



Philosophical Review

Personal Identity

Author(s): Derek Parfit

Reviewed work(s):

Source: *The Philosophical Review*, Vol. 80, No. 1 (Jan., 1971), pp. 3-27

Published by: Duke University Press on behalf of Philosophical Review

Stable URL: <http://www.jstor.org/stable/2184309>

Accessed: 06/06/2012 19:56

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Duke University Press and *Philosophical Review* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Review*.

<http://www.jstor.org>

PERSONAL IDENTITY¹

WE CAN, I think, describe cases in which, though we know the answer to every other question, we have no idea how to answer a question about personal identity. These cases are not covered by the criteria of personal identity that we actually use.

Do they present a problem?

It might be thought that they do not, because they could never occur. I suspect that some of them could. (Some, for instance, might become scientifically possible.) But I shall claim that even if they did they would present no problem.

My targets are two beliefs: one about the nature of personal identity, the other about its importance.

The first is that in these cases the question about identity must have an answer.

No one thinks this about, say, nations or machines. Our criteria for the identity of these do not cover certain cases. No one thinks that in these cases the questions "Is it the same nation?" or "Is it the same machine?" must have answers.

Some people believe that in this respect they are different. They agree that our criteria of personal identity do not cover certain cases, but they believe that the nature of their own identity through time is, somehow, such as to guarantee that in these cases questions about their identity must have answers. This belief might be expressed as follows: "Whatever happens between now and any future time, either I shall still exist, or I shall not. Any future experience will either be *my* experience, or it will not."

This first belief—in the special nature of personal identity—has, I think, certain effects. It makes people assume that the principle of self-interest is more rationally compelling than any moral principle. And it makes them more depressed by the thought of aging and of death.

¹ I have been helped in writing this by D. Wiggins, D. F. Pears, P. F. Strawson, A. J. Ayer, M. Woods, N. Newman, and (through his publications) S. Shoemaker.

I cannot see how to disprove this first belief. I shall describe a problem case. But this can only make it seem implausible.

Another approach might be this. We might suggest that one cause of the belief is the projection of our emotions. When we imagine ourselves in a problem case, we do feel that the question "Would it be me?" must have an answer. But what we take to be a bafflement about a further fact may be only the bafflement of our concern.

I shall not pursue this suggestion here. But one cause of our concern is the belief which is my second target. This is that unless the question about identity has an answer, we cannot answer certain important questions (questions about such matters as survival, memory, and responsibility).

Against this second belief my claim will be this. Certain important questions do presuppose a question about personal identity. But they can be freed of this presupposition. And when they are, the question about identity has no importance.

I

We can start by considering the much-discussed case of the man who, like an amoeba, divides.²

Wiggins has recently dramatized this case.³ He first referred to the operation imagined by Shoemaker.⁴ We suppose that my brain is transplanted into someone else's (brainless) body, and that the resulting person has my character and apparent memories

² Implicit in John Locke, *Essay Concerning Human Understanding*, ed. by John W. Yolton (London, 1961), Vol. II, Ch. XXVII, sec. 18, and discussed by (among others) A. N. Prior in "Opposite Number," *Review of Metaphysics*, 11 (1957-1958), and "Time, Existence and Identity," *Proceedings of the Aristotelian Society*, LVII (1965-1966); J. Bennett in "The Simplicity of the Soul," *Journal of Philosophy*, LXIV (1967); and R. Chisholm and S. Shoemaker in "The Loose and Popular and the Strict and the Philosophical Senses of Identity," in *Perception and Personal Identity: Proceedings of the 1967 Oberlin Colloquium in Philosophy*, ed. by Norman Care and Robert H. Grimm (Cleveland, 1967).

³ In *Identity and Spatio-Temporal Continuity* (Oxford, 1967), p. 50.

⁴ In *Self-Knowledge and Self-Identity* (Ithaca, N. Y., 1963), p. 22.

of my life. Most of us would agree, after thought, that the resulting person is me. I shall here assume such agreement.⁵

Wiggins then imagined his own operation. My brain is divided, and each half is housed in a new body. Both resulting people have my character and apparent memories of my life.

What happens to me? There seem only three possibilities: (1) I do not survive; (2) I survive as one of the two people; (3) I survive as both.

The trouble with (1) is this. We agreed that I could survive if my brain were successfully transplanted. And people have in fact survived with half their brains destroyed. It seems to follow that I could survive if half my brain were successfully transplanted and the other half were destroyed. But if this is so, how could I *not* survive if the other half were also successfully transplanted? How could a double success be a failure?

We can move to the second description. Perhaps one success is the maximum score. Perhaps I shall be one of the resulting people.

The trouble here is that in Wiggins' case each half of my brain is exactly similar, and so, to start with, is each resulting person. So how can I survive as only one of the two people? What can make me one of them rather than the other?

It seems clear that both of these descriptions—that I do not survive, and that I survive as one of the people—are highly implausible. Those who have accepted them must have assumed that they were the only possible descriptions.

What about our third description: that I survive as both people?

It might be said, "If 'survive' implies identity, this description makes no sense—you cannot be two people. If it does not, the description is irrelevant to a problem about identity."

I shall later deny the second of these remarks. But there are ways of denying the first. We might say, "What we have called 'the two resulting people' are not two people. They are one person.

⁵ Those who would disagree are not making a mistake. For them my argument would need a different case. There must be some multiple transplant, faced with which these people would both find it hard to believe that there must be an answer to the question about personal identity, and be able to be shown that nothing of importance turns upon this question.

I do survive Wiggins' operation. Its effect is to give me two bodies and a divided mind."

It would shorten my argument if this were absurd. But I do not think it is. It is worth showing why.

We can, I suggest, imagine a divided mind. We can imagine a man having two simultaneous experiences, in having each of which he is unaware of having the other.

We may not even need to imagine this. Certain actual cases, to which Wiggins referred, seem to be best described in these terms. These involve the cutting of the bridge between the hemispheres of the brain. The aim was to cure epilepsy. But the result appears to be, in the surgeon's words, the creation of "two separate spheres of consciousness,"⁶ each of which controls one half of the patient's body. What is experienced in each is, presumably, experienced by the patient.

There are certain complications in these actual cases. So let us imagine a simpler case.

Suppose that the bridge between my hemispheres is brought under my voluntary control. This would enable me to disconnect my hemispheres as easily as if I were blinking. By doing this I would divide my mind. And we can suppose that when my mind is divided I can, in each half, bring about reunion.

This ability would have obvious uses. To give an example: I am near the end of a maths exam, and see two ways of tackling the last problem. I decide to divide my mind, to work, with each half, at one of two calculations, and then to reunite my mind and write a fair copy of the best result.

What shall I experience?

When I disconnect my hemispheres, my consciousness divides into two streams. But this division is not something that I experience. Each of my two streams of consciousness seems to have been straightforwardly continuous with my one stream of consciousness up to the moment of division. The only changes in each stream are the disappearance of half my visual field and the loss of sensation in, and control over, half my body.

⁶ R. W. Sperry, in *Brain and Conscious Experience*, ed. by J. C. Eccles (New York, 1966), p. 299.

PERSONAL IDENTITY

Consider my experiences in what we can call my “right-handed” stream. I remember that I assigned my right hand to the longer calculation. This I now begin. In working at this calculation I can see, from the movements of my left hand, that I am also working at the other. But I am not aware of working at the other. So I might, in my right-handed stream, wonder how, in my left-handed stream, I am getting on.

My work is now over. I am about to reunite my mind. What should I, in each stream, expect? Simply that I shall suddenly seem to remember just having thought out two calculations, in thinking out each of which I was not aware of thinking out the other. This, I submit, we can imagine. And if my mind was divided, these memories are correct.

In describing this episode, I assumed that there were two series of thoughts, and that they were both mine. If my two hands visibly wrote out two calculations, and if I claimed to remember two corresponding series of thoughts, this is surely what we should want to say.

If it is, then a person’s mental history need not be like a canal, with only one channel. It could be like a river, with islands, and with separate streams.

To apply this to Wiggins’ operation: we mentioned the view that it gives me two bodies and a divided mind. We cannot now call this absurd. But it is, I think, unsatisfactory.

There were two features of the case of the exam that made us want to say that only one person was involved. The mind was soon reunited, and there was only one body. If a mind was permanently divided and its halves developed in different ways, the point of speaking of one person would start to disappear. Wiggins’ case, where there are also two bodies, seems to be over the borderline. After I have had his operation, the two “products” each have all the attributes of a person. They could live at opposite ends of the earth. (If they later met, they might even fail to recognize each other.) It would become intolerable to deny that they were different people.

Suppose we admit that they are different people. Could we still claim that I survived as both, using “survive” to imply identity?

We could. For we might suggest that two people could compose

a third. We might say, “I do survive Wiggins’ operation as two people. They can be different people, and yet be me, in just the way in which the Pope’s three crowns are one crown.”⁷

This is a possible way of giving sense to the claim that I survive as two different people, using “survive” to imply identity. But it keeps the language of identity only by changing the concept of a person. And there are obvious objections to this change.⁸

The alternative, for which I shall argue, is to give up the language of identity. We can suggest that I survive as two different people without implying that I am these people.

When I first mentioned this alternative, I mentioned this objection: “If your new way of talking does not imply identity, it cannot solve our problem. For that is about identity. The problem is that all the possible answers to the question about identity are highly implausible.”

We can now answer this objection.

We can start by reminding ourselves that this is an objection only if we have one or both of the beliefs which I mentioned at the start of this paper.

The first was the belief that to any question about personal identity, in any describable case, there must be a true answer. For those with this belief, Wiggins’ case is doubly perplexing. If all the possible answers are implausible, it is hard to decide which of them is true, and hard even to keep the belief that one of them must be true. If we give up this belief, as I think we should, these problems disappear. We shall then regard the case as like many others in which, for quite unpuzzling reasons, there is no answer to a question about identity. (Consider “Was England the same nation after 1066?”)

Wiggins’ case makes the first belief implausible. It also makes

⁷ Cf. David Wiggins, *op. cit.* p. 40.

⁸ Suppose the resulting people fight a duel. Are there three people fighting, one on each side, and one on both? And suppose one of the bullets kills. Are there two acts, one murder and one suicide? How many people are left alive? One? Two? (We could hardly say, “One and a half.”) We could talk in this way. But instead of saying that the resulting people are the original person—so that the pair is a trio—it would be far simpler to treat them as a pair, and describe their relation to the original person in some new way. (I owe this suggested way of talking, and the objections to it, to Michael Woods.)

PERSONAL IDENTITY

it trivial. For it undermines the second belief. This was the belief that important questions turn upon the question about identity. (It is worth pointing out that those who have only this second belief do not think that there must *be* an answer to this question, but rather that we must decide upon an answer.)

Against this second belief my claim is this. Certain questions do presuppose a question about personal identity. And because these questions *are* important, Wiggins' case does present a problem. But we cannot solve this problem by answering the question about identity. We can solve this problem only by taking these important questions and prizing them apart from the question about identity. After we have done this, the question about identity (though we might for the sake of neatness decide it) has no further interest.

Because there are several questions which presuppose identity, this claim will take some time to fill out.

We can first return to the question of survival. This is a special case, for survival does not so much presuppose the retaining of identity as seem equivalent to it. It is thus the general relation which we need to prize apart from identity. We can then consider particular relations, such as those involved in memory and intention.

“Will I survive?” seems, I said, equivalent to “Will there be some person alive who is the same person as me?”

If we treat these questions as equivalent, then the least unsatisfactory description of Wiggins' case is, I think, that I survive with two bodies and a divided mind.

Several writers have chosen to say that I am neither of the resulting people. Given our equivalence, this implies that I do not survive, and hence, presumably, that even if Wiggins' operation is not literally death, I ought, since I will not survive it, to regard it *as* death. But this seemed absurd.

It is worth repeating why. An emotion or attitude can be criticized for resting on a false belief, or for being inconsistent. A man who regarded Wiggins' operation as death must, I suggest, be open to one of these criticisms.

He might believe that his relation to each of the resulting people fails to contain some element which is contained in survival. But how can this be true? We agreed that he *would* survive if he

stood in this very same relation to only *one* of the resulting people. So it cannot be the nature of this relation which makes it fail, in Wiggins' case, to be survival. It can only be its duplication.

Suppose that our man accepts this, but still regards division as death. His reaction would now seem wildly inconsistent. He would be like a man who, when told of a drug that could double his years of life, regarded the taking of this drug as death. The only difference in the case of division is that the extra years are to run concurrently. This is an interesting difference. But it cannot mean that there are *no* years to run.

I have argued this for those who think that there must, in Wiggins' case, be a true answer to the question about identity. For them, we might add, "Perhaps the original person does lose his identity. But there may be other ways to do this than to die. One other way might be to multiply. To regard these as the same is to confuse nought with two."

For those who think that the question of identity is up for decision, it would be clearly absurd to regard Wiggins' operation as death. These people would have to think, "We could have chosen to say that I should be one of the resulting people. If we had, I should not have regarded it as death. But since we have chosen to say that I am neither person, I *do*." This is hard even to understand.⁹

My first conclusion, then, is this. The relation of the original person to each of the resulting people contains all that interests us—all that matters—in any ordinary case of survival. This is why we need a sense in which one person can survive as two.¹⁰

One of my aims in the rest of this paper will be to suggest such a sense. But we can first make some general remarks.

II

Identity is a one-one relation. Wiggins' case serves to show that what matters in survival need not be one-one.

⁹ Cf. Sydney Shoemaker, in *Perception and Personal Identity: Proceedings of the 1967 Oberlin Colloquim in Philosophy*, loc. cit.

¹⁰ Cf. David Wiggins, *op. cit.*, p. 54.

Wiggins' case is of course unlikely to occur. The relations which matter are, in fact, one-one. It is because they are that we can imply the holding of these relations by using the language of identity.

This use of language is convenient. But it can lead us astray. We may assume that what matters *is* identity and, hence, has the properties of identity.

In the case of the property of being one-one, this mistake is not serious. For what matters is in fact one-one. But in the case of another property, the mistake *is* serious. Identity is all-or-nothing. Most of the relations which matter in survival are, in fact, relations of degree. If we ignore this, we shall be led into quite ill-grounded attitudes and beliefs.

The claim that I have just made—that most of what matters are relations of degree—I have yet to support. Wiggins' case shows only that these relations need not be one-one. The merit of the case is not that it shows this in particular, but that it makes the first break between what matters and identity. The belief that identity *is* what matters is hard to overcome. This is shown in most discussions of the problem cases which actually occur: cases, say, of amnesia or of brain damage. Once Wiggins' case has made one breach in this belief, the rest should be easier to remove.¹¹

To turn to a recent debate: most of the relations which matter can be provisionally referred to under the heading “psychological continuity” (which includes causal continuity). My claim is thus that we use the language of personal identity in order to imply

¹¹ Bernard Williams' “The Self and the Future,” *Philosophical Review*, LXXIX (1970), 161-180, is relevant here. He asks the question “Shall I survive?” in a range of problem cases, and he shows how natural it is to believe (1) that this question must have an answer, (2) that the answer must be all-or-nothing, and (3) that there is a “risk” of our reaching the *wrong* answer. Because these beliefs are so natural, we should need in undermining them to discuss their causes. These, I think, can be found in the ways in which we misinterpret what it is to remember (cf. Sec. III below) and to anticipate (cf. Williams' “Imagination and the Self,” *Proceedings of the British Academy*, LII [1966], 105-124); and also in the way in which certain features of our egoistic concern—e.g., that it is simple, and applies to all imaginable cases—are “projected” onto its object. (For another relevant discussion, see Terence Penelhum's *Survival and Disembodied Existence* [London, 1970], final chapters.)

such continuity. This is close to the view that psychological continuity provides a criterion of identity.

Williams has attacked this view with the following argument. Identity is a one-one relation. So any criterion of identity must appeal to a relation which is logically one-one. Psychological continuity is not logically one-one. So it cannot provide a criterion.¹²

Some writers have replied that it is enough if the relation appealed to is always in fact one-one.¹³

I suggest a slightly different reply. Psychological continuity is a ground for speaking of identity when it is one-one.

If psychological continuity took a one-many or branching form, we should need, I have argued, to abandon the language of identity. So this possibility would not count against this view.

We can make a stronger claim. This possibility would count in its favor.

The view might be defended as follows. Judgments of personal identity have great importance. What gives them their importance is the fact that they imply psychological continuity. This is why, whenever there is such continuity, we ought, if we can, to imply it by making a judgment of identity.

If psychological continuity took a branching form, no coherent set of judgments of identity could correspond to, and thus be used to imply, the branching form of this relation. But what we ought to do, in such a case, is take the importance which would attach to a judgment of identity and attach this importance directly to each limb of the branching relation. So this case helps to show that judgments of personal identity do derive their importance from the fact that they imply psychological continuity. It helps to show that when we can, usefully, speak of identity, this relation is our ground.

This argument appeals to a principle which Williams put forward.¹⁴ The principle is that an important judgment should be asserted and denied only on importantly different grounds.

¹² "Personal Identity and Individuation," *Proceedings of the Aristotelian Society*, LVII (1956-1957), 229-253; also *Analysis*, 21 (1960-1961), 43-48.

¹³ J. M. Shorter, "More about Bodily Continuity and Personal Identity," *Analysis*, 22 (1961-1962), 79-85; and Mrs. J. M. R. Jack (unpublished), who requires that this truth be embedded in a causal theory.

¹⁴ *Analysis*, 21 (1960-1961), 44.

Williams applied this principle to a case in which one man is psychologically continuous with the dead Guy Fawkes, and a case in which two men are. His argument was this. If we treat psychological continuity as a sufficient ground for speaking of identity, we shall say that the one man is Guy Fawkes. But we could not say that the two men are, although we should have the same ground. This disobeys the principle. The remedy is to deny that the one man is Guy Fawkes, to insist that sameness of the body is necessary for identity.

Williams' principle can yield a different answer. Suppose we regard psychological continuity as more important than sameness of the body.¹⁵ And suppose that the one man really is psychologically (and causally) continuous with Guy Fawkes. If he is, it would disobey the principle to deny that he is Guy Fawkes, for we have the same important ground as in a normal case of identity. In the case of the two men, we again have the same important ground. So we ought to take the importance from the judgment of identity and attach it directly to this ground. We ought to say, as in Wiggins' case, that each limb of the branching relation is as good as survival. This obeys the principle.

To sum up these remarks: even if psychological continuity is neither logically, nor always in fact, one-one, it can provide a criterion of identity. For this can appeal to the relation of *non-branching* psychological continuity, which is logically one-one.¹⁶

The criterion might be sketched as follows. "*X* and *Y* are the same person if they are psychologically continuous and there is no person who is contemporary with either and psychologically continuous with the other." We should need to explain what we mean by "psychologically continuous" and say how much continuity the criterion requires. We should then, I think, have described a sufficient condition for speaking of identity.¹⁷

We need to say something more. If we admit that psychological

¹⁵ For the reasons given by A. M. Quinton in "The Soul," *Journal of Philosophy*, LIX (1962), 393-409.

¹⁶ Cf. S. Shoemaker, "Persons and Their Pasts," to appear in the *American Philosophical Quarterly*, and "Wiggins on Identity," *Philosophical Review*, LXXIX (1970), 542.

¹⁷ But not a necessary condition, for in the absence of psychological continuity bodily identity might be sufficient.

continuity might not be one-one, we need to say what we ought to do if it were not one-one. Otherwise our account would be open to the objections that it is incomplete and arbitrary.¹⁸

I have suggested that if psychological continuity took a branching form, we ought to speak in a new way, regarding what we describe as having the same significance as identity. This answers these objections.¹⁹

We can now return to our discussion. We have three remaining aims. One is to suggest a sense of “survive” which does not imply identity. Another is to show that most of what matters in survival are relations of degree. A third is to show that none of these relations needs to be described in a way that presupposes identity.

We can take these aims in the reverse order.

III

The most important particular relation is that involved in memory. This is because it is so easy to believe that its description must refer to identity.²⁰ This belief about memory is an important cause of the view that personal identity has a special nature. But it has been well discussed by Shoemaker²¹ and by Wiggins²². So we can be brief.

It may be a logical truth that we can only remember our own experiences. But we can frame a new concept for which this is not a logical truth. Let us call this “*q*-memory.”

¹⁸ Cf. Bernard Williams, “Personal Identity and Individuation,” *Proceedings of the Aristotelian Society*, LVII (1956-1957), 240-241, and *Analysis*, 21 (1960-1961), 44; and also Wiggins, *op. cit.*, p. 38: “if coincidence under [the concept] *f* is to be *genuinely* sufficient we must not withhold identity . . . simply because transitivity is threatened.”

¹⁹ Williams produced another objection to the “psychological criterion,” that it makes it hard to explain the difference between the concepts of identity and exact similarity (*Analysis*, 21 [1960-1961], 48). But if we include the requirement of causal continuity we avoid this objection (and one of those produced by Wiggins in his note 47).

²⁰ Those philosophers who have held this belief, from Butler onward, are too numerous to cite.

²¹ *Op. cit.*

²² In a paper on Butler’s objection to Locke (not yet published).

To sketch a definition²³ I am *q*-remembering an experience if (1) I have a belief about a past experience which seems in itself like a memory belief, (2) someone did have such an experience, and (3) my belief is dependent upon this experience in the same way (whatever that is) in which a memory of an experience is dependent upon it.

According to (1) *q*-memories seem like memories. So I *q*-remember *having* experiences.

This may seem to make *q*-memory presuppose identity. One might say, "My apparent memory of *having* an experience is an apparent memory of *my* having an experience. So how could I *q*-remember my having other people's experiences?"

This objection rests on a mistake. When I seem to remember an experience, I do indeed seem to remember *having* it.²⁴ But it cannot be a part of what I seem to remember about this experience that I, the person who now seems to remember it, am the person who had this experience.²⁵ That I am is something that I automatically assume. (My apparent memories sometimes come to me simply as the belief that *I* had a certain experience.) But it is something that I am justified in assuming only because I do not in fact have *q*-memories of other people's experiences.

Suppose that I did start to have such *q*-memories. If I did, I should cease to assume that my apparent memories must be about my own experiences. I should come to assess an apparent memory by asking two questions: (1) Does it tell me about a past experience? (2) If so, whose?

Moreover (and this is a crucial point) my apparent memories would now come to me as *q*-memories. Consider those of my ap-

²³ I here follow Shoemaker's "quasi-memory." Cf. also Penelhum's "retro-cognition," in his article on "Personal Identity," in the *Encyclopedia of Philosophy*, ed. by Paul Edwards.

²⁴ As Shoemaker put it, I seem to remember the experience "from the inside" (*op. cit.*).

²⁵ This is what so many writers have overlooked. Cf. Thomas Reid: "My memory testifies not only that this was done, but that it was done by me who now remember it" ("Of Identity," in *Essays on the Intellectual Powers of Man*, ed. by A. D. Woozley [London, 1941], p. 203). This mistake is discussed by A. B. Palma in "Memory and Personal Identity," *Australasian Journal of Philosophy*, 42 (1964), 57.

parent memories which do come to me simply as beliefs about my past: for example, "I did that." If I knew that I could *q*-remember other people's experiences, these beliefs would come to me in a more guarded form: for example, "Someone—probably I—did that." I might have to work out who it was.

I have suggested that the concept of *q*-memory is coherent. Wiggins' case provides an illustration. The resulting people, in his case, both have apparent memories of living the life of the original person. If they agree that they are not this person, they will have to regard these as only *q*-memories. And when they are asked a question like "Have you heard this music before?" they might have to answer "I am sure that I *q*-remember hearing it. But I am not sure whether I remember hearing it. I am not sure whether it was I who heard it, or the original person."

We can next point out that on our definition every memory is also a *q*-memory. Memories are, simply, *q*-memories of one's own experiences. Since this is so, we could afford now to drop the concept of memory and use in its place the wider concept *q*-memory. If we did, we should describe the relation between an experience and what we now call a "memory" of this experience in a way which does not presuppose that they are had by the same person.²⁶

This way of describing this relation has certain merits. It vindicates the "memory criterion" of personal identity against the charge of circularity.²⁷ And it might, I think, help with the problem of other minds.

²⁶ It is not logically necessary that we only *q*-remember our own experiences. But it might be necessary on other grounds. This possibility is intriguingly explored by Shoemaker in his "Persons and Their Pasts" (*op. cit.*). He shows that *q*-memories can provide a knowledge of the world only if the observations which are *q*-remembered trace out fairly continuous spatiotemporal paths. If the observations which are *q*-remembered traced out a network of frequently interlocking paths, they could not, I think, be usefully ascribed to persisting observers, but would have to be referred to in some more complex way. But in fact the observations which are *q*-remembered trace out single and separate paths; so we can ascribe them to ourselves. In other words, it is epistemologically necessary that the observations which are *q*-remembered should satisfy a certain general condition, one particular form of which allows them to be usefully self-ascribed.

²⁷ Cf. Wiggins' paper on Butler's objection to Locke.

But we must move on. We can next take the relation between an intention and a later action. It may be a logical truth that we can intend to perform only our own actions. But intentions can be redescribed as *q*-intentions. And one person could *q*-intend to perform another person's actions.

Wiggins' case again provides the illustration. We are supposing that neither of the resulting people is the original person. If so, we shall have to agree that the original person can, before the operation, *q*-intend to perform their actions. He might, for example, *q*-intend, as one of them, to continue his present career, and, as the other, to try something new.²⁸ (I say "*q*-intend as one of them" because the phrase "*q*-intend *that* one of them" would not convey the directness of the relation which is involved. If I intend that someone else should do something, I cannot get him to do it simply by forming this intention. But if I am the original person, and he is one of the resulting people, I can.)

The phrase "*q*-intend as one of them" reminds us that we need a sense in which one person can survive as two. But we can first point out that the concepts of *q*-memory and *q*-intention give us our model for the others that we need: thus, a man who can *q*-remember could *q*-recognize, and be a *q*-witness of, what he has never seen; and a man who can *q*-intend could have *q*-ambitions, make *q*-promises, and be *q*-responsible for.

To put this claim in general terms: many different relations are included within, or are a consequence of, psychological continuity. We describe these relations in ways which presuppose the continued existence of one person. But we could describe them in new ways which do not.

This suggests a bolder claim. It might be possible to think of experiences in a wholly "impersonal" way. I shall not develop

²⁸ There are complications here. He could form *divergent q*-intentions only if he could distinguish, in advance, between the resulting people (e.g., as "the left-hander" and "the right-hander"). And he could be confident that such divergent *q*-intentions would be carried out only if he had reason to believe that neither of the resulting people would change their (inherited) mind. Suppose he was torn between duty and desire. He could not solve this dilemma by *q*-intending, as one of the resulting people, to do his duty, and, as the other, to do what he desires. For the one he *q*-intended to do his duty would face the same dilemma.

this claim here. What I shall try to describe is a way of thinking of our own identity through time which is more flexible, and less misleading, than the way in which we now think.

This way of thinking will allow for a sense in which one person can survive as two. A more important feature is that it treats survival as a matter of degree.

IV

We must first show the need for this second feature. I shall use two imaginary examples.

The first is the converse of Wiggins' case: fusion. Just as division serves to show that what matters in survival need not be one-one, so fusion serves to show that it can be a question of degree.

Physically, fusion is easy to describe. Two people come together. While they are unconscious, their two bodies grow into one. One person then wakes up.

The psychology of fusion is more complex. One detail we have already dealt with in the case of the exam. When my mind was reunited, I remembered just having thought out two calculations. The one person who results from a fusion can, similarly, *q*-remember living the lives of the two original people. None of their *q*-memories need be lost.

But some things must be lost. For any two people who fuse together will have different characteristics, different desires, and different intentions. How can these be combined?

We might suggest the following. Some of these will be compatible. These can coexist in the one resulting person. Some will be incompatible. These, if of equal strength, can cancel out, and if of different strengths, the stronger can be made weaker. And all these effects might be predictable.

To give examples—first, of compatibility: I like Palladio and intend to visit Venice. I am about to fuse with a person who likes Giotto and intends to visit Padua. I can know that the one person we shall become will have both tastes and both intentions. Second, of incompatibility: I hate red hair, and always vote Labour. The other person loves red hair, and always votes Conservative. I can

PERSONAL IDENTITY

know that the one person we shall become will be indifferent to red hair, and a floating voter.

If we were about to undergo a fusion of this kind, would we regard it as death?

Some of us might. This is less absurd than regarding division as death. For after my division the two resulting people will be in every way like me, while after my fusion the one resulting person will not be wholly similar. This makes it easier to say, when faced with fusion, "I shall not survive," thus continuing to regard survival as a matter of all-or-nothing.

This reaction is less absurd. But here are two analogies which tell against it.

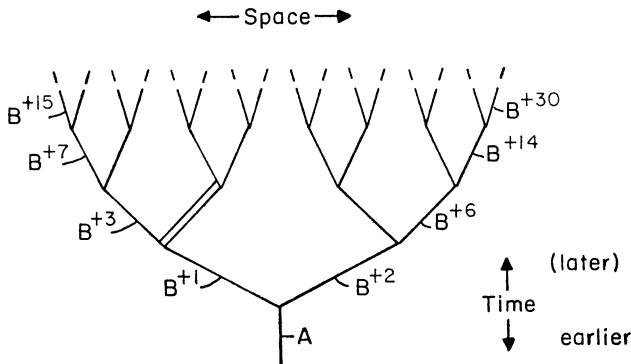
First, fusion would involve the changing of some of our characteristics and some of our desires. But only the very self-satisfied would think of this as death. Many people welcome treatments with these effects.

Second, someone who is about to fuse can have, beforehand, just as much "intentional control" over the actions of the resulting individual as someone who is about to marry can have, beforehand, over the actions of the resulting couple. And the choice of a partner for fusion can be just as well considered as the choice of a marriage partner. The two original people can make sure (perhaps by "trial fusion") that they do have compatible characters, desires, and intentions.

I have suggested that fusion, while not clearly survival, is not clearly failure to survive, and hence that what matters in survival can have degrees.

To reinforce this claim we can now turn to a second example. This is provided by certain imaginary beings. These beings are just like ourselves except that they reproduce by a process of natural division.

We can illustrate the histories of these imagined beings with the aid of a diagram. (This is given on the next page.) The lines on the diagram represent the spatiotemporal paths which would be traced out by the bodies of these beings. We can call each single line (like the double line) a "branch"; and we can call the whole structure a "tree." And let us suppose that each "branch" corresponds to what is thought of as the life of one individual.



These individuals are referred to as “A,” “B + 1,” and so forth.

Now, each single division is an instance of Wiggins’ case. So *A*’s relation to both *B + 1* and *B + 2* is just as good as survival. But what of *A*’s relation to *B+30*?

I said earlier that what matters in survival could be provisionally referred to as “psychological continuity.” I must now distinguish this relation from another, which I shall call “psychological connectedness.”

Let us say that the relation between a *q*-memory and the experience *q*-remembered is a “direct” relation. Another “direct” relation is that which holds between a *q*-intention and the *q*-intended action. A third is that which holds between different expressions of some lasting *q*-characteristic.

“Psychological connectedness,” as I define it, requires the holding of these direct psychological relations. “Connectedness” is not transitive, since these relations are not transitive. Thus, if *X* *q*-remembers most of *Y*’s life, and *Y* *q*-remembers most of *Z*’s life, it does not follow that *X* *q*-remembers most of *Z*’s life. And if *X* carries out the *q*-intentions of *Y*, and *Y* carries out the *q*-intentions of *Z*, it does not follow that *X* carries out the *q*-intentions of *Z*.

“Psychological continuity,” in contrast, only requires overlapping chains of direct psychological relations. So “continuity” is transitive.

To return to our diagram. *A* is psychologically continuous with *B + 30*. There are between the two continuous chains of overlap-

ping relations. Thus, *A* has *q*-intentional control over *B* + 2, *B* + 2 has *q*-intentional control over *B* + 6, and so on up to *B* + 30. Or *B* + 30 can *q*-remember the life of *B* + 14, *B* + 14 can *q*-remember the life of *B* + 6, and so on back to *A*.²⁹

A, however, need *not* be psychologically connected to *B* + 30. Connectedness requires direct relations. And if these beings are like us, *A* cannot stand in such relations to every individual in his indefinitely long “tree.” *Q*-memories will weaken with the passage of time, and then fade away. *Q*-ambitions, once fulfilled, will be replaced by others. *Q*-characteristics will gradually change. In general, *A* stands in fewer and fewer direct psychological relations to an individual in his “tree” the more remote that individual is. And if the individual is (like *B* + 30) sufficiently remote, there may be between the two *no* direct psychological relations.

Now that we have distinguished the general relations of psychological continuity and psychological connectedness, I suggest that connectedness is a more important element in survival. As a claim about our own survival, this would need more arguments than I have space to give. But it seems clearly true for my imagined beings. *A* is as close psychologically to *B* + 1 as I today am to myself tomorrow. *A* is as distant from *B* + 30 as I am from my great-great-grandson.

Even if connectedness is not more important than continuity, the fact that one of these is a relation of degree is enough to show that what matters in survival can have degrees. And in any case the two relations are quite different. So our imagined beings would need a way of thinking in which this difference is recognized.

V

What I propose is this.

First, *A* can think of any individual, anywhere in his “tree,” as “a descendant self.” This phrase implies psychological continuity. Similarly, any later individual can think of any earlier

²⁹ The chain of continuity must run in one direction of time. *B* + 2 is not, in the sense I intend, psychologically continuous with *B* + 1.

individual on the single path³⁰ which connects him to *A* as “an ancestral self.”

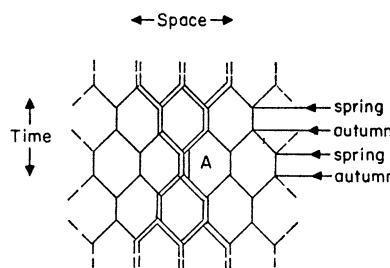
Since psychological continuity is transitive, “being an ancestral self of” and “being a descendant self of” are also transitive.

To imply psychological connectedness I suggest the phrases “one of my future selves” and “one of my past selves.”

These are the phrases with which we can describe Wiggins’ case. For having past and future selves is, what we needed, a way of continuing to exist which does not imply identity through time. The original person does, in this sense, survive Wiggins’ operation: the two resulting people are his later selves. And they can each refer to him as “my past self.” (They can share a past self without being the same self as each other.)

Since psychological connectedness is not transitive, and is a matter of degree, the relations “being a past self of” and “being a future self of” should themselves be treated as relations of degree. We allow for this series of descriptions: “my most recent self,” “one of my earlier selves,” “one of my distant selves,” “hardly one of *my* past selves (I can only *q*-remember a few of his experiences),” and, finally, “not in any way one of *my* past selves—just an ancestral self.”

This way of thinking would clearly suit our first imagined beings. But let us now turn to a second kind of being. These reproduce by fusion as well as by division.³¹ And let us suppose that they fuse every autumn and divide every spring. This yields the following diagram:



³⁰ Cf. David Wiggins, *op.cit.*

³¹ Cf. Sydney Shoemaker in “Persons and Their Pasts,” *op.cit.*

PERSONAL IDENTITY

If *A* is the individual whose life is represented by the three-lined “branch,” the two-lined “tree” represents those lives which are psychologically continuous with *A*’s life. (It can be seen that each individual has his own “tree,” which overlaps with many others.)

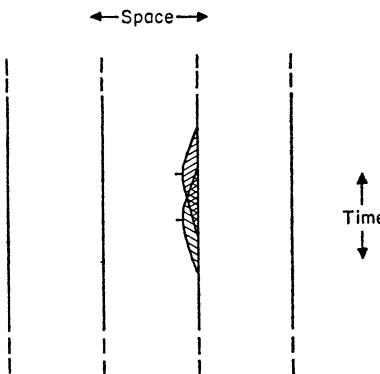
For the imagined beings in this second world, the phrases “an ancestral self” and “a descendant self” would cover too much to be of much use. (There may well be pairs of dates such that every individual who ever lived before the first date was an ancestral self of every individual who ever will live after the second date.) Conversely, since the lives of each individual last for only half a year, the word “I” would cover too little to do all of the work which it does for us. So part of this work would have to be done, for these second beings, by talk about past and future selves.

We can now point out a theoretical flaw in our proposed way of thinking. The phrase “a past self of” implies psychological connectedness. Being a past self of is treated as a relation of degree, so that this phrase can be used to imply the varying degrees of psychological connectedness. But this phrase can imply only the degrees of connectedness between different lives. It cannot be used within a single life. And our way of delimiting successive lives does not refer to the degrees of psychological connectedness. Hence there is no guarantee that this phrase, “a past self of,” could be used whenever it was needed. There is no guarantee that psychological connectedness will not vary in degree within a single life.

This flaw would not concern our imagined beings. For they divide and unite so frequently, and their lives are in consequence so short, that within a single life psychological connectedness would always stand at a maximum.

But let us look, finally, at a third kind of being.

In this world there is neither division nor union. There are a number of everlasting bodies, which gradually change in appearance. And direct psychological relations, as before, hold only over limited periods of time. This can be illustrated with a third diagram (given on the next page). In this diagram the two shadings represent the degrees of psychological connectedness to their two central points.



These beings could not use the way of thinking that we have proposed. Since there is no branching of psychological continuity, they would have to regard themselves as immortal. It might be said that this is what they are. But there is, I suggest, a better description.

Our beings would have one reason for thinking of themselves as immortal. The parts of each "line" are all psychologically continuous. But the parts of each "line" are not all psychologically connected. Direct psychological relations hold only between those parts which are close to each other in time. This gives our beings a reason for *not* thinking of each "line" as corresponding to one single life. For if they did, they would have no way of implying these direct relations. When a speaker says, for example, "I spent a period doing such and such," his hearers would not be entitled to assume that the speaker has any memories of this period, that his character then and now are in any way similar, that he is now carrying out any of the plans or intentions which he then had, and so forth. Because the word "I" would carry none of these implications, it would not have for these "immortal" beings the usefulness which it has for us.³²

To gain a better way of thinking, we must revise the way of thinking that we proposed above. The revision is this. The distinction between successive selves can be made by reference, not to the branching of psychological continuity, but to the degrees

³² Cf. Austin Duncan Jones, "Man's Mortality," *Analysis*, 28 (1967-1968), 65-70.

of psychological connectedness. Since this connectedness is a matter of degree, the drawing of these distinctions can be left to the choice of the speaker and be allowed to vary from context to context.

On this way of thinking, the word “I” can be used to imply the greatest degree of psychological connectedness. When the connections are reduced, when there has been any marked change of character or style of life, or any marked loss of memory, our imagined beings would say, “It was not I who did that, but an earlier self.” They could then describe in what ways, and to what degree, they are related to this earlier self.

This revised way of thinking would suit not only our “immortal” beings. It is also the way in which we ourselves could think about our lives. And it is, I suggest, surprisingly natural.

One of its features, the distinction between successive selves, has already been used by several writers. To give an example, from Proust: “we are incapable, while we are in love, of acting as fit predecessors of the next persons who, when we are in love no longer, we shall presently have become. . . .”³³

Although Proust distinguished between successive selves, he still thought of one person as being these different selves. This we would not do on the way of thinking that I propose. If I say, “It will not be me, but one of my future selves,” I do not imply that I will be that future self. He is one of my later selves, and I am one of his earlier selves. There is no underlying person who we both are.

To point out another feature of this way of thinking. When I say, “There is no person who we both are,” I am only giving my decision. Another person could say, “It will be you,” thus deciding differently. There is no question of either of these decisions being a mistake. Whether to say “I,” or “one of my future selves,” or “a descendant self” is entirely a matter of choice. The matter of fact, which must be agreed, is only whether the disjunction applies. (The question “Are *X* and *Y* the same person?” thus becomes “Is *X* at least an ancestral [or descendant] self of *Y*?”)

³³ *Within a Budding Grove* (London, 1949), I, 226 (my own translation).

VI

I have tried to show that what matters in the continued existence of a person are, for the most part, relations of degree. And I have proposed a way of thinking in which this would be recognized.

I shall end by suggesting two consequences and asking one question.

It is sometimes thought to be especially rational to act in our own best interests. But I suggest that the principle of self-interest has no force. There are only two genuine competitors in this particular field. One is the principle of biased rationality: do what will best achieve what you actually want. The other is the principle of impartiality: do what is in the best interests of everyone concerned.

The apparent force of the principle of self-interest derives, I think, from these two other principles.

The principle of self-interest is normally supported by the principle of biased rationality. This is because most people care about their own future interests.

Suppose that this prop is lacking. Suppose that a man does not care what happens to him in, say, the more distant future. To such a man, the principle of self-interest can only be propped up by an appeal to the principle of impartiality. We must say, "Even if you don't care, you ought to take what happens to you then equally into account." But for this, as a special claim, there seem to me no good arguments. It can only be supported as part of the general claim, "You ought to take what happens to everyone equally into account."³⁴

The special claim tells a man to grant an *equal* weight to all the parts of his future. The argument for this can only be that all the parts of his future are *equally* parts of *his* future. This is true. But it is a truth too superficial to bear the weight of the argument. (To give an analogy: The unity of a nation is, in its nature, a matter of degree. It is therefore only a superficial truth

³⁴ Cf. Thomas Nagel's *The Possibility of Altruism* (Oxford, 1970), in which the special claim is in effect defended as part of the general claim.

PERSONAL IDENTITY

that all of a man's compatriots are *equally* his compatriots. This truth cannot support a good argument for nationalism.)³⁵

I have suggested that the principle of self-interest has no strength of its own. If this is so, there is no special problem in the fact that what we ought to do can be against our interests. There is only the general problem that it may not be what we want to do.

The second consequence which I shall mention is implied in the first. Egoism, the fear not of near but of distant death, the regret that so much of one's *only* life should have gone by—these are not, I think, wholly natural or instinctive. They are all strengthened by the beliefs about personal identity which I have been attacking. If we give up these beliefs, they should be weakened.

My final question is this. These emotions are bad, and if we weaken them we gain. But can we achieve this gain without, say, also weakening loyalty to, or love of, other particular selves? As Hume warned, the "refined reflections which philosophy suggests . . . cannot diminish . . . our vicious passions . . . without diminishing . . . such as are virtuous. They are . . . applicable to all our affections. In vain do we hope to direct their influence only to one side."³⁶

That hope *is* vain. But Hume had another: that more of what is bad depends upon false belief. This is also my hope.

DEREK PARFIT

All Souls College, Oxford

³⁵ The unity of a nation we seldom take for more than what it is. This is partly because we often think of nations, not as units, but in a more complex way. If we thought of ourselves in the way that I proposed, we might be less likely to take our own identity for more than what it is. We are, for example, sometimes told, "It is irrational to act against your own interests. After all, it will be *you* who will regret it." To this we could reply, "No, not me. Not even one of my future selves. Just a descendant self."

³⁶ "The Sceptic," in "Essays Moral, Political and Literary," *Hume's Moral and Political Philosophy* (New York, 1959), p. 349.