RAYMOND MARTIN

IDENTITY'S CRISIS

(Received in revised form 14 March, 1987)

Occasionally, but not often, philosophy stumbles onto something genuinely new — a new problem, or a subtle change in an old problem that brings a new set of issues into focus. When this happens circumstances are ripe for transformations not just of what we believe, but also of what we think is worth considering, and how we think we ought to proceed.

I believe that something genuinely new has happened during the last two decades in the philosophical debate over personal identity, and that as a consequence the current debate has lost its way — is in the throes, one might say, of a kind of identity crisis. So far no one seems to have noticed that there is a crisis, let alone suggested how to solve it. The point of the present paper is to argue that there is a crisis, to diagnosis its source, and to raise the question of how we should proceed.

I

What's genuinely new in the debate over personal identity is that while it used to be assumed routinely that the preservation of identity is primarily what matters in survival, now several important philosophers have argued convincingly that other things matter more than identity. According to the philosophers who hold this new view — a group which includes Robert Nozick, Sydney Shoemaker and Derek Parfit — identity was only mistakenly thought to be primarily what matters in survival, and this because identity is all but invariably correlated with other characteristics of persons which *are* primarily what matters in survival.[1]

This shift in focus from a concern with identity to a concern with what matters in survival is the direct result of the attempt to come to terms with so-called fission cases, first introduced into the contem-

Philosophical Studies **53** (1988) 295–307.

porary discussion by Wiggins in 1967.[2] Consider, for instance, the much discussed case of a person whose brain is divided, each half of which is then transplanted into its own new body. Suppose that before division each half of the donor's brain is psychologically redundant, so that immediately after division each of the resulting persons has all and only the psychological characteristics of the donor — the same character traits, the same beliefs, the same tastes, and so on. And suppose that the resulting persons have bodies which are exactly similar to each other and to the donor's body.

Many philosophers believe that in such a case neither of the resulting persons is identical to the donor.[3] The reason is that identity is a transitive relationship, and the resulting persons, at least once they begin to lead independent lives, are not identical to each other. Since each resulting person has an equally good claim to be identical to the donor, neither can be identical to the donor unless both are. It follows that neither is identical to the donor.

The fission operation does not preserve the donor's identity. Even so, the outcome of the operation may be better, from the donor's point of view, than all of the available alternatives which include the preservation of identity. Shoemaker provides a convincing example:

> Suppose that half of my brain and all of the rest of my body are ridden with cancer, and that my only hope for survival is for my healthy half-brain to be transplanted to another body. There are two transplantation procedures available. The first, which is inexpensive and safe (so far as the prospects of the recipient are concerned) involves first transplanting the healthy hemisphere and then destroying (or allowing to die) the diseased hemisphere that remains. The other, which is expensive and risky (the transplant may not take, or it may produce a psychologically damaged person) involves first destroying the diseased hemisphere and then transplanting the other. Which shall I choose? Notice that if I choose the first procedure there will be, for a short while, two persons psychologically continuous with the original person (me), and therefore that on the non-branching psychological continuity theory the recipient of the healthy hemisphere cannot count as me. If I choose the second procedure, on the other hand, then at no point will the recipient (the post-operative possessor of the healthy hemisphere) have a "competitor" for the status of being me, so it seems that he can count as me (if the transplantation takes). Should I therefore choose the expensive and risky procedure? This seems absurd. The thing to do is to choose the first procedure, even though (I think) it guarantees that the transplant recipient will not be me.[4]

Such examples, by separating identity from other characteristics that normally accompany it and give it its inflated sense of importance, have convinced a number of philosophers that the preservation of identity is not primarily what matters in survival.

If identity is not primarily what matters in survival, then what does primarily matter? Well, aside from identity, what's left? For most philosophers what's left that might matter is bodily continuity, in whole or in part, and/or psychological continuity, in whole or in part. I don't wish in the present paper to try to answer the question of what matters in survival, but rather to provide a motivation for considering three methodologically prior questions: First, what question ought we to be asking about what matters in survival? Second, why ought we to be asking it? And finally, how ought we to be trying to answer it?

II

When it used to be generally assumed that identity *is* primarily what matters in survival, it was relatively clear what was the central philosophical question about personal identity and why this question was important. There was also widespread, if not quite universal, agreement about how this question should be investigated. The central question was that of specifying the conditions under which identity is preserved. The main reason this question was important was because the preservation of identity was primarily what matters in survival. And the most common way philosophers investigated this central question was by generating hypothetical puzzle cases to test various putative analyses of personal identity.

Now that it is no longer generally assumed that identity is primarily what matters in survival, it is no longer clear what the central philosophical questions are, or why they are important, or how they should be considered. Yet the debate generally goes on as if these issues were clear, as if there were no crisis, and in spite of many symptoms of a new found confusion. One such symptom is that it can be difficult now even to say *meaningfully* what the main questions are that should be considered. Another is that it is often unclear whether these new questions are the most important questions to consider. Finally, it often seems that the new questions are being discussed in a philosophical framework more appropriate to the old than to the new questions.

I want now to illustrate these claims by considering Derek Parfit's discussion of identity and what matters in survival in *Reasons and Persons* (Oxford, 1984). I've chosen Parfit's account because he more

than anyone else has stressed the importance of the difference between personal identity and what matters in survival and has tried to formulate the new questions that we should be asking, and because I believe that *Reasons and Persons*, despite serious flaws, contains the most philosophically rewarding account of personal identity and of what matters in survival that currently exists.

Parfit's view, in brief, is that personal identity consists in non-branching psychological connectedness and continuity. There is psychological connectedness when a person remembers doing or experiencing something that someone earlier did or experienced, or when a person's intention to do something leads to its later being done, or when psychological states, such as beliefs and desires, persist over time. There is psychological continuity when there is a sufficiently strong overlapping chain of psychological connectedness.[5] What matters in survival is not personal identity but rather psychological connectedness and/or continuity (hereafter, simply "psychological continuity") with the right kind of cause. And the right kind of cause can be any kind of cause.

Parfit argues that those who claim that *physical* continuity is necessary to identity and primarily what matters in survival cannot justifiably require the continuity of the whole body, but *at most* only that enough of the brain continues to exist to insure the preservation of psychological continuity. But they cannot require even this much. The only justification for singling out the brain is that it is the brain that normally preserves psychological continuity. Otherwise, the brain would be no more important than any other organ. But if something else reliably preserved psychological continuity, then it would be irrational to insist that even partial brain continuity is necessary for identity or primarily what matters in survival.

We can imagine a futuristic device — Parfit imagines a futuristic traveling device which he calls a "Teletransporter" — that reliably preserves psychological continuity without preserving physical or bodily continuity, not even continuity of any part of the brain.[6] Since such a device is conceptually possible, it is irrational to insist that brain continuity is necessary for identity or primarily what matters in survival.

Parfit argues that it is not even necessary that whatever preserves psychological continuity preserves it reliably: it is the effect that

matters, not its cause. Imagine, for instance, that having just used the Teletransporter to travel successfully to Mars, you learn that the Teletransporter has been malfunctioning lately. Had you known that before the trip, it would have been rational to refuse to go. Once on Mars, the unreliability of the Teletransporter doesn't matter to the prospects of your surviving the trip. You are psychologically continuous with your pre-trip self, and that is all that matters. The reliability or unreliability of what caused you to survive the trip has no bearing on whether you did survive it.

Parfit, in his discussion of what matters in survival, sometimes asks "What *does* matter?", but more often asks "What *rationally ought to* matter?" He clearly regards the latter question as more important. But one problem with Parfit's account is that it is doubtful whether the question *that Parfit wants to ask* about what rationally ought to matter in survival can even be meaningfully expressed.[7] And another problem is that even if Parfit's question can be asked, it is doubtful that it has the kind of importance that Parfit attributes to it.

Consider first the difficulties Parfit has expressing the question of what rationally ought to matter in survival:

> As the case of My Division shows personal identity is not what matters. It is merely true that, in most cases, personal identity coincides with what matters. [1] What does matter *in the way in which* personal identity is, mistakenly, thought to matter? [2] What is it rational to care about in our concern about our own future?
>
> The question can be restated. [3] Assume, for simplicity, that it could be rational to be concerned only about one's own self-interest. Suppose that I am an Egoist, and that I could be related in one of several ways to some resulting person. What is the relation that would justify egoistic concern about this, resulting person? If the rest of this person's life will be well worth living, in what way should I want to be related to this person? If the rest of his life will be much worse than nothing, in what way should I want *not* to be related to this person? In short, what is the relation that, for an Egoist, should fundamentally matter? This relation will also be what, for all of us, should fundamentally matter, in our concern for our own future. But since we may be concerned about the fate of the resulting person, *whatever* his relation is to us, it is clearest to ask what, for an Egoist, should matter.[8]

Parfit's "My Division" case is a fission case relevantly similar to the one sketched at the beginning of this paper. Parfit intends his question of what rationally ought to matter in survival to include within its scope not only standard cases where identity and what matters in survival go hand in hand, but also cases, like his My Division case, where identity is *not* preserved, but where what matters in survival *is* preserved. My

question is this: Given Parfit's interpretation of the fission cases, is the question that he wants to ask meaningful?

Parfit has given us several versions of his question. Some, or all, of these may appear to be meaningful formulations of the question that he wants to ask. Let's see. Consider his first formulation [1]: What does matter *in the way in which* personal identity is, mistakenly, thought to matter? It is clear from his subsequent formulations that Parfit is asking not what *does* matter, but rather what *rationally ought to* matter. What is not clear is what is meant by the expression, "*in the way in which* personal identity is thought to matter." Before we can answer, we have to know *which way* is the way in which personal identity is thought to matter. The most straightforward interpretation of *which way* identity is thought to matter is that it is thought to matter in the "identity-way," that is, personal identity matters in the way in which preserving ourselves as the same persons that we now are matters.

But on this interpretation of Parfit's question, it either makes no sense or else it yields what for Parfit is the wrong answer to his question, at least when it is asked of fission cases like the ones considered. On this interpretation of his question, Parfit is asking us what matters in the way in which preserving ourselves as the same persons that we now are matters in a situation in which we will not be preserved as the same persons that we now are. If this question makes sense, the answer to it would seem to be "nothing," that is, nothing matters in *that* way. But this is not the answer that Parfit wants.

Consider then Parfit's second formulation [2] of his question: What is it rational to care about in our concern about our own future? The problem here is that the question, when asked of the fission cases, is best regarded as elliptical. The full question is this: What is it rational to care about in our concern about *our own* future in a situation in which we will not have "a future," that is, in which there will not be any continuation of *our own* future? Again, the answer seems to be: "nothing." And again, this is not the answer that Parfit wants.

Consider Parfit's third formulation [3] of his question, again as it applies to fission cases in which, on Parfit's view, identity is lost but what matters in survival is preserved. Parfit's question would seem to be this: How should a rational Egoist want to be related to some person who will be around after the Egoist is gone and whose life will be worth

living? How we answer this question depends crucially on what we understand by "rational Egoist"? Presumably, whatever else a rational Egoist may be, he is someone who cares only about *his own self-interest.* So, when asked of the fission cases, the question is: "What should a rational Egoist who cares only about his own self-interest care about after he no longer has any interests of his own?" Again, the answer, arguably, is "nothing," since nothing that happens after the Egoist ceases to exist can affect the Egoist's self-interest.

But there may be another answer. It is not obvious that nothing that happens after the Egoist ceases to exist can affect the Egoists self-interest. After all, someone might argue, a rational Egoist might while she is alive care about many things that will happen only after she is gone. For instance, she might care about how her property will be dispersed, about whether her children will be happy, and so on. Will this save Parfit's question? I don't think so. If we interpret the possible cares of a rational Egoist this inclusively, the floodgates are open. A rational Egoist who cares only about her own self-interest, on this interpretation of what it means to care only about one's own self-interest, might care about virtually anything.

Why is it so difficult to ask the question Parfit wants to ask? For this reason: Parfit is trying to ask a *qualified* version of the question of what sort of remains a person should care about leaving after he no longer exists, and the obvious ways to make the intended qualification imply that the person who *will no longer continue* to exist *will continue* to exist. But there is no way to answer a question so qualified and no obvious way to reformulate it. The only way to ask the question that Parfit wants to ask seems to be to introduce the notion of "*self*-interest," or something like it, into a context from which it has been explicitly excluded.

Perhaps there is a non-obvious way to ask the question that Parfit wants to ask. After all, the only reason Parfit gets into difficulty formulating his question is that the donor in certain fission examples is faced with a situation in which there will be several replicas of herself anyone of which would have been her but for the fact that the others exist. Why then can't Parfit simply ask his question this way: What should a rational egoist care about in caring about what would have preserved her identity but for the fact of fission?

While this question is not entirely free of difficulties, it may seem better than the ones that Parfit asks. There is an interpretation of this question according to which it both makes sense and doesn't yield the unwanted answer, "nothing." But on this interpretation the question will yield for Parfit the same answer as the question of what preserves personal identity. And that is not the answer that Parfit wants. On Parfit's view *non-branching* psychological continuity is what preserves personal identity. The answer that Parfit wants to the question he is trying to ask is not *non-branching* psychological continuity, but rather simply psychological continuity, whether branching or not.[9]

III

Let's assume that Parfit *can* ask the question about what rationally ought to matter in survival that he wants to ask. That would still not solve the main problems that concern us. The difficulty in asking a meaningful question is only part of the problem. Another part is that of showing that the question one asks is among the most important questions to ask about what matters in survival.

One way to show that the question of what rationally ought to matter in survival is an important question is to show that it has an important answer. Parfit tries to do this. He tries to show that only psychological continuity ought to matter in survival. His answer implies that *physical* continuity *ought not* to matter in survival — that it is *irrational* to regard physical continuity as an important part of what matters in survival. If Parfit could show that physical continuity ought not to matter in survival, then the question of what ought to matter in survival would be an important question, for it would have a startling answer with sweeping philosophical implications. But can Parfit show this?[10]

Parfit has a Socratic view of the relative values of the mind and the body. The mind counts for virtually everything. The body has only instrumental value, and *it has that only as the vehicle for psychological continuity.* Parfit says that physical continuity "is the least important element in a person's continued existence."[11]

Parfit concedes that it might not be irrational to care *a little* about the physical continuity of one's body, just as one might for sentimental reasons prefer his original wedding ring to a new ring that is exactly

similar. But he never explains why a little bit of sentimental attachment isn't irrational — perhaps just a little bit irrational. Nor does he explain why if it is not irrational to be *a little bit* sentimental about something as relatively insignificant as a wedding ring, it may not be irrational to be *very* sentimental about one's body! The ring, after all, is an external object and has only symbolic significance. One's body, on the other hand, is the actual — not merely symbolic — vehicle for virtually all that matters in life.

If it is not irrational to be sentimental about one's wedding ring, it must be because of the ring's historical properties. Your wedding ring, let's suppose, has an intimate association with events now past that had great positive value in and for your life. The ring may help you to feel psychologically close to those events, and that feeling of closeness may be important to you. Hence, the ring may have great instrumental value. But if a mere wedding ring can have great instrumental value, over and above its value as a vehicle for psychological continuity, why not also, and even more so, one's own body?

The word "sentimental" has pejorative connotations it is best to avoid. The important question is not whether it might be rational to be very sentimental about one's body, but rather whether it might be rational to put a high value on one's body in circumstances where there are two guarantees: that as soon as one loses one's current body, one acquires an exact replica; and that only one replica exists at a time.

Parfit's answer is that it is irrational in such circumstances to put a high value on one's current body — an exact replica is as good as the original. So, for instance, a person who chooses not to use the equally reliable Teletransporter to travel to Mars, but chooses instead to pay much more for a conventional space trip, is irrational. But Parfit's defense of this claim assumes that the only reason a person could have for preferring her body to a replica is her belief that her body is essential to her identity and hence that the replica would be somebody else. Parfit argues that this is not a good reason for preferring one's body to a replica. But even if Parfit is right about this, his defense fails because there is another reason a person could have for preferring her body to a replica.

We saw that a wedding ring may help someone to feel psychologically closer to events now past that had great positive value in and

for his life. Hence the ring may have great instrumental value, over and above its value as a vehicle for Parfitian psychological continuity. But the ring example is not an exceptional case. It illustrates a general and pervasive value phenomenon: we tend to strongly prefer originals to replicas. Consider, for instance, our attitudes toward important historical artifacts, such as Galileo's telescope, great works of art, such as the Pieta of Michaelangelo that is in the Vatican, and certain highly valued objects in nature, such as the Sequoia stands in California made up of majestic trees that may be the oldest living things on earth. Would you value replicas, made yesterday, of each of these as much as you would value the originals? Set aside the greater resale and prestige value of the originals. Suppose, for instance, that no one would know you had or experienced any of these things except yourself. Just considering how it would affect you directly, would you value the replicas as much as the originals?

I find when I consider such questions that I would value the originals much more, for reasons similar to those that emerged in the ring example. The originals help me to feel psychologically close to the past or to certain past events in ways that the replicas could not, at least if I knew they were replicas. And that feeling of closeness is important to me. Hence, the originals have great instrumental value that the replicas would lack, and for reasons that have nothing to do with false metaphysics.

So also with our bodies. To treat our bodies as mere vehicles for Parfitian psychological continuity is to ignore other important instrumental values that our bodies may have. Some may value their bodies solely because they believe their bodies are essential to preserving their identities. Many will value their bodies because their bodies have been and are the vehicle for virtually all that has been significant in their lives. Thus, physical continuity can matter, and for perfectly acceptable reasons, even if it is not necessary for identity. So, Parfit is mistaken.[12]

The conclusion just reached allows us to see important middle ground between the views of physical continuity theorists and Parfitian-style psychological continuity theorists. The physical continuity theorists have traditionally argued that physical continuity is essential to identity and thus an important part of what matters in survival. The psychological-continuity theorists argue that only non-branching, psychological

continuity — and not physical continuity — is essential to identity. They sometimes also argue, as does Parfit, that only psychological continuity matters in survival. But one can accept the *psychological*-continuity theorist's claim that only non-branching *psychological* continuity is essential to *identity*, and also the *physical*-continuity theorist's claim that *physical* continuity *can matter importantly in survival.* This middle position explains why attitudes that some of us have about what matters in survival that are irrational on the view of a psychological-continuity theorist such as Parfit are in fact perfectly rational.

If the argument I have just given is correct, then the question of what *ought to matter* in survival is probably not as important a question as Parfit thinks it is. If Parfit could have shown that the only thing that ought to matter in survival is psychological continuity, then the question of what ought to matter in survival would have been quite important. There would have been a nice symmetry and economy to the package Parfit was presenting us: the very thing that gave identity its inflated sense of importance in the old days when it was generally agreed that identity was what mattered in survival, namely psychological continuity, is also the only thing that rationally *ought* to matter in survival. That's basically what Parfit tried to show, and if he had succeeded, it would have conferred a kind of retrospective importance on the question of what *ought* to matter in survival. But if, as I've suggested, there are few interesting *a priori* or conceptual constraints on what rationally ought to matter in survival, then the question of what rationally ought to matter in survival is no longer so enticing. The part of Parfit's answer to the question of what rationally ought to matter in survival that is both interesting *and* correct may just boil down to the admonition not to let things matter for the wrong metaphysical reasons.

IV

We still lack an appropriate question to ask about what matters in survival. The fission cases provide the basis for a convincing argument that identity is not, at least in many situations, what primarily matters in survival. The question remains, "What does matter?" We've seen that if we try to convert this question into the question of what *rationally ought to matter*, then an appropriately restricted version of this question

is difficult even to formulate meaningfully and perhaps not such an important question anyway.

So it may seem that we should just stick with the question of what *does* matter in survival. But there are problems also in adopting this question as a question of central philosophical importance. One problem is that it is too broad because it lacks the kind of qualification that Parfit was trying to provide. We do, or could, leave all sorts of remains that matter, from children to books to successive stages of ourselves, to successive stages of what would be ourselves but for the fact of fission. But our interests in remains of these sorts are not all of equal philosophical interest. Another problem is that the question of what does matter in survival is an empirical question that philosophers have no special competence to answer. The question of what matters in survival, at least if it means what matters to the population at large, is a question that should be left primarily to psychologists.

Where does all of this leave us? It leaves us with evidence that the debate over personal identity is in the throes of an identity crisis of its own. The old questions have died, and the new ones have not yet been born. Or less dramatically: the old questions have been demoted, and the new ones and the rationale for asking them not yet properly formulated. In the good old days when identity and what matters in survival went hand in hand, we knew, or at least we thought we knew, what the central philosophical questions were, and why they were important, and how they should be investigated. Now, it seems, we do not know any of these things. We are making philosophical progress.[13]

NOTES

[1] See Robert Nozick, *Philosophical Explanations*, Harvard, 1981, Ch. 1; Sydney Shoemaker, 'Personal Identity. A Materialist Account', in Sydney Shoemaker and Richard Swinburne, *Personal Identity*, Basil Blackwell, 1984, pp. 67—132; and Derek Parfit, *Reasons and Persons*, Oxford, 1984, Ch. 12.

[2] David Wiggins, *Identity and Spatio-Temporal Continuity*, Oxford, Basil Blackwell, 1967, p. 50.

[3] John Perry and David Lewis have taken a different view. See Perry, 'Can the Self Divide?', *Journal of Philosophy* **69**, 1972, pp. 463—488; and Lewis, 'Survival and Identity', in Amelie Rorty, Ed., *The Identities of Persons*, University of California, 1976, pp. 17—40. But see also Parfit's reply, 'Lewis, Perry and What Matters', in Rorty, *ibid.*, pp. 91—108.

[4] Shoemaker, *op. cit.*, p. 119.

[5] Parfit claims that there is a sufficiently strong overlapping chain of psychological connectedness "if the number of connections, over any day, is *at least half* the number of direct connections that hold, over every day, in the lives of nearly every actual person," in Parfit, *op. cit.*, p. 206.

[6] Michael Slote and Neera Badhwar have both suggested to me, independently of each other, that the Teletransporter does not necessarily break physical or bodily continuity. The reason is that if persons or their psychologies can be discontinuous and can survive sudden and complete replacement of their "component parts," which in a sense they can on Parfit's view, then why not also physical objects and/or bodies? If Slote and Badhwar are right, Parfit is wrong, but not in ways which undermine the points that I make in the current paper. These latter points are based on fission cases and arise even if there is a correct bodily criterion of personal identity.

[7] I discovered, after submitting this paper for publication, that Joseph Margolis briefly makes a similar criticism of Parfit, which of course I accept, from which he draws a very different conclusion, which I don't accept: see his 'Critical Notice: *Reasons and Persons*', v. 47, December, 1986, *Philosophy and Phenomenological Research*, pp. 320–321.

[8] Parfit, *ibid.*, pp. 282–3, numerals and brackets added.

[9] There are other ways to try to reformulate Parfit's question. For instance, one could specify some data concerning what matters to people in certain choice situations covering both standard and fission cases, and then ask how such data are best explained, hopefully yielding psychological continuity as the answer. The problems here are, first, that unless one identifies the data ostensibly, there may be no acceptable way to characterize which data it is one's theory explains, and, secondly, that there may be no plausible explanation of why it is important to explain just this set of data. There are similar problems with other attempts at reformulation.

[10] What follows in this section is a shortened and somewhat revised version of my argument in Part IV of 'Memory, Connecting and What Matters in Survival', *Australasian Journal of Philosophy* **65**, 1987, pp. 82–97.

[11] Parfit, *op. cit.*, p. 284.

[12] It may seem that Parfit could avoid my criticisms simply by enriching his notion of psychological continuity to include such states as the feeling of closeness one may get as a consequence of knowingly experiencing an original. But the problem is not that Parfit's notion of psychological continuity is too thin, but rather that Parfit overlooks the instrumental value of physical continuity.

[13] Earlier versions of this paper were read at the University of Maryland and at Dalhousie University. I am grateful to several people at both institutions for helpful questions and comments. I argue for a solution to the problem that I raise in this paper in 'Personal Identity and Survival: What *Really* Matters?', forthcoming.

Department of Philosophy,
University of Maryland,
College Park, MD 20742,
U.S.A.