

Personal Identity and Trivial Survival

by

ANDREA SAUCHELLI 

Lingnan University

Abstract: Your replica is created on Mars and you, on Earth, are destroyed. Parfit claims that your replica may still have what prudentially matters for you – provided that you are psychologically connected and continuous with your replica. If someone accidentally destroys the tapes containing your psychological profile used in the production of your replica and this same action fortuitously produces a functionally equivalent tape, Ehring claims that Parfit should maintain that the resulting new individual may still have what matters. Nihilism about what matters follows, or so Ehring claims. I argue that Ehring is wrong and that the difference between the two ways of creating a replica is not trivial – there is no trivial survival.

Keywords: personal identity, Parfit, what matters, survival, memory

1. Relation R, Trivial Survival and Nothing Matters

DOUGLAS EHRLING ASCRIBES TO Derek Parfit the belief that, although “psychological connectedness” can hold in various ways, his (Parfit’s) preferred version is that *any* causal connection can suffice for this relation to hold (Ehring, 1987, p. 51). Strong psychological connectedness between persons (or person stages, depending on the favourite metaphysics of time and persistence) each existing at different times obtains when there are enough psychological connections between them. These connections are quasi-memories, quasi-intentions and actions, character traits and their expressions (Parfit, 1984/87, pp. 206–207). Psychological continuity between two persons existing at different times holds when enough overlapping chains of direct psychological connections hold between them. When psychological continuity does not branch within t_1 and t_2 , the relation of personal identity between two persons within t_1 and t_2 consists only in psychological continuity. Ehring says that Parfit believes that whereas personal identity consists in non-branching psychological continuity, what matters in survival is relation R, where this relation is “psychological continuity and/or connectedness”. In a work more recent than *Reasons and Persons*, however, Parfit claims that this formulation is misleading: we should understand relation R as being psychological connectedness *and* continuity (Parfit, 2007, n. 30). This clarification implies other changes in what Parfit previously stated. For example, in *Reasons and Persons*, Parfit (1984/87, p. 267) claims that personal identity consists in non-branching

relation R holding *in the right way*. However, given his clarification in (2007) of what R is, it cannot be right to say that personal identity consists in only non-branching R since personal identity may hold between P at t_1 and Q at t_2 even in the case of no psychological connections between P and Q – think about Reid’s case of the brave officer and Parfit’s reply to it. It may well be the case that a general does not remember being flogged as a boy (an alleged case of weak or no psychological connectedness due to a lack of a direct memory connection), but the general and the boy may be the same person provided that intermediate steps (e.g., the brave-officer stage) are sufficiently psychologically connected to both. So, personal identity may hold between the boy and the general even though they may not be (strongly) psychologically connected with each other.

Parfit also claims that there are different ways in which personal identity and the what-matters relation may hold: in the normal way (e.g., due to the continuity of the same body or brain), due to reliable (e.g., a well-functioning tele-transporter), or unreliable causes (e.g., a tele-transporter that works only occasionally). In the case of personal identity, Parfit (1984, p. x) claims that we do not have to choose which way is right, so long as we are reductionists (i.e., so long as we believe that personal identity consists in only psychological and physical connections).¹ Once we realize that non-reductionism is false, the interesting question is whether the relation of personal identity coincides with the what-matters relation. As Parfit has clarified, the what-matters-in-survival debate was meant to address the rationality of our prudential/self-concern over time (2007). Whether a future outcome turns out to be desirable or not is irrelevant to the question of whether we have rational *prudential* concern for the future of the individual(s) involved.² For example, one may have rational prudential concern for a future one knows will be unpleasant (undesirable), and because one believes that it will be one’s own future. Parfit’s thesis is that our prudential reasons do not depend on the holding of everything personal identity consists in but on only a part of it – I will not rehearse here his arguments for this conclusion.³ On Parfit’s view, the relation of personal identity and the what-matters relation do not coincide. Personal identity between P at t_1 and Q at t_2 consists in non-branching psychological continuity between P at t_1 and Q at t_2 , and the way in which we regard that personal identity holds is generally in terms of its normal cause (i.e., spatio-temporal continuity of parts of our brain).⁴

1 See Parfit (1999) for a clarification of “reductionism”.

2 See the discussion in Parfit (2007, pp. 20–27) and Setiya (2015) for criticism.

3 See Gustafsson (2018), Johansson (2010), Persson (2016) and Ehring (2018, 2019) for recent discussions.

4 See Parfit (2012) in which he claims that he has defended a normal-cause view of personal identity (not of what matters).

Relation R, which I will assume to be the what-matters relation, on the other hand, may hold for any cause (the *any-cause R relation*), even due to a teletransporter that works occasionally. This last point about R has been misunderstood by many critics, as we will see later in this article.

Ehring's (1987, 2013) arguments against Parfit's main thesis that any-cause-R is the relation that has what matters rest on three cases:

Case 1. In the future, people "transfer" their psychological states to a tape and subsequently such information is impressed or realized into a new body. The original body is then destroyed. The new individual will (at least apparently) remember the original person's experiences, etc. – let us say that the *original* and the *replica* have the same *psychological profile*. Ehring claims that in this case, Parfit would say that the any-cause relation R holds between the original and the replica (the expression "case X has what matters" shall be understood as meaning "relation R holds between the original and the new individual referred to in case X".)

Case 2. After recording the information on a tape (see case 1), on one occasion, the tape is accidentally destroyed but in a way that activates one of the devices that produces new bodies (e.g., the tape is dropped into the device, causing a malfunction). Highly improbable as it is, this accident causes the device to produce a tape that is functionally similar to the original tape that has been dropped (and destroyed). An individual that seems to have the same psychological profile as the original is then created from this new tape. Again, Ehring claims that the any-cause relation R holds also between the new individual and the original.

Case 3. As in case 2, the tape is accidentally destroyed. This time, a new tape is not fortuitously created as a result of the previous event. However, a new tape, encoding the same psychological profile of the original, is accidentally created as a result of an unrelated accident. As before, a new body is then infused with the psychological profile contained on the new tape. The resulting person "has apparent memories of [the original in case 1 and 2], but these memories are not caused in any way by the experiences of [the original person]" (Ehring, 2013, pp. 143–144). Ehring suggests that Parfit should say that this case does not have what matters.

Ehring claims that R has in case 1 a non-accidental causal dependence and in case 2 an accidental causal dependence. Ehring also claims that Parfit believes in the principle saying that the holding of the mattering relation cannot depend on trivial facts. More precisely, the Non-triviality principle holds that if there is only a trivial difference between two cases, there should not thereby be a difference concerning what matters (Ehring, 1987, p. 51; 2013, p. 146). Since the difference between 2 and 3 is trivial, Ehring argues that if 2 has what matters, it is irrational to regard 3 as not having it. In case 2, Ehring believes that Parfit should say that the original person survives as the new individual (and thus the case has what matters). Based on these descriptions of the outcomes, Ehring proposes an argument against Parfit's main thesis that any-cause-R is the relation that properly has what matters. Suppose that what matters is any-cause-R. On Parfit's view, cases 1 and 2 have what matters, but 3 does not. However, the difference between

2 and 3 is trivial, so 3 should have what matters too (what matters should not depend on trivial factual differences).⁵ Hence, any-cause-R does not satisfy the Non-triviality principle. Since the relation that has what matters must satisfy this principle, any-cause-R is not the what -matters relation. In a final crescendo, Ehring also argues that if personal identity consists in non-branching R, and personal identity and R are the only reasonable candidates for what matters, since neither can be the reasonable candidate for what matters, nothing matters in ordinary survival (Ehring, 2013, p. 148).

2. Not So Trivial

One reply to Ehring's argument is that it is wrong to claim that case 2 has what matters. Ehring anticipates this move and says that one reason for distinguishing the two cases is that in 1, but not in 2, the causal relation between the two psychological profiles is "information sensitive". The idea is that only in case 1 does the original's psychological profile play a direct causal role in determining the replica's psychological profile. This point can be articulated in counterfactual terms: had the original psychological profile been different, the replica's profile would have been different too (we may not even be justified in calling this new individual a replica of the original). In case 2, the new individual's psychological profile does not have such causal dependence. This reply allows us to disarm the appeal to triviality in the move from case 2 to case 3 since, with respect to the causal dependence just described, 3 may not be different from 2, but 2 and 1 are different. Ehring's reply is that this reasoning fails for two reasons. The first is that "it is far from clear that the normal causal processes associated with memory and intention have this 'information sensitivity'" (Ehring, 2013, p. 144). I take this to be a revealing point of what is frequently presupposed in the debate, the fault being also in some of Parfit's formulations of his views. I will develop a criticism of this point more fully in the next main section because my criticism requires a more elaborated understanding of the causal relations between some of the relevant mental states. The second reason against distinguishing case 1 from 2 is that, even if the processes associated with memory normally work by means of such a causal dependency, it "is not necessary for what matters that every link in the relevant causal chain display this characteristic" (Ehring, 2013, p. 144).⁶

⁵ More precisely, Ehring claims that, in case 3, R does not hold, but since the difference between 2 and 3 is trivial, if it is rational to be prudentially concerned with the fate of the replica in 2, we should be concerned also in 3, thus, given the Non-triviality principle, R should hold there too (Ehring, 2013, pp. 144 and 146, n. 14).

⁶ Ehring distinguishes between "information sensitivity" and "causal dependency", but I discuss the cases involving them together as I assume that the first depends on (a proper form of) the second.

This second argument does not seem to help Ehring – and the example he gives to explain his point is revealing of what I take to be a misunderstanding of Parfit’s position. Ehring claims that if there is a link in, for example, A’s memory connection at t_1 that fails to display the normal causal dependency with respect to a seeming memory of B at t_2 , the relation between A and B may still have what matters (Ehring, 2013, pp. 144–145). Ehring is right to claim that it may not be necessary that *each* and *every* mental state of A at t_1 has to have the normal sort of causal dependency for A to still have what matters at t_2 , but this reasoning does not show that *all* of A’s connections with B may not have such a sort of normal causal dependency and constitute what matters. On a charitable understanding of the example at (2013, pp. 144–145), Ehring may be taken as proposing a case in which, despite a short time gap t (say, a glitch of a few seconds) in the holding of all of the relevant psychological connections, which are allegedly resumed after t , the case can still have what matters. Again, although the precise details of how this gap is supposed to influence the relevant psychological connections before and after the temporal gap are not specified, we may say that, in this case, Ehring seems to simply state the controversial point (i.e., that the case has what matters) but that, after all, the example does not show what it purports to do – one of the reasons being that, on reflection and given a proper understanding of Parfit’s R (see section 3), R does not hold between the individuals before and after the gap. For, in this case, psychological connections and continuity do not hold between the individuals before and after t , given that the holding of the relevant relations depends on, among other things, the right causal connections (which in the case at issue seem to be interrupted, by hypothesis). Besides, contrary to what Ehring (2013, p. 145) says, it is simply false that *most* of the psychological connections in case 2 or 3 have the normal kind of causal dependency (e.g., the “information sensitive” causal connection).

When Ehring (2013, p. 145) claims that there is no reason to believe that information-dependent-causal-chains of psychological connections are required for what matters, we may understand him as suggesting that, in principle, we could abandon the causal requirement completely. However, surprisingly, Ehring claims that without such a requirement, the what-matters relation would be too liberal. In support of this idea, he proposes the following example. Suppose that A has an accident and dies, while at the same time B, an individual on a distant planet, fortuitously instantiates A’s psychological profile as a result of a causally unrelated event. In this case, B is not related to A by the what-matters relation – any-cause-R does not hold between B and A. In his explanation of why any-cause-R does not hold in this example, Ehring (2013, p. 145) also claims that A’s prudential concern for B is not justified if B is simply a causally unrelated “similar replica”. If what-matters does not hold because A and B are only similar and not numerically

identical – exact similarity is not sufficient for the what-matters relation – then we may argue that the difference in terms of information dependency causal connections between 1 and 2 is *not* trivial or unimportant for the relation of what matters. In fact, it is hard to tell what if not an appropriate causal connection or a related (non-branching, non-interrupted) spatio-temporal continuity could ensure that B is not only similar to A but also numerically identical.⁷ This last case is a minor variation on 3, and Ehring claims that both cases do not have what matters. If this is true, Ehring needs to say on what grounds 2 has what matters while 3 does not without recurring to the information-sensitive counterfactual connection discussed above – and he cannot resort to spatio-temporal continuity because it does not hold in the relevant cases. The holding of the information-dependent-causal-chains of connections seems to be the only element so far recognized by Ehring that can ensure the fact that the various cases under discussion involve numerical identity and not just similarity. In short, either the difference between 1 and 2 is not trivial (and thus his reasoning should be dismissed), or he cannot use the previous reasoning to dismiss the option that R may hold between A and B even if they are merely similar (an equally implausible outcome).⁸

To address the last standing objection raised by Ehring (“it is far from clear that the normal causal processes associated with memory and intention have this ‘information sensitivity’”), to further my criticism of one of Ehring’s arguments above, and to clarify some of Parfit’s own claims about relation R, we need to investigate what these causal processes associated with (quasi-)memory are.

3. Memory and R

A common feature of recent causal memory theories (CMT) is their use of the concept of a memory trace, which is also present in Parfit’s account.⁹ Although there is no consensus on how to define such a notion, I assume that memory traces can have a mental and a physical description.¹⁰ As physical structures, they

7 This was Bernard Williams’s point in his (1956) and one of the reasons Shoemaker (1970) and others included a causal requirement in their accounts of personal identity (via their characterization of various relevant quasi-mental states).

8 Referring to the nihilistic conclusion of Ehring (2013), at this stage of the argumentation, does not seem warranted either. After all, we are still debating whether the evaluation of the cases that will justify such a conclusion is plausible.

9 A causal theory of memory can be devised without referring to memory traces. According to some, they should be separated (Robins, 2016). Theories of memory requiring some form of causal connection between an experience and a subsequent episode of recall of that experience are standard in the current debate.

10 See Martin and Deutscher (1966), Tulving (2007), Sutton (1998) and, in particular, Bernecker (2010, pp. 131–132).

are structural modifications of synapses the effect of which is the facilitation of the activation of neurons in a neural network (Tulving, 2007, p. 66). Because these traces preserve mental content, they can be regarded as involving mental representations too. On Sven Bernecker's CMT, the appropriate causal connection between an experience and its subsequent recall is partly characterized by these conditions (Bernecker, 2010, p. 151):

- (i) S's representing at t_1 that p^* and S's representation at t_2 that p are connected by a persisting memory trace or a contiguous series of memory traces.
- (ii) If S had not represented at t_1 that p^* he would not represent at t_2 that p .¹¹

The preservation of mental content through time (or, at least, of a sufficiently similar content) and subsequent occurrences of related states of recall depend on memory traces (Bernecker, 2010, p. 131). According to Bernecker's externalism about memory content, memory traces have the content they do in virtue of systematic relations the subject bears to certain aspects of her physical states and social environment. On this account, artificially created "memories" or mental configurations created without any causal connection with previous mental traces would generate states that are not memories in the proper sense. In discussing cases of tele-transportation, Bernecker claims that the replication of traces can be compatible with subsequent episodes of remembering only if the causal chain connecting the past and present representations is contiguous – which in turn depends on the details of the case at issue. For example, he claims that if such a replication process is akin to downloading information from a trace onto a computer, sending a file to another device, copying the information into an empty trace, then the causal chain may well be temporally and spatially contiguous (Bernecker, 2010, p. 140). What Bernecker calls the "counterfactuality of memory causation" – condition (ii) above – resembles Ehring's information sensitivity counterfactuality. However, according to Bernecker, far from being trivial, condition (ii) is a necessary condition for having a memory state at t_2 of a previous experience.

Let us once again examine Ehring's cases. Depending on the details of case 1, it may be true that, on a certain understanding of memory traces, there can be memory connections between the original and the new individual – who would be a proper replica of the original. Provided that a sufficient number of psychological connections can be generated by the procedure, the replica may even have

¹¹ There are other conditions discussed in his account, but it is sufficient to focus only on the above for the purposes of this essay. Michaelian (2011, pp. 330–337) proposes further refinements of this theory that seem highly plausible.

strong psychological connectedness and thus be psychologically continuous with the original. Relation R would hold even in case the original were not destroyed. In case 2, however, memory connections do not hold: condition (ii) is *not* satisfied and the memory traces generated by the experiences of the original are not involved in the generation of memory-resembling experiences in the new individual. As a result, psychological connectedness and continuity do not hold between the original and the new individual since these two relations do not hold, R does not hold either. The last case is relevantly similar: psychological connectedness and continuity do not hold between the original and the replica – once again, the what-matters relation does not hold either.

In reply to Ehring's last standing criticism, it is very clear that the main family of theories of memory today state exactly the opposite of what he says. I would suggest that even theories that do not rely on memory traces or on causal connections between experiences and later episodes of remembering should include at least a counterfactual condition similar to Bernecker's condition (ii) to distinguish between genuine cases of recollection from cases of re-learning, mere confabulations and hallucinations. So, the difference between Ehring's case 1 and 2 is far from trivial; in fact, it makes all of the difference for the holding of relation R. After all, whether something is trivial/not important depends on the context. Since here the context is the holding of the what-matters relation and that this relation consists in the holding of memory connections also, it is not trivial to the case at issue whether such memory connections hold or not. In particular, there is no plausible understanding of "triviality" that applies here to the difference between the holding or not of at least condition (ii). As stated previously, R holds only if enough psychological connections such as (quasi-)memory connections hold, but such connections do not hold at all in case 2. So, in case 2, the original and the new individual are not related by R and there is no need to slide into nihilism (it is not necessary to discuss case 3 either.)

4. Coda

What I think has misled Ehring and other critics of the thesis that identity is not what prudentially matters is also some of Parfit's own formulations of relation R. For reasons explained previously, Parfit should have made clear that, if the holding of R consists in the holding of (strong) psychological connectedness *and* continuity, then (1) R depends on the holding of psychological connections, and (2) the holding of some of these connections (e.g., memory connections) require certain appropriate types of causal dependence. When Parfit claims that R is psychological connectedness and continuity *with any cause*, he might be understood as suggesting that, for example, a memory connection between A and B can hold

independently of any specific causal connection between A and B, so long as there is any kind of causal connection between A and B. This understanding is wrong – after all, Parfit explicitly refers to a causal theory of memory along the lines of Shoemaker (1970) and Martin and Deutscher (1966).¹² The previous erroneous understanding would be wrong even if we ascribe to Parfit a purely internalist view of memory content; for example, even if we ascribe to Parfit the belief that the content of B's representation at t_2 of one of A's experiences at t_1 only depends on the intrinsic physical states of A at t_1 and B at t_2 (Parfit, 1984/87, pp. 219–223). For example, on a purely internalist interpretation of Parfit's view, (1) the content of B's representation can be the same (or sufficiently similar) to A's previous representation, but (2) for B's representation to qualify as a case of memory of A's experience, their two experiences or their subjects must be still relevantly causally connected (Parfit, 1984/87, p. 226).

In short, I think that Parfit's view on R should be put as follows: What matters is the holding of relation R in any way that is compatible with the holding of (strong) chains of psychological connectedness, continuity, and of the relevant psychological connections on which they depend. I think that this is what Parfit thought all along: his example of the holding of any-cause-R was an unreliable tele-transporter machine. Although unreliably, such a device operates on the basis of the psychological profile of the original individual, thus maintaining memory connections between the original and the replica. The holding of R is not preserved in Ehring's case 2, so the difference in terms of counterfactual dependency is not and cannot plausibly be regarded as trivial.

Acknowledgements

Many thanks to Rafael De Clercq and Chiu Wai Wai for their comments and suggestions on an early version of this article.

References

- BERNECKER, S. (2010) *Memory*. Oxford: Oxford University Press.
 EHRLING, D. (1987) "Survival and Trivial Facts." *Analysis* 47: 50–54.
 EHRLING, D. (2013) "Why Parfit Did Not Go Far Enough." *Philosophical Studies* 165: 133–149.
 EHRLING, D. (2018) "Why Parfit Cannot Generalize from Fission." *Analytic Philosophy* 59(3): 413–425.
 EHRLING, D. (2019) "Johansson on Fission." *Acta Analytica* 34(2): 155–163.

12 See Shoemaker (2004, p. 581) for the claim that Parfit's famous example of the "Venetian memories" (Parfit, 1984/87, p. 220) is not a genuine case of quasi-memory.

- GUSTAFSSON, J. (2018) "The Unimportance of Being Any Future Person." *Philosophical Studies* 175: 745–750.
- JOHANSSON, J. (2010) "Parfit on Fission." *Philosophical Studies* 150: 21–35.
- MARTIN, C. B. and DEUTSCHER, M. (1966) "Remembering." *Philosophical Review* 75(2): 161–196.
- MICHAELIAN, K. (2011) "Generative Memory." *Philosophical Psychology* 24(3): 323–342.
- PARFIT, D. (1984/87) *Reasons and Persons*. Oxford: Clarendon Press.
- PARFIT, D. (1999) "Experiences, Subjects, and Conceptual Schemes." *Philosophical Topics* 26(1–2): 217–270.
- PARFIT, D. (2007) "Is Personal Identity What Matters?" *Marc Sanders Foundation*, 31 December, <http://www.marcsandersfoundation.org/wp-content/uploads/paper-Derek-Parfit.pdf>.
- PARFIT, D. (2012) "We are not human beings." *Philosophy* 87(1): 6–28.
- PERSSON, I. (2016) "Parfit on Personal Identity." *Theoria* 82(2): 148–165.
- ROBINS, S. (2016) "Representing the Past." *Philosophical Studies* 173: 2993–3013.
- SETIYA, K. (2015) "Selfish Reasons." *Ergo* 2(18): 445–472.
- SHOEMAKER, S. (1970/2003) "Persons and their Pasts." In S. Shoemaker (ed.), *Identity, Cause, and Mind*, pp. 19–48. Oxford: Oxford University Press.
- SHOEMAKER, S. (2004) "Brown-Brownson Revisited." *Monist* 87(4): 573–593.
- SUTTON, C. (1998) *Philosophy and Memory Traces*. Cambridge: Cambridge University Press.
- TULVING, E. (2007) "Coding and Representation." In H. L. Roediger, Y. Dudai and S. M. Fitzpatrick (eds), *Science of Memory: Concepts*, pp. 65–68. Oxford: Oxford University Press.
- WILLIAMS, B. (1956/57) "Personal Identity and Individuation." *Proceedings of the Aristotelian Society* 57: 229–252.