer, proposed by Hume and further developed in Kant, is that the number of people is determined by facts about psychological unity and disunity. These answers may agree in ordinary cases, but diverge in unusual ones. Other answers are also possible.

Commissurotomy--the severing of the corpus callosum connecting the cerebral hemispheres--can produce a certain degree of mental disunity, illustrated in extreme cases by such peculiar behavior as simultaneously pulling one's trousers up with one hand and pulling them down with the other. Some even say that commissurotomy can produce two separate streams of consciousness. There is dispute about how much mental disunity there really is in these cases. But many

philosophers say that if there were enough disunity, of the right sort, there would be two people--two thinking, conscious beings--sharing a single organism, each with its own unified consciousness and its own set of beliefs, experiences, and actions. Dissociative identity disorder (multiple personality) raises similar issues: we can ask whether Dr Jekyll and Mr Hyde would be two people or one. (At least this is so if one of them continues to exist while the other is active; otherwise this case is a puzzle for the persistence question.) Does the existence of two or more separate streams of consciousness entail that there are two or more conscious beings?

In these cases, or in imaginary but possible extensions of them, there can be psychological disunity of the sort that is ordinarily present in cases where there are two or more people: there may be two independent sets of beliefs and experiences, and two separate streams of consciousness. Yet at the same time there is biological unity of the sort we find when there is just one person: there is always just one biological organism present, and one brain. If the number of people or thinking beings is determined by facts about psychological unity (unity of consciousness, for example), there will be two or more people in these cases. If it is determined by nonpsychological facts, such as biological unity, there will be one. Philosophers disagree about this, though the psychological-unity account is currently more popular.

The personal-ontology question

More generally, we can ask about our basic metaphysical nature. What are the most general and fundamental characteristics of human people? This is one of the more abstruse questions about personal identity, but it is of central importance. We can call it the personal-ontology question. It is really a collection of more specific questions, such as these:

What, at the most basic level, are we made of? Are we made up entirely of matter, as stones are? Or are we made up at least partly of something other than

matter?

If we are made up of matter, what matter is it? Is it just the matter that makes up our bodies? Do we extend all the way out to our skin and no further, or might we be larger or smaller than our bodies? Where, in other words, do our spatial boundaries lie? More fundamentally, what fixes those boundaries? If we do extend right out to our skin, why that far and no further? Does this have to do with the extent of a certain sort of conscious awareness, as Locke thought, or does the answer lie in brute biology?

Are we ontologically independent beings? Or is each of us a mere state or aspect of something else? Think of a knot in a rope. It is not ontologically independent. It cannot exist without the rope: you can't take it away in your pocket and leave the entire rope behind. It is a state or aspect of the rope. The same goes for events, such as the rope's gradually wearing out. The rope itself, though, does not seem to be a state or an aspect of anything else: nothing appears to stand to the rope as the rope stands to the knot. It seems to be an ontologically independent being--what metaphysicians call a substance. The question, then, is whether we are substances, like ropes, or whether we are states or aspects or events, like knots. Is there something--an organism or a mass of matter, perhaps--that stands to you as a rope stands to a knot?

An answer to these questions and others like them would tell us our basic metaphysical nature. The persistence and population questions also fall under the personal-ontology question. Here are the main proposed answers to this question:

We might be biological organisms (and thus material substances). This view, held by Aristotle, is currently defended under the name of animalism. Most philosophers, both past and future, however, have rejected it.

We might be immaterial substances or souls, as Plato, Descartes, and Leibniz held. This view faces notorious difficulties in explaining how the soul relates to the body--what enables my soul to animate this particular organism and no other, for

instance--and has few advocates today.

Many contemporary philosophers say that we are material things but not organisms. One such view is that we are spatial parts of organisms, such as brains. A more popular view is that we are temporal parts of organisms: you are spatially the same size as your animal body but temporally shorter, in that the animal extends further into the past or future than you do. Yet another currently popular view is that we are nonorganisms made of just the same matter as our animal bodies. The thinking behind this is that the same matter can make up two different objects at once. Specifically, the matter making up a typical human organism also makes up a certain nonorganism, and such things are what we are.

Hume proposed that we are bundles of mental states and events. Our parts are not cells or atoms, he thought, but memories and dreams. We are not substances, but events or processes: Hume compares a person to a theater production. Advocates of this view need to explain which mental states make up a person, or a conscious being more generally: what ties the bundles together, as it were. They must also defend the baffling claim that a bundle of mental states is the sort of thing that can think or be conscious.

A few philosophers espouse the paradoxical view that we don't really exist at all. When we say 'I', we fail to refer to anything. Your atoms may be real enough; perhaps even the thoughts and experiences we call yours are real; but those atoms or mental states are not parts or states of any conscious being.

To answer the personal-ontology question, we need to know what sort of thing it is--if anything--that thinks and is conscious. When I wonder what I am, what is it that wonders?

What matters in identity

Then there is the question of the practical importance of facts about our persistence through time. We ordinarily care deeply about what happens to us in

the future, differently from the way we care about the future welfare of others. Why is this? More specifically, what reason do we have for this special selfish concern? What justifies it? This is the question of what matters in identity.

Imagine that surgeons are going to destroy all of you but your brain and then transplant that organ into my head, and that neither of us has any choice about this. The resulting person, we may suppose, will have all your memories and other psychological features and none of mine, and will think he is you. To make things simple, suppose that we are both entirely selfish. Which of us will have a reason to care about the welfare of the resulting being? Suppose he will be in terrible pain after the operation unless one of us pays a large sum in advance. Which of us will have a reason to pay? And why? (In the same way, we can ask whether the resulting being will be morally responsible for my actions or for yours--or perhaps both, or neither.)

The answer may seem to turn simply on whether the resulting person would be you or I: if he is you, you have a reason to care about his welfare and I don't; if he is me, the reverse is true. And the reason for this may seem blindingly obvious: I care in a special way about what happens to myself in the future simply because he is me. Each person has a special reason to care about her own future welfare, and hers alone. The only one whose future welfare I cannot rationally ignore is myself. Likewise, only I myself can be morally responsible for my past actions. What matters in identity, we might say, is identity.

But some deny this. They say that I could have an entirely selfish reason to care about someone else's welfare, for his own sake. The reason why I care, selfishly, about a certain person's future welfare is not that he is me, but that he relates to me in some other way. It may be, for instance, that I care selfishly about a certain future person because he has inherited my mental life, and can remember my actions and no one else's. Ordinarily, the only person who has inherited my mental life and can remember my actions is me, and so I ordinarily care selfishly only about myself. But

the fact that he is me is not the reason why I care about him. I care about him because of his psychological connections with me. If someone else, other than me, were connected with me in that psychological way, he would have what matters in identity, and I ought to transfer my selfish concern to him. Perhaps that person would then also be responsible for my actions. We will return to this theme in a later section.

Understanding the Persistence Question

Let us turn now to the persistence question: what is necessary and sufficient for a past or future being to be you, or what it takes for a human person to continue existing from one time to another. The concept of identity over time, or persistence, is a notorious source of confusion, and the persistence question is often misunderstood.

The question asks what it takes for a being existing at another time to be numerically identical with you. For this and that to be numerically identical is for them to be one and the same: one thing rather than two. Numerical identity is different from qualitative identity or exact similarity: identical twins may be qualitatively identical (nearly), but not numerically identical. A past or future person need not, at that past or future time, be exactly like you are now in order to be you--that is, in order to be numerically identical with you. You don't remain qualitatively the same throughout your life. You change.

Of course, someone might say (as Hume did) that a past or future being could not be you unless he or she were then qualitatively just like you are now. This would mean that no person could survive any change whatever: even blinking your eyes would be fatal, resulting in your ceasing to exist and being replaced by someone else. In that case there would be no point in asking the persistence question. Virtually all discussions of personal identity over time assume that it is

possible for a person to change.

This point enables us to distinguish the persistence question from another one that sounds similar. We sometimes ask what it takes for someone to remain the same person from one time to another--as opposed to simply asking what it is for someone to continue existing. If I were to change in certain ways--if I lost a great deal of memory, say, or my personality changed dramatically--we might describe this by saying that I should no longer be the person I was before, but a different one.

The question of what it takes for someone to remain the same person in this sense is not the persistence question. It is not about numerical identity at all. If it were, it would answer itself, for I could not possibly come to be a numerically different person from the one I am now. Nothing can start out as one thing and end up as another thing: a thing can change qualitatively, but it cannot change numerically. This has nothing to do with personal identity in particular, but reflects the logic of identity in general. Those who say that certain events could make you a different person from the one you were before mean that you would still exist, but would have changed qualitatively in some profound and important way. They mean that you might come to be a radically different kind of person. If the resulting person were not numerically identical with you, it would not be the case that you yourself were a different person; rather, you would have ceased to exist and been replaced by someone else.

Proposed Answers to the Persistence Question

There have been three main sorts of answers to the persistence question. The most popular, the psychological-continuity view, says that some psychological relation is both necessary and sufficient for one to persist. You are that future being that in some sense inherits its mental features--beliefs, memories, preferences, the capacity for rational thought and consciousness, or the like--from you; and you are

that past being whose mental features you have inherited in this way.

This view comes in different versions, varying in the sort of inheritance that figures in them: whether a future being, in order to be you, needs to acquire its mental features from you via a continuously functioning brain, for instance, or whether some less direct transfer, such as Star Trek teleportation, might suffice. They also vary in the sort of mental features that need to be inherited--whether mental contents, such as memories, or core mental capacities such as the capacity for thought and consciousness, for instance. And they vary in how much of your mental life they require to be preserved in order for you to survive. Presumably it would not suffice, for some future person to be me, for him to inherit only one of my belief-states. But how much is enough? Yet another disagreement has to do with fission cases, of which more later. Most philosophers writing on the persistence question since the mid-20th century have followed Locke in endorsing some version of the psychological-continuity view.

The second answer is the brute-physical view or bodily criterion, according to which our identity through time consists in some brute physical relation. You are that past or future being that has your body, or that is the same biological organism as you are, or the like. Whether you survive or perish has nothing to do with psychological facts. This is typically held in combination with the view that we are biological organisms (animalism), for organisms, including human animals, appear to persist by virtue of brute physical facts. Here too there are variants.

Hybrid views are also possible: one might propose that we need both mental and physical continuity to survive, or that either would suffice without the other. For present purposes we can treat these as versions of the psychological-continuity view.

Both the psychological-continuity and brute-physical views agree that there is something that it takes for us to persist--that our identity through time consists in or necessarily follows from something other than itself. A third view denies this.

Mental and physical continuity are evidence for our persistence, it says, but do not always guarantee it, and are not required. No sort of continuity is both absolutely necessary and completely sufficient for us to persist. There are no informative and non-trivial persistence conditions for people. This is called the simple view or anticriterialism. It is often, though not always, combined with the view that we are immaterial souls.

The Psychological-Continuity View

The psychological-continuity view is intuitively appealing. Suppose your brain, or your cerebrum, were transplanted into my head, resulting in a being with most of your mental features and few if any of mine. He would, of course, believe that he was you and not me. Many people find it obvious that he would be right, precisely because of his psychological connections with you. (The brute-physical view, by contrast, seems to imply that he would be me, and that you would stay behind with an empty head if your cerebrum were transplanted. Most people find this strongly counterintuitive.) But it is notoriously difficult to get from this conviction to a plausible answer to the persistence question.

What psychological relation might our identity through time consist in? Memory seems to play an important role: the one who got your transplanted brain would seem to be you at least partly because he or she would remember your life. So one version of the psychological-continuity view is that a past or future being is you just in the case that you can now remember an experience that being had then, or that being can then remember an experience you are having now. More precisely,

Necessarily, a person x existing at time t is identical to a being y existing at another time, t^* , if and only if x can remember, at t , an experience y has at t^* or y can remember, at t^* , an experience x has at t .

This sort of view, the memory criterion, is often attributed to Locke (though it is doubtful whether he or anyone else actually held it).

The memory criterion faces two well-known problems. First, suppose a young student is arrested for drunken excesses. As a middle-aged neuroscientist, she retains a vivid memory of this event. In her dotage, however, she remembers her science career, but has entirely forgotten the arrest and all the other events of her youth. According to the memory criterion, the young student would be the middle-aged scientist, the scientist would be the old woman, but the old woman would not be the young student. This is an impossible result: if x and y are one and y and z are one, x and z cannot be two.

The second problem is that it seems to belong to the very idea of remembering an experience that you can remember only your own. You can no more remember someone else's experiences than you could be a married bachelor. To remember being arrested (or the experience of it) is to remember yourself being arrested. That makes it trivial and uninformative to say that you are the person whose experiences you can remember--that memory continuity is sufficient for personal identity. It is uninformative because you could not know whether someone genuinely remembered a past experience without already knowing whether she was the one who had it. Suppose we want to know whether Blott, who exists now, is the same as Clott, whom we know to have existed at some time in the past. The memory criterion tells us that Blott is Clott if Blott can now remember an experience of Clott's that occurred at that past time. But even if Blott seems to remember one of Clott's experiences from that time, this counts as genuine memory only if Blott really is Clott. We should already have to know who is who before we could apply the theory that is supposed to tell us.

(Note, however, that this is a problem only for the claim that memory connections are sufficient for identity, not for the claim that they are necessary. There is nothing trivial or uninformative in saying that a future being can be you

only if he or she can then remember an experience you are having now.)

One response to the first problem is to modify the memory criterion by switching from direct to indirect memory connections: the old woman is the young student because in her old age she can recall experiences the scientist had at a time when the scientist remembered the student's life. The second problem is traditionally (and controversially) met by inventing a new concept, quasi-memory, which is just like memory but without the identity requirement: even if it is self-contradictory to say that I remember an experience of someone else's, I could still quasi-remember it. Neither move gets us far, however, for even the modified memory criterion faces a more obvious objection: there are many times in my past that I cannot remember or quasi-remember at all, and to which I am not linked even indirectly by an overlapping chain of memories. There is no time when I could recall anything that happened to me while I was dreamlessly sleeping last night. The memory criterion has the absurd implication that I did not exist at any time when I was completely unconscious.

A better solution appeals to causal dependence. We can define two notions, psychological connectedness and psychological continuity. A being is psychologically connected, at some future time, with me as I am now just if he is in the psychological states he is in then in large part because of the psychological states I am in now. Having a current memory (or quasi-memory) of an earlier experience is one sort of psychological connection--the experience causes the memory of it--but there are others. Importantly, one's current mental states can be caused in part by mental states one was in at a times when one was unconscious: for example, most of your current beliefs are the same ones you had while you slept last night. We can then define the second notion thus: I am now psychologically continuous with a past or future being just if my current mental states relate to those he is in then by a chain of psychological connections.

This enables us to avoid the most obvious objections to the memory criterion by

saying that a person who exists at one time is identical with something existing at another time if and only if the first is, at the first time, psychologically continuous with the second as she is at the second time.

Fission

A serious difficulty for the psychological-continuity view is the fact that you could be psychologically continuous with two future people at once. If your cerebrum were transplanted, the resulting person would be psychologically continuous with you (even if, as recent neuroscience has shown, there would also be important psychological differences). If we destroyed one of your cerebral hemispheres, the resulting being would also be psychologically continuous with you.

(Hemispherectomy--even the removal of the left hemisphere, which controls speech--is considered a drastic but acceptable treatment for otherwise-inoperable brain tumors.) What if we did both at once, destroying one of your cerebral hemispheres and transplanting the other? Then too, the one who got the transplanted hemisphere would be psychologically continuous with you, and according to the psychological-continuity view he or she would be you.

But now let both hemispheres be transplanted, each into a different empty head. The resulting beings--call them Lefty and Righty--will each be psychologically continuous with you. The psychological-continuity view as we have stated it implies that any future being who is psychologically continuous with you must be you. It follows that you are Lefty and also that you are Righty. But that cannot be, for Lefty and Righty are two, and one thing cannot be numerically identical with two things. (If you and Lefty are one, and you and Righty are one, Lefty and Righty cannot be two.) This is the fission problem.

Psychological-continuity theorists have proposed two different solutions to this problem. One, sometimes called the multiple-occupancy view, says that if there is fission in your future, then there are, so to speak, two of you even now. What we

think of as you is really two people, who are now exactly similar and located in the same place, doing the same things and thinking the same thoughts. The surgeons merely separate them.

The multiple-occupancy view is almost invariably combined with the thesis that people and other persisting things are made up of temporal parts. For each person there is, for instance, such a thing as her first half, which is just like the person only briefer--something analogous to the first half of a football match. On this account, the multiple-occupancy view is that Lefty and Righty coincide before the operation by sharing their pre-operative temporal parts, and diverge later by having different temporal parts located afterwards. Lefty and Righty are like two roads that coincide for a stretch and then fork, sharing some of their spatial parts but not others. At the places where the roads overlap, they will look just like one road. Likewise, the idea goes, at the times before the operation when Lefty and Righty share their temporal parts, they will look just like one person--even to themselves. Whether people really are made up of temporal parts, however, is a disputed metaphysical question.

The other solution to the fission problem abandons the intuitive claim that psychological continuity by itself suffices for one to persist. It says, rather, that you are identical with a past or future being who is psychologically continuous with you only if no other being is then psychologically continuous with you. (There is no circularity in this. We need not know the answer to the persistence question in order to know how many people there are at any one time; that comes under the population question.) This means that neither Lefty nor Righty is you. They both come into existence when your cerebrum is divided. If both your cerebral hemispheres are transplanted, you cease to exist--though you would survive if only one were transplanted and the other destroyed.

This proposal, the non-branching view, has the surprising consequence that if your brain is divided, you will survive if only one half is preserved, but you will die if

both halves are. That is just the opposite of what most of us expect: if your survival depends on the functioning of your cerebrum (because that is what underlies psychological continuity), then the more of that organ we preserve, the greater ought to be your chance of survival. In fact the non-branching view implies you would perish if one of your hemispheres were transplanted and the other left in place: you can survive hemispherectomy only if the excised hemisphere is immediately destroyed.

The non-branching view makes the question of what matters in identity especially acute. Faced with the prospect of having one of your cerebral hemispheres transplanted, there would seem to be no reason to prefer that the other be destroyed. Most of us would rather have both preserved, even if they end up in different heads. Yet on the non-branching view that is to prefer death over continued existence. Some philosophers infer from this that that is precisely what we ought to prefer. Insofar as we are rational, we don't want to continue existing. Or at least we don't want it for its own sake. I only want there to be, in the future, someone psychologically continuous with me, whether or not he is strictly me. Likewise, even the most selfish person has a reason to care about the welfare of the beings who would result from her undergoing cerebral fission, even if, as the non-branching view implies, neither would be her. In the fission case, the sorts of practical concerns you ordinarily have for yourself seem to apply to someone who isn't strictly you. This suggests more generally that facts about who is numerically identical with whom have no practical importance. What matters practically is, rather, who is psychologically continuous with whom.

The Beginning and End of Life

Advocates of the psychological-continuity view commonly discuss its implications about who is who in imaginary science-fiction scenarios. But the view also has important implications about when we begin and end in real life.

We ordinarily suppose that each of us was once an embryo and then a fetus. We suppose that the fetus my mother once carried was born, grew up, and later wrote this article. (This does not imply that a fetus in the womb is already a person, but at most that it is a potential person: something that can come to be a person. We might start out as non-people and only later become people, just as we are first children and later adults.) But the psychological-continuity view implies that none of us was ever an embryo. Whatever exactly psychological continuity amounts to, you can never be psychologically continuous with a being that has no psychological states at all: you cannot inherit your current psychological states from a being that has none to pass on. And the embryo from which you developed had no psychological states. If psychological continuity is necessary for you to persist, as the psychological-continuity view says, then you were never an embryo. The embryo from which you developed was not you. Not only was it not a person; it was not even something that could come to be a person. You did not come into being when the embryo did, but some time later.

How much later? The psychological-continuity view implies that you could not have existed at a time when you had no psychological states or properties at all. The earliest you could have begun to exist is when the fetus you developed from acquired its first mental states. No one is certain when that was, but embryologists agree that the fetus has no mental states--not even the capacity to feel pain--before mid-gestation at the earliest. So the psychological-continuity view implies that none of us was ever a four-month-old fetus. The fetus you see in an ultrasound scan is not something that can one day come to be a person.

According to the psychological-continuity view, when you and I come into existence depends on what sort of psychological continuity our identity over time consists in. Many psychological-continuity theorists say that we do not come into being when the first mental capacities develop in the course of fetal development, but only with the advent of more sophisticated mental properties, such as the

capacity for self-consciousness. On their view, we do not begin to exist until we are born, or perhaps several months or even years afterwards.

The psychological-continuity view also has important implications about when we come to an end. It implies, for instance, that you could not exist in a persistent vegetative state. A being in such a state has no mental states, and thus has not in any way inherited yours. There is complete psychological discontinuity between you as you are now and the human vegetable that would result if you were to lapse into such a state. If all your mental states were destroyed, you would cease to exist, even if the rest of you kept on working as before. The resulting being, despite being alive by most definitions of the word (and in all current legal systems), would not be you. If the doctors withdraw the tubes that feed it, they do not kill anyone, or allow anyone to die. Some being or other dies when the feeding tubes are withdrawn, of course, but not anything that was ever a person. This too goes against what many of us believe.

The reason we tend to think that each of us was once a fetus and might end up in a vegetative state is that in each of these cases there is a good deal of brute physical continuity of an appropriate sort: the same sort of physical continuity we normally exhibit from one day to the next. So these considerations seem to support the brute-physical view as against the psychological-continuity view. More generally, the brute-physical view appears to give the right verdicts about who is who in all real-life cases.

The Persistence Question and Personal Ontology

We have been evaluating answers to the persistence question by considering their implications about who is who in cases where they disagree. We saw that the psychological-continuity view gives attractive results in science-fiction stories (such as the cerebrum transplant), but has implausible consequences about when we begin and end in real life. With the brute-physical view it is the other way round.

But this sort of debate over cases is not the only relevant consideration. A good answer to the persistence question must also be compatible with an acceptable answer to the personal-ontology question. What sort of things could we be if the psychological-continuity view were true? That is, what sort of things could persist by virtue of psychological continuity?

The psychological-continuity view appears to rule out our being organisms. It says that our persistence consists in some sort of psychological continuity. As we have seen, this means that you would go along with your transplanted cerebrum, because the one who ended up with that organ, and no one else, would be psychologically continuous with you. And it implies that if you were to lapse into a persistent vegetative state, you would cease to exist, because no one would then be psychologically continuous with you. But the persistence of a human organism does not consist in any sort of psychological continuity. If we transplanted your cerebrum, the human organism--your body--would not go along with that organ. It would stay behind with an empty head. If you were an organism, then you would stay behind with an empty head, contrary to the psychological-continuity view. Likewise, no human organism ceases to exist by lapsing into a persistent vegetative state. If you were an organism, you could survive as a human vegetable. What the psychological-continuity view says about our persistence through time is not true of human organisms. So if that view is true, we could not be organisms.

Psychological-continuity theorists must deny that we are organisms, and find another account of what we are. Some say that we are temporal parts of organisms. (If your cerebrum were transplanted, you would be made up of temporal parts of two organisms.) Others say that we are material things made of the same matter as organisms, or bundles of mental states. But the view that we are not organisms faces an awkward problem.

The Too-Many-Minds Problem

Even if you are not an organism, your body is. That organism--a human animal--thinks and is conscious. In fact it would seem to be psychologically indistinguishable from you. So if you are not that animal, but something else, it follows that there is a conscious, intelligent being other than you, now sitting in your chair and thinking your thoughts. This means that there are at least twice as many thinking beings as the census reports: for each of us, there is another thinking, conscious being, namely the animal we call one's body. Worse, you ought to wonder which of the two thinkers is you. You may believe that you are the nonanimal (because you accept the psychological-continuity view, perhaps). But the animal has the same grounds for believing that it is a nonanimal as you have for supposing that you are. Yet it is mistaken. For all you know, you might be the one making this mistake. If you were the animal and not the person, you would never be any the wiser.

An analogy may help here. Imagine a three-dimensional duplicating machine. When you step into the in box, it reads off your information and assembles a perfect duplicate of you in the out box. The process causes temporary unconsciousness, but is otherwise harmless. Two beings wake up, one in each box. The boxes are indistinguishable. Because each being will have the same apparent memories and perceive identical surroundings, each will think that he or she is you, and will have the same evidence for this belief. But only one will be right. How could you ever know, afterwards, whether you were the original or the newly-created duplicate?

In the same way, the psychological-continuity view appears to leave you without any grounds for supposing that you are the being who persists by virtue of psychological continuity, rather than the animal. So even if the psychological-continuity view is true, it seems that you could never know whether it applied to you: for all you can tell, you may instead be an organism with brute physical persistence conditions. This is the too-many-minds or too-many-thinkers problem. It threatens

to make any view according to which we are not organisms look absurd. Only animalism, the view that we are organisms (and that there are no beings who persist by virtue of psychological continuity), appears to escape it.

Psychological-continuity theorists have proposed two ways of solving the problem. One is to argue that despite appearances, human animals are not psychologically indistinguishable from us. Although our animal bodies share our brains, are physically just like us, and show all the outward signs of consciousness and intelligence, they themselves do not think and are not conscious.

If human organisms cannot be conscious, then presumably no biological organism of any sort could have any mental properties at all. Why not? It may be, as Descartes and Leibniz argued, because organisms are material things. Only an immaterial thing could think or be conscious. You and I must therefore be immaterial. Though this would solve the to-many-thinkers problem, it raises many others, and few philosophers nowadays accept it.

One notable attempt to explain why organisms should be unable to think that is compatible with our being material things appeals to the nature of mental states and properties. It says that whatever thinks or is conscious must persist by virtue of psychological continuity. That is because it belongs to the nature of a mental state that it tend to have certain characteristic causes and effects in the being that is in the state, and not in any other being. (This is a version of the functionalist theory of mind.)

For instance, your preference for chocolate over vanilla must tend to cause you, and no one else, to choose chocolate. Now if an organism were to have such a preference, that state might cause another being to choose chocolate, for an organism's cerebrum might be transplanted into another organism. That would violate the proposed account of mental states. It follows, on that account, that no organism could have a preference; and similar reasoning goes for other mental states. The persistence conditions of organisms are incompatible with their having

mental properties. But a material thing that would go along with its transplanted cerebrum--a being of which the psychological-continuity view was true--could have mental states. It would follow that you and I, who obviously have mental states, persist by virtue of psychological continuity, and thus are not organisms. This would solve the too-many-thinkers problem, and at the same time show that the psychological-continuity view is true. It is, however, a minority view.

The other alternative for psychological-continuity theorists is to concede that human organisms think as we do, so that you are one of two beings now thinking your thoughts, but try to explain how we can still know that we are not those organisms. One strategy for doing this focuses on the nature of personhood and first-person reference. It proposes that not just any being with mental properties of the sort that you and I have--rationality and self-consciousness, for instance--counts as a person. A person must also persist by virtue of psychological continuity. It follows from this that human animals, despite being psychologically just like from us, are not people.

The proposal goes on to say that personal pronouns such as 'I' refer only to people. This means that when your animal body says or thinks 'I', it does not refer to itself. Rather, it refers to you, the person who says it at the same time. When the animal says 'I am a person', it does not thereby express the false belief that it is a person, but rather the true belief that you are. It follows that the animal is not mistaken about which thing it is, because it has no first-person beliefs about itself at all. And you are not mistaken either. You can infer that you are a person from the linguistic facts that you are whatever you refer to when you say 'I', and that 'I' never refers to anything but a person. You can know that you are not the animal thinking your thoughts because it is not a person, and personal pronouns never refer to nonpeople.

Although this proposal avoids the surprising claim that organisms cannot have mental properties, it gives a highly counterintuitive view of what it is to be a person.

And it still implies that there are twice as many intelligent, conscious beings as we thought. It remains a controversial view.

Conclusion

The answer to the persistence and personal-ontology questions is likely to depend on more general metaphysical matters. Which account of personal identity you find attractive will depend on your view about the metaphysics of material things in general.

Consider, for example, the view that all persisting things are made up of temporal parts, mentioned earlier. As we saw, psychological-continuity theorists who hold this view can say that if your cerebrum were divided and each half transplanted into a different head, you would survive twice over, as it were. They need not accept the non-branching view, according to which you would cease to exist, but would survive if just one cerebral hemisphere were preserved. So temporal-parts theorists will find the psychological-continuity view more attractive than those who reject temporal parts.

The temporal-parts theory also implies that animalists face their own version of the too-many-thinkers problem. Suppose you are an animal, and that you are made up of temporal parts. Then there is a temporal part of you that is just like you are now, except that it extends in time only from midnight last night until midnight tonight: your today-part. It is now sitting in your chair, fully conscious and thinking your thoughts. How could you ever know that you are not your today-part, and have only hours to live? This looks just as troubling for the animalist as the problem of how you can know you are not your animal body is for the psychological-continuity theorist. And any solution the animalist can give to her version of the problem will suit the psychological-continuity theorist equally well for hers. (Temporal-parts theorists solve the too-many-thinkers problem by appealing to the account of personhood and first-person reference sketched in the previous

section.)

So if the temporal-parts theory is true, the problems facing the psychological-continuity theory appear considerably less than they would otherwise be. Whether persisting things really do have temporal parts, however, is not a question about personal identity. It is a general question about the fundamental nature of the world. The main problems of personal identity cannot be debated in isolation.

Further Reading

- Baker, L. R. (2000). *Persons and bodies: a constitution view*. Cambridge: Cambridge University Press.
- Garrett, B. (1998). *Personal identity and self-consciousness*. London and New York: Routledge.
- Hudson, H. (2001). *A materialist metaphysics of the human person*. Ithaca: Cornell University Press.
- Kolak, D. and Martin, R. (eds.) (1991). *Self and identity: contemporary philosophical issues*. New York: Macmillan.
- Martin, R. and Barresi, J. (eds.) (2003). *Personal identity*. Oxford: Blackwell.
- Martin, R. (2006). *The rise and fall of soul and self: an intellectual history of personal identity*. New York: Columbia University Press.
- Noonan, H. (2003). *Personal identity* (2nd edn.). London and New York: Routledge.
- Olson, E. T. (1997). *The human animal: personal identity without psychology*. New York: Oxford University Press.
- Olson, E. T. (2007). *What are we? a study in personal ontology*. New York: Oxford University Press.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Perry, J. (ed.) (1975). *Personal identity*. Berkeley: University of California Press.
- Shoemaker, S. and Swinburne, R. (1984). *Personal identity*. Oxford: Blackwell.
- Shoemaker, S. (1999). Self, body and coincidence. *Proceedings of the Aristotelian Society, Supplementary Volume 73*, 287-306.
- Unger, P. (1990). *Identity, consciousness and value*. New York: Oxford University Press.