

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

EXPERIENCE, AGENCY, AND PERSONAL IDENTITY*

BY MARYA SCHECHTMAN

I. INTRODUCTION

Questions of personal identity are raised in many different philosophical contexts. In metaphysics the question at issue is that of personal identity over time, or of what relation a person at one time must bear to a person at another time in order for them to be, literally, the same person. There are two standard responses to this question in the current literature. One considers a person as essentially a biological entity and defines the identity of the person over time in terms of the continued existence of a single organism. The other follows John Locke in accepting a distinction between persons and human beings and defines the identity of a person in terms of the continued flow of psychological life.¹ This second response—the “psychological continuity theory”—has a great deal of appeal and enjoys a great many supporters. Both positions are still very much alive in the current discussion, suggesting that at the very least each response expresses some important aspect of our thought about what we are and how we continue. In what follows I will leave aside the dispute *between* these two accounts of identity and will focus only on the psychological side. Understanding what fuels the idea that personal identity should be defined in psychological terms, and what a viable psychological account of identity would look like, is an important goal in its own right and will put us in a better position to understand the relation between psychological and biological accounts of identity.

A central element in the motivation and defense of the psychological continuity theory of personal identity comes from the intimate connection between the concepts of identity and survival. One obvious way of understanding what it means for a person to survive is for there to be someone in the future who is she. Since surviving is, to put it mildly, a matter of the utmost practical importance, it seems that whatever relation defines personal identity over time should also bear the importance we attach to

* I am indebted to many friends and colleagues for their input in the course of writing this essay. I would like especially to thank David DeGrazia, Anthony Laden, Ray Martin, Marc Stors, and the editors of *Social Philosophy and Policy*.

¹ There are, of course, other positions that are and have been defended. Historically it has been popular to define personal identity in terms of the persistence of an immaterial soul, and this view still has its defenders. Others see identity as irreducible and unanalyzable. Continuity of organism and continuity of psychological life are, however, generally acknowledged as the basic candidates for an account of personal identity in the current discussion, and the vast majority of authors in the area defend one or the other of them.

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

survival. Psychological continuity theorists argue that when we ask ourselves which kind of continuation—physical or psychological—seems to bear this importance, the latter emerges as the better candidate. This view trades on the intuition that a situation where a person's body continues to function but his psychological life comes to an end (as in, e.g., a case of irreversible coma or total possession) seems like a failure to survive, while a situation in which psychological life continues despite the death of a person's body (as in, e.g., some of the traditional religious depictions of an afterlife) seems like a case of personal continuation.

It remains, however, for psychological continuity theorists to describe in more detail the kind of psychological continuation that constitutes survival. The initial and most natural idea is that persons should be construed as experiencing subjects, and that personal identity should be defined in terms of a deep unity of consciousness throughout an entire life, but this unity turns out to be difficult to define coherently. For this and other reasons, some identity theorists who support a psychologically based view have shifted from a picture of the person as a subject to a picture of the person as an agent, and from a definition of personal identity in terms of a deep phenomenological unity to a definition in terms of the unity of agency. This shift leaves us with the question of whether unity of agency is a strong enough relation to provide an account of what it is for a person, literally, to survive.

Some resources for answering this question can be found in other areas of philosophy—in particular, in discussions of free will, practical reason, and the theory of action, which I will refer to collectively with the term “practical philosophy.” A prominent strand of thought there suggests that having deeply held projects and commitments is essential to being a person, and describes how a person can literally disintegrate from a lack of such projects and commitments. While this discussion does make a strong case that unity of agency must be part of any acceptable psychological account of the persistence of the person, the original intuition that some sort of deep unity of consciousness is required for personal survival runs deep, and a view couched entirely in terms of unity of agency is unlikely to be fully satisfying. Fortunately, the insights gleaned by following out the shift toward a view of the person as agent can help point to a way that something like phenomenological unity of consciousness might be introduced without so much difficult metaphysical baggage.

My main goal here is to follow out the way in which a picture of persons as essentially agents has found its way into the metaphysical discussion of personal identity, and to assess where this leaves us in our attempts to understand the conditions of personal persistence. I begin by showing how the internal pressures of the metaphysical discussion lead to the introduction of an understanding of persons as agents, and how work in practical philosophy can help support the plausibility of this understanding as a response to the metaphysical question of personal

persistence. After this I discuss why a response to this question wholly in terms of agency is likely to remain unsatisfying, and I point very briefly to a possible strategy for reintroducing an understanding of the person as subject as part of a more integrated account.

II. WHAT MATTERS IN SURVIVAL AND PETER UNGER'S DISTINCTION

As I have already mentioned, the central considerations supporting the psychological continuity theory are based on the idea that it is psychological rather than physical continuation that captures what matters to us in survival. As discussion of this view has unfolded, many psychological continuity theorists have found reason to put aside the question of personal identity per se and ask directly about the conditions of survival and what we find important in it. The main impetus for this switch is a difficulty with using psychological continuity as a criterion of personal identity. Psychological continuity—at least as usually defined by these theories—does not have the logical form of an identity relation. The continuity relation, unlike identity itself, can be one-many, intransitive, and a matter of degree. Derek Parfit famously argues that this does not undermine the psychological continuity theory but rather reveals that what really underlies our interest in the question of personal identity over time is the issue of survival.² In ordinary circumstances, survival will involve identity; the person as whom one survives will be oneself. However, in certain imaginable science fiction scenarios, Parfit argues, it may be possible to survive as someone else. This might happen, for instance, if a person could split amoeba-like into two qualitatively identical people. Such a division hardly seems like death, he argues, yet the two resulting people, since they are clearly not identical with one another, cannot both be identical to the original person on pain of intransitivity. This “fission” scenario is thus a case in which a person enjoys what we take to be significant in survival despite the fact that there is no one in the future to whom she is identical.³

Parfit holds that our interest in questions of identity is an artifact of the fact that in ordinary circumstances survival requires identity with some future person. Once we see that we can have what matters in survival without identity, we can see that our main interest is with survival rather than identity itself. The question of whether or not we are going to sur-

² In fact, Parfit shifts his use of the term “survival” over the course of his work. In Derek Parfit, “Personal Identity,” *Philosophical Review* 80 (1971): 3–27, he suggests that survival can come apart from identity, and suggests that we are interested in questions of survival. In his book *Reasons and Persons* (Oxford: Clarendon Press, 1984), Parfit seems to imply that, strictly speaking, survival requires identity, and that we are interested in “what matters in survival.” I will be using the terminology in the former sense. For a full discussion of the evolution in Parfit’s use of this term, see Marvin Belzer, “Self-Conception and Personal Identity,” elsewhere in this volume, Section IIC.

³ Parfit, *Reasons and Persons*, chap. 12.

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

vive undoubtedly bears a special significance for us. As John Perry puts it: “You learn that someone will be run over by a truck tomorrow; you are saddened, feel pity, and think reflectively about the frailty of life; one bit of information is added, that someone is you, and a whole different set of emotions rise in your breast.”⁴ What we are after with questions of diachronic personal identity (that is, questions about the conditions for the persistence of a single person over time), says Parfit, is the question of what relation to the future carries this importance—“what matters” in survival. Many theorists of personal identity who would define it in psychological terms have thus shifted their emphasis from identity to what matters in survival.

It might seem as if the shift of focus to “what matters” automatically turns the discussion to a more metaphorical or psychological notion than that with which we started, since we are now asking what makes our continuation a good rather than what makes us, literally, still alive. Most psychological continuity theorists would deny this, however. The survival they are trying to define is, they insist, quite literal, and the sense of “what matters” at issue should be read that way. This point is made most explicitly by Peter Unger, who says that there are different things we could mean by “what matters in survival,” and that it is important for us to keep clear on which of them is under consideration. He distinguishes between three different uses of this term. One, to which I will return later, concerns the question of which features of survival *constitute* it as survival. For present purposes, the most important distinction is between the other two senses, which he calls the “desirability” and the “prudential” senses of “what matters.” According to the desirability sense,

‘what matters in survival’ will mean much the same as this: what it is that one gets out of survival that makes continued survival a desirable thing for one, a better thing, at least, than is utter cessation. On this desirability use, if one has what matters in survival, then, from a self-interested perspective, one has reason to continue rather than opt for sudden painless termination.⁵

This is contrasted with the prudential use, which determines,

from the perspective of a person’s concern for herself, or from a slight rational extension of that perspective, what future being there is or, possibly, which future beings there are, for whom the person rationally should be “intrinsically” concerned. Saying that this rational concern is “intrinsic” means, roughly, that, even apart from questions

⁴ John Perry, “The Importance of Being Identical,” in Amélie Oksenberg Rorty, ed., *The Identities of Persons* (Berkeley: University of California Press, 1976), 67.

⁵ Peter Unger, *Identity, Consciousness, and Value* (New York: Oxford University Press, 1990), 93.

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

of whether or not he might advance the present person's projects, there is this rational concern for the welfare of the future being.⁶

Unger suggests that there is a question of whether a future will be mine to experience—whether I will undergo what happens in it—which at least seems quite independent of my values, goals, ideals, or character traits.

If we are looking to answer literal, metaphysical questions about our survival, Unger argues, we should be looking for the relation that underlies the more fundamental, prudential connection between different moments of a person's life. The desirability sense of what matters may be an important thing to define in its own right, but it is not, he says, "highly relevant to questions of our survival."⁷ To test for the prudential sense of what matters, Unger suggests that we use the "avoidance of future pain test." We are to imagine ourselves connected to a future person in some way, and then imagine that person being subjected to horrible torture or excruciating pain. The question we are to ask is whether our horror of the future pain is sympathetic in nature or rather the sort of concern we would feel if we were anticipating our own future pain. The latter response suggests that we have what matters in survival in the fundamental sense at issue in metaphysical discussions of personal survival.

Unger himself believes that attention to this question will lead us to a more biologically based account of personal survival, albeit one that requires the continuity of minimal psychological capacities (what he calls "core" psychology). Traditional psychological continuity theorists have rejected Unger's solution, but they seem to accept his basic insight. They suggest that the avoidance of future pain test reveals that what matters fundamentally in survival is the legitimate *anticipation* of future experiences. To survive we must bear a connection to the future that makes it rational for us to expect to have experiences then. These theorists do not accept Unger's solution because they deny, for reasons given by Locke, that the continuity of the body can itself be what legitimates anticipation of future experiences. (Locke famously argues through the use of hypothetical cases that continuity of body—or soul—with no continuity of consciousness cannot provide what we seek in survival.) For such anticipation we need a deep phenomenological connection to the future—a unity of consciousness over time. Continuity of the body might turn out to be one source of such a phenomenological connection; it might even turn out to be the only source. This would, however, be a contingent fact, and even if it were true it would not be the continuity of the body per se, but rather the unity of consciousness it produced, that would make our intrinsic concern for some future person rational. On this view, then, a *person* is most fundamentally an experiencing subject, and survival must

⁶ Ibid., 94.

⁷ Ibid., 93.

involve the unity of an experiencing subject over time. This is, in many ways, a very natural and appealing picture of what is involved in our survival as persons, but it is difficult to develop a coherent theory based on this picture, as the next section will explain.

III. DEREK PARFIT'S ARGUMENT AGAINST THE PERSON AS EXPERIENCING SUBJECT

A compelling initial picture of what constitutes literal survival for those who take a psychological approach to this question is, as we have seen, a unity of consciousness over time.⁸ Derek Parfit has argued quite forcefully, however, that this natural first idea of what defines survival cannot be defended—or even defined—in the end. He provides several arguments explicitly directed at the idea that personal identity cannot be defensibly defined in terms of the unity of a subject of experiences, but most of the fundamental insights behind these arguments can be found in the Teletransportation and Branch-Line cases with which he opens his discussion of personal identity in *Reasons and Persons*.⁹ This pair of cases uncovers both the hold this conception of personal continuation has on us, and its problematic nature. In Teletransportation, a subject enters a booth where his body is scanned and dematerialized. A molecule-for-molecule duplicate of the original body is built on the destination planet, a duplicate who by hypothesis replicates the intrinsic character of the original person's psychological states exactly. The question is whether this is an efficient means of travel to distant planets, or death and replacement by an imposter. Parfit acknowledges that we are quite likely to think of it as the latter. Despite the exact similarity of the newly formed individual to the original person, we may well fear that the replica will lack a deep attachment to that person present in ordinary survival. We fear that while in ordinary survival one will *feel* one's future experiences, in Teletransportation one will feel nothing in the future—one's point of view, which would continue in the ordinary case, is snuffed out.

The presupposition underlying this fear is that there is a deep unity of consciousness within a given life that may be absent in our connection to a replica. It is not immediately obvious exactly what this deep unity of consciousness is, but the basic idea can be understood on the model of co-consciousness at a time. At a given time there is a brute difference between a set of experiences being present in a single consciousness and their being parts of distinct consciousnesses. This is why, for example, a decision about whether to tolerate a small pain in my shoulder now to avoid an excruciating pain in my foot now is different in kind from a

⁸ As mentioned in note 2, for terminological simplicity in what follows I will refer to having what matters in this fundamental sense simply as "survival," meaning that in my use survival does not automatically imply identity.

⁹ Parfit, *Reasons and Persons*, 199–201.

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

decision about whether to tolerate a small pain in my shoulder now to avoid *your* experiencing an excruciating pain in *your* foot now. The fear of Teletransportation reveals the supposition that we have a relation to our future experiences which is relevantly like the relation we have to our own present ones. This is why the decision about whether to suffer a small pain now to avoid an excruciating pain *later* is also different from the decision about whether to suffer a small pain now to keep someone else from suffering an excruciating pain later.¹⁰ This difference is, of course, what the avoidance of future pain test proposed by Unger is testing for. When we fear that we will not survive Teletransportation, we are expressing the possibility that this connection to the future may not hold in cases of replication. In more concrete terms, we suspect that my attitude toward the replica's potential pains should be more like my attitude toward the potential pains of others than toward my own.

Since a teletransported replica is, by stipulation, completely identical to the original person in terms of the *contents* of her consciousness, whatever difference there is between ordinary survival and having a replica cannot be defined in terms of these contents. The previous discussion makes it clear, however, that it is supposed to be a difference in the *quality of experience*. To make the case that there could be such a difference, we thus need to be able to understand a phenomenological distinction between really continuing and being replaced by a replica which is not defined in terms of the contents of consciousness. Parfit's claim is that there is no defensible way to define such a distinction. The difference we are seeking cannot be explained by the fact that in one case we retain our original body, since the psychological account is based on the intuition that we could in theory have what counts as survival without continuation of the body. If the sameness of body plays a crucial role, it will have to be because continuing in a single body is experientially different from being replaced by a replica, and we still need to say something about the nature of this phenomenological difference.¹¹ Parfit thinks that we might have been able to make sense of this mysterious unity of consciousness over time if we accepted the view that we are simple, immaterial souls (although the Lockean arguments make this questionable in any event), but argues that there is no reason to believe such a thing, and many reasons not to.¹²

¹⁰ Of course, in either case we may decide that suffering the small pain is the better thing to do. We can care as much or more about the pains of others as we do about our own. The point is just that the considerations that go into the decision making are different in each case.

¹¹ Of course, at this point we might revert to a biological criterion, saying that our special concern about our own futures is not based on a deep unity of consciousness, but only on the fact that the organism currently having experience will be present, sentient, and in pain in the future. There are many considerations in favor of this position, and Parfit also offers many arguments against it. For the present, I am only interested in following out the intuitions behind psychological accounts of identity. I will, therefore, only consider what would be required to make this approach work, and see where this leads us.

¹² Parfit, *Reasons and Persons*, 227–28.

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

Of course, the inability to define the deep unity of consciousness that we assume in our own lives—and fear might be absent in replication—may not be enough to convince us that it does not exist. It seems as if we actively experience this continuity and connection in our own lives. Moreover, it seems all too possible to imagine the flow of our own psychological life terminated or diverted despite the existence of a replica with qualitatively identical contents of consciousness. Kathleen Wilkes makes this point quite effectively in a footnote in *Real People: Personal Identity without Thought Experiments*. Discussing the transporter depicted on the science fiction series *Star Trek*, she says:

Captain Kirk, so the story goes, disintegrates at place p and reassembles at place p^* . But perhaps, instead, he dies at p and a doppelganger emerges at p^* . What is the difference? One way of illustrating the difference is to suppose there is an afterlife: a heaven, or hell, increasingly supplemented by yet more Captain Kirks all cursing the day they ever stepped into the molecular disintegrator.¹³

Once the viewpoint of the original Captain Kirk is reintroduced, it seems natural to assume there is a unified flow of consciousness between the original Kirk and the disembodied Kirk—a flow of consciousness that does not exist between the original Kirk and his doppelganger. If we now simply take out the assumption of an afterlife, we can see that the connection between the original and doppelganger remains phenomenologically deficient.

Parfit himself is very sensitive to the pull of this idea, and this is why he introduces the variation on his Teletransportation case that he calls the “Branch-Line case.” In the Branch-Line case, the traveler is scanned and a blueprint made, but the original body is not dematerialized. Instead, it is damaged in such a way that it can live only a few more days. At the same time, a healthy replica is built on the distant planet (here Mars) as in simple Teletransportation. There are now two people who are psychologically continuous (in Parfit’s sense) with the original person: the replica on Mars and the person with the original body on Earth. It seems clear, Parfit acknowledges, that the connection between the person on Earth (now traveling on a branch line of psychological continuation) and the replica does not contain what matters most fundamentally in survival. Parfit imagines himself on the branch line talking to his replica on Mars:

Since my Replica knows that I am about to die, he tries to console me with the same thoughts with which I recently tried to console a dying friend. It is sad to learn, on the receiving end, how unconsoling these

¹³ Kathleen Wilkes, *Real People: Personal Identity without Thought Experiments* (Oxford: Clarendon Press, 1994), 46n.

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

thoughts are. My Replica then assures me that he will take up my life where I leave off. He loves my wife, and together they will care for my children. And he will finish the book that I am writing. . . . All these facts console me a little. Dying when I know that I shall have a Replica is not quite as bad as, simply, dying. Even so, I shall soon lose consciousness, forever.¹⁴

The replica is as like the branch-line person in terms of contents of consciousness as a person is like herself in the near future, but there is all the difference in the world for this person between her psychological life continuing several more decades and its being terminated while the replica continues for those decades—or so it seems. This is why, Parfit acknowledges, “it is natural to assume that my prospect, on the Branch Line, is almost as bad as ordinary death.” Nonetheless, he goes on to say, “[a]s I shall argue later, I ought to regard having a Replica as being about as good as ordinary survival.”¹⁵

He offers a variety of considerations in favor of this claim. These include a collection of arguments for a reductionist view of persons, arguments against the possibility of directly experiencing a deep unity of consciousness over time in our own lives, and thought experiments aimed at making it more plausible that the person on the branch line could survive as the replica. The case most directly addressing this issue is, importantly, not a science fiction case, but one that is part of a real life. Parfit describes a sleeping pill he has taken that takes a short time to work. After taking the pill, the person remains fully conscious for a while, but when he awakens he has no memory of the time between taking the pill and falling asleep. Parfit reports that having once taken such a pill, he apparently solved a practical problem before falling asleep, for he found a note the next morning under his razor advising him of the solution, despite having no memory of solving the problem or writing the note. He argues that the person who has taken such a pill can know, like the person on the branch line, that his stream of consciousness will end before morning, but he is unlikely to face this prospect as if he is facing death. Yet, Parfit argues, in terms of intrinsic relations of consciousness or phenomenology the relation of the person who takes this sleeping pill to the person in the morning is no different from the relation of the branch-line person to his replica on Mars. There are, indeed, a host of differences between the two cases, but they do not seem to be *intrinsic* differences, and so if we want to say that these make the difference between surviving and not surviving we will have given up on the idea that the importance of survival is linked to a deep diachronic unity of consciousness.

¹⁴ Parfit, *Reasons and Persons*, 201.

¹⁵ *Ibid.*

Cambridge University Press

978-0-521-61767-3 - Personal Identity

Edited by Ellen Frankel Paul, Fred D. Miller and Jeffrey Paul

Excerpt

[More information](#)

While there is much to understand—and to take exception to—in Parfit's arguments, the underlying theme is, I think, fairly simple and quite persuasive. What his arguments do at their most effective is direct us to commonplace examples of survival and undermine our conviction that there is any deep unity of consciousness there. The sleeping-pill case, for instance, tells us less about a heart-damaged person with a replica on Mars than about what happens to us all the time. We do not need to take sleeping pills to have gaps in our consciousness, or to know sometimes that what we are thinking and experiencing in the present will be entirely lost to future consciousness—this is why people make lists and keep journals. This challenge is reinforced by the sort of empirical work used by Daniel Dennett and Andy Clark, among others, to reveal that our unity of consciousness—even at a time—is not all that casual introspection might lead us to believe.¹⁶

Whatever the upshot of detailed engagement with Parfit's many arguments, he raises a powerful and serious challenge to one deep-seated understanding of what is involved in our survival, and its importance. A tempting and natural way to think about our continuation is in terms of a flow of consciousness that connects current experience to future experience in something like the way our current experiences are connected to one another. It is this flow of consciousness that seems to allow us to really have experiences in the future—to *be* there in the sense that is most important. Parfit's repeated challenge to this picture, emblemized in the Teletransportation and Branch-Line cases, is essentially David Hume's—that if we take the time to really look, we can find no such unity in our own lives. This is why Parfit finally says of Teletransportation: "I want the person on Mars to be me in a specially intimate way in which no future person will ever be me. My continued existence never involves this deep further fact. What I fear will be missing is *always* missing."¹⁷ He thus draws the following conclusion: it is not, as he claimed earlier, that Teletransportation is about as good as ordinary survival, but rather that "ordinary survival is *about as bad as*, or little better than, Teletransportation. *Ordinary survival is about as bad as being destroyed and having a Replica.*"¹⁸

If Parfit succeeds in his challenge—which many think he does—the only psychological connection between one's present and one's future is a relation of similarity or association between the contents of consciousness at different times. Since our sense of the importance of survival seems to be based—at least largely—on the assumption of a deep, phenomenological connection to the future, that importance needs to be reassessed in light of Parfit's challenge. The most extreme possibility is that the psychological relations we really do bear to the future have no intrinsic

¹⁶ See, for instance, Andy Clark, *Being There* (Cambridge, MA: MIT Press, 1998); and Daniel Dennett, *Consciousness Explained* (Boston: Little, Brown, and Co., 1991).

¹⁷ Parfit, *Reasons and Persons*, 208.

¹⁸ *Ibid.*