

# Is personal identity a matter of psychological continuity?

Zak Edwards

November 26, 2014

I aim in this paper to forward a neo-Lockean defense of *psychological continuity* – albeit in a manner perhaps dissimilar to ordinary understanding – as solely constitutive of (any reasonable conception of) *personal identity*.

Since Locke’s formative developments in the field<sup>1</sup>, the metaphysical problem of personal identity has been ordinarily construed as the problem of specifying necessary and sufficient conditions for a relation of co-personality between a plurality of *person-stages*, indexed temporally and distinct as such to varying degrees.<sup>2</sup> Such a co-personality relation, as it has here been presented, is taken often to be a strict Leibnizian identity relation, impervious to degree and satisfying with mathematical precision the criteria for logical equivalence<sup>3</sup>. Precisely this manner of understanding the connectedness between a person’s past and present ‘*selves*’ is, I claim, unnecessarily problematic and born principally out of comfortable inclination. The aim of this paper, therefore, is to present a cogent dismissal of one-to-one, ‘all-or-nothing’ *non-Reductionist*<sup>4</sup> *identity* as *what matters* with concern to *personal identity*, in favour of a Parfitian conception of variable *survival* as the sole significant constituent of any philosophically tenable notion of personal identity.<sup>5</sup>

---

<sup>1</sup>*cf.* Locke [5].

<sup>2</sup>The term *person-stage* is here co-opted from Slors ([14], p. 7).

I use ‘the problem of personal identity’ and ‘the *diachronic* problem of personal identity’ synonymously. One can, I claim, extrapolate a satisfactory deflationary conception of the *synchronic* dual from the account of diachronicity offered in this paper.

<sup>3</sup>Parfit calls this the *determinacy* of identity; *cf.* Parfit [9], p. 15.

Relations referred to throughout as *Leibnizian* are precisely those relations that satisfy Leibniz’ ontological principles (detailed originally in Leibniz [3]), in addition to the standard logical conditions for equivalence. That is:

$$\begin{array}{ll} \forall\varphi\forall\psi\{\varphi R\psi \implies \forall P[P(\varphi) \iff P(\psi)]\} & \text{(Indiscernibility of Identicals)} \\ \forall\varphi\forall\psi\{\forall P[P(\varphi) \iff P(\psi)] \implies \varphi R\psi\} & \text{(Identity of Indiscernibles)} \\ \forall\varphi(\varphi R\varphi) & \text{(Reflexivity)} \\ \forall\varphi\forall\psi(\varphi R\psi \implies \psi R\varphi) & \text{(Symmetry)} \\ \forall\varphi\forall\psi\forall\phi[(\varphi R\psi \wedge \psi R\phi) \implies \varphi R\phi] & \text{(Transitivity)} \end{array}$$

<sup>4</sup>As I have here employed the terms, and with respect to the intended scope of this essay, the *reductionist/non-reductionist* distinction is directly analogous with the *Bundle/Ego theory* distinction, as it is often elsewhere termed: *cf.* Parfit [8], p. 20 ff. The motivation for this terminological distinction is clear; for the Parfitian reductionist, there are no *selves* or *persons* as temporally immutable entities, apart from certain streams of physical and mental events ‘bundled’ together. Conversely, the Ego theory aims to defend the *self* as fundamentally distinct from precisely these events.

One is forced to presuppose a kind of physicalistic monism in writing of Reductionism and non-Reductionism as comprising the entire metaphysical domain of personal identity. Indeed, to provide a refutation of personhood as constituted by an immaterial entity – one thinks firstly of the *soul* – is beyond the scope of this paper. (Interesting and convincing efforts toward this objective have been forwarded; *cf.* Urbaniak and Rostalska’s [11] refutation of Swinburne’s modal argument for the existence of a *soul*.) It is not far-fetched, however, to regard such an entity as an ontological superfluity, provided, of course, that this paper’s conclusions are correct.

<sup>5</sup>As such, the present paper can be considered a further defense of those views espoused in *chapters 10–13* of *Reasons*

Seemingly paradoxically, or at the very least counter-intuitively, *personal identity* as we shall understand it – as its own phenomenon, with due respect to a *common-sense* understanding – will be argued to be constituted *not* by any strict conception of *identity*. Why, then, should we wish to present a non-Leibnizian characterisation of co-personality as a sufficient relation theory of personal identity? Precisely because, I claim, this characterisation functions as a sufficient explanation of the same personal identity problem that Leibnizian identity theories purport to solve.<sup>6</sup> If indeed there is a '*further thing*' – the immutable *self*, the Platonic subject of experiences separate from the *being*, or indeed the Cartesian *Pure Ego*, as referred to by Parfit<sup>7</sup> – *over and above* the sum of one's psychological components which can be dismissed with no loss of explanatory capacity, to postulate such a *thing* is *at best* ontologically redundant.

It shall here be necessary to introduce a tentative distinction between *bodily* and *psychological* continuity.<sup>8</sup> Bodily continuity can be broadly considered as follows: *the spatiotemporal continuity of a human body constitutes personal identity*. Psychological continuity, on the other hand, can *presently* be considered the thesis that personal identity *consists in a continuous psychological relation between past and present person-stages*.

One can consider a number of thought experiments, illuminating both the ostensible problems continuity (both *bodily* and *psychological*) faces and the explanatory supremacy of Parfitian reasoning in dissolving precisely these problems. Perhaps the most simple such experiment is considered by Parfit ([7], p. 497), whom I quote:<sup>9</sup>

Suppose [...] that I need surgery. All of my brain cells have a defect which, in time, would be fatal. But a surgeon can replace all these cells. He can insert new cells that are exact replicas of the existing cells except that they have no defect.

The above consideration, I claim, renders the thesis of personal identity as constituted by bodily continuity untenable. The reason for this is clear: one who should undergo such an operation is, subsequent to the operation, devoid of any bodily continuity with that individual prior to the operation; yet, one would imagine the resultant person to be *precisely the same person* as before, in *every meaningful sense* of personhood. *Apropos* of no apparent convincing counter-argument, I submit that this serves a definitive blow to the theory of bodily continuity. Far from offering any pernicious criticism of Reductionist psychological continuity, the case for it is strengthened.

Precisely by excluding the plausibility of bodily continuity theories, one takes a further step towards the negative thesis that numerical identity *cannot matter* for personal identity. Should I undergo the above operation, there ultimately remains none of my original body, but nonetheless, I remain *me* in

---

and *Persons* [7].

<sup>6</sup>That is, provided even a conception of psychological continuity that incurs a partial dissociation from personal identity.

<sup>7</sup>Parfit [7], p. 231:

Many Non-Reductionists believe that *we are separately existing entities*. [...] On the best known version of this view, a person is a *purely mental* entity; a Cartesian Pure Ego, or spiritual substance.

<sup>8</sup>*Tentative* insofar as I leave the criteria for psychological continuity intentionally vague for the purposes of later revision.

<sup>9</sup>I have simplified Parfit's views here. Parfit proceeds to consider two cases which I omit; one on which the surgeon performs a hundred incremental operations, gradually replacing the brain, and another where all parts of the brain are removed, and replicas are inserted in one operation. The conclusion is the same. As he states earlier in the text, the question of who *I* shall be after the operation is 'empty'; '*These are not here different possibilities, one of which must be true. These are merely different descriptions of the same outcome*' (Parfit [7], p. 302).

the *right kind of way*. Notably, this is analogous to Locke’s example of the *prince and the cobbler*<sup>10</sup>, without the restraint<sup>11</sup> of, for example, Shoemaker’s *Brown-Brownson* case<sup>12</sup>. This invites criticisms of implausibility; Nagel goes as far as to say it’s ‘psychologically impossible to believe’.<sup>13</sup> I maintain that levying a claim of implausibility against a metaphysical thesis does not constitute a counterargument.

Motivated by an experiment imagined by Shoemaker ([12], p. 22), Wiggins ([15], p. 53) reinvents an oft-discussed case of *fission* concerning a person *divided* into two distinct individuals. I quote:

Suppose that we split Brown’s brain and house the two halves in different bodies. [...] We are supposing that the transplanted persons, Brown I and Brown II, claim to remember exactly the same things, that they are equally intelligent, and that they are equally at home in their new bodies. [...] if we say each is the same person as Brown, we shall have to say that Brown I is the same person as Brown II.

Further restricting the domain of continuity, this proves clearly to be a problem for the non-Reductionist psychological account of personal identity; for indeed, if *Brown I* and *Brown II* are psychologically alike in every considerable manner, the non-Reductionist advocate of psychological continuity must account for the identity of personhood in conflict with the logical impossibility of numerical identity between multiple persons. Concerning this case, Parfit remarks ([6], p. 8):

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

Parfit continues to say we should ‘*give up the language of identity*’. Indeed, we can say that what makes us *selves, individuals, persons* over time is just that we are a causally connected and continuous stream of psychological (or perhaps psycho-physical) events; but the use of such language invites conflation with the irreducible, enduring *Ego* posited baselessly by detractors of Reductionism.

Thus, to reiterate, identity is not what matters – in the above *split-brain* case, for instance, the question of who ‘*is*’ *me* (strictly in what I have termed the Leibnizian sense) is not of our principal concern. But, if this is indeed the case, what *can* be said to matter? To answer this question, Parfit introduces the concept of *survival*: and what matters for survival is precisely psychological continuity; that a person’s psychology should be a temporally indexed stage of a *succession on stages* related by *similarity* and *lawful causal dependence*.

---

<sup>10</sup>cf. Locke [5], *Book II, Chapter XXVII*, “Of Ideas of Identity and Diversity”:

For should the soul of a prince, carrying with it the consciousness of the prince’s past life, enter and inform the body of a cobbler, as soon as deserted by his own soul, every one sees he would be the same *person* with the prince, accountable only for the prince’s actions [...].

<sup>11</sup>That is, the *obvious* physical restraint. One need not suggest there is anything *over and above* the physical to find Parfit’s (and indeed Locke’s) characterisation more ontologically parsimonious in contrast to Shoemaker’s.

<sup>12</sup>cf. Shoemaker [12], pp. 23–24:

Two men, a Mr. Brown and a Mr. Robinson, had been operated on for brain tumors, and brain extractions had been performed on both of them. At the end of the operations, however, the assistant inadvertently put Brown’s brain in Robinson’s head, and Robinson’s brain in Brown’s head. One of these men immediately dies, but the other, the one with Robinson’s body and Brown’s brain, eventually regains consciousness. Let us call the latter “Brownson”.

Shoemaker is “strongly inclined” to believe ‘Brownson’ *is indeed Brown*; the significant discrepancy here, in contrast to Locke and Parfit, is that we preserve some perspicuous form of physical continuity. (This perspicuity seems to incur a correct conclusion, albeit from a premise based on intuition and plausibility rather than constitutive integrity.)

<sup>13</sup>cf. Parfit [7], p. 304.

The notion of *survival* can be fortified with the help of Parfit's R-relation. It is tempting to equate the R-relation with psychological continuity; however, we should opt rather to equate it with the more nuanced notion of psychological *connectedness*. For Parfit, psychological connectedness is the holding of particular *direct* intransitive psychological connections; based upon this, therefore, psychological continuity consists in there being a chain of strong overlapping connections.

To illustrate the importance of this distinction, consider the well-known case of the *brave soldier*<sup>14</sup>, advanced by Reid as an objection to Locke. In Reid's thought experiment, the logical property of transitivity is violated on the memory account (*i.e.*, the psychological continuity account). Psychological connectedness avoids Reid's objection simply on account of the fact that it is not firstly bound by transitivity; it does, by necessity, admit of degree. It is certainly conceivable that connectedness, interpreted in this manner, can hold to various degrees, and this manner of interpretation is tenable *simpliciter* and not merely a circumvention of Reid's objection.

The cell replacement and split-brain cases are analogous respectively to one's common intuition concerning the diachronic personhood of *one resultant person* who is *numerically identical* to an initial person, and *multiple resultant persons*, none of which are numerically identical to the initial person. I submit that it has been shown that both cases support survival as the sole tenable factor with regard to substantive questions concerning personal identity. However, there is a further possible consideration, which invites a singular degree of confusion; the possibility of multiple resultant persons, one of which is numerically identical – indisputably so – to the initial person.

In *Reasons and Persons*, Parfit introduces the *teletransportation* experiment, and subsequently a variation thereupon, principally to illustrate the capacity for his reductionist theory to accommodate branching (or *one-to-many*) forms of personal identity, a feature explicitly inconsistent with the non-Reductionist view. In Parfit's original science fiction of teletransportation, a exact blueprint of my body is constructed by a Scanner on Earth, and reconstituted three minutes later by a Replicator on Mars. Consequently, the *person* on Mars is imbued with *my* memories, character traits, *etc.* This straightforward teletransportation – which we may term '*simple*' – incurs the question of sameness of personhood; can it reasonably be said that the person on Mars is *me*? Such a person should exhibit connectedness and continuity of psychology; thus, I advance the positive claim that such an individual *would indeed be me* in any *meaningful, substantive sense*.<sup>15</sup> This position is intuitive, and I suspect commonplace; many would consider this but an effective form of travel. A trivial consequence of simple teletransportation is a further counterpoint against the perspicuous conception bodily continuity.

What we may term the '*complex*' case of teletransportation concerns a certain 'failure' of the

---

<sup>14</sup>*cf.* Reid [10]:

Suppose a brave officer to have been flogged when a boy at school for robbing an orchard, to have taken a standard from the enemy in his first campaign, and to have been made a general in advanced life; suppose, also, which must be admitted to be possible, that, when he took the standard, he was conscious of his having been flogged at school, and that, when made a general, he was conscious of his taking the standard, but had absolutely lost the consciousness of his flogging. These things being supposed, it follows, from Mr. Locke's doctrine, that he who was flogged at school is the same person who took the standard, and that he who took the standard is the same person who was made a general. Whence it follows, if there be any truth in logic, that the general is the same person with him who was flogged at school. But the general's consciousness does not reach so far back as his flogging; therefore, according to Mr. Locke's doctrine, he is not the person who was flogged. Therefore the general is, and at the same time is not, the same person with him who was flogged at school.

<sup>15</sup>This nuanced condition is crucial, for I do not mean to suggest, or present Parfit as suggesting, that 'I' am *identical*, logically so, to *both* resultant persons. This alternative claim would, of course, face the insurmountable criticism of logical flaw.

Scanner on Earth. As a consequence of this failure, my body on Earth is not instantaneously destroyed – I remain on Earth for, say, some length of time greater than three minutes – yet a person is created on Mars, just as before. In this scenario, then, there exists *two* persons, qualitatively identical yet numerically distinct; the Leibnizian faces an immediate problem, as two numerically distinct persons, under the strict condition of logical identity, cannot be the *same* person. Following Parfit, I submit that, congruent with our position regarding the simple case, I would *survive in both persons*.<sup>16</sup> But I would not *be* both persons, nor be one or the other, nor be neither; for, as Parfit explains, ‘*we can suggest that I survive as two different people, without implying that I am these people*’ ([6], p. 203).

The complex teletransportation appears, I claim, to be definitively fatal to the tenability of Leibnizian identity as constitutive of personal identity. However, strict identity does not remain undefended from cases of fission; one such attempt at saving identity comes from Lewis.

In *Survival and Identity* [4], Lewis presents the thesis that identity is *what matters* as the “commonsense platitude”.<sup>17</sup> Towards the cause of defending identity, Lewis attempts to show that, in cases of fission, there exists a multitude of co-located wholly present persons, and thus, *pace* Parfit, there is no destruction nor creation of persons. The *multitude of persons* in typical cases of fission is usually imagined as, but not restricted to, two persons; thus, the proponent of the co-locality of persons is committed *a fortiori* to a host of perplexing theses, perhaps to the thesis that there exists infinitely many co-local persons.<sup>18</sup> For precisely these reasons, I claim we ought regard Lewis’ attempt as unsuccessful.

One such virtue of Parfitian reductionism, then, is its ability to make sense of the otherwise nigh-insurmountable teletransportation dilemma.<sup>19</sup>

Having identified survival as *what matters* with concern to the *substantive* problem of personal identity, a simple deduction can show that Leibnizian identity cannot also *matter* in this Parfitian sense. I propose an extended version of Parfit’s essential argument:

- P1 Leibnizian identity is *one-one* and *does not admit of degree*.
  - P2 What matters for *survival* is *psychological continuity*.
  - P3 Psychological continuity *need not be one-to-one* and *may admit of degree*.
  - P4 What matters for *personal identity* is *survival*.
- 
- C Leibnizian identity is *not what matters* for personal identity. (P1, P2, P3, P4)

P1, of course, is merely an axiomatic consequence of the Leibnizian principle of the *indiscernibility of identicals*. P2 is justified on the simple basis of definition, for Parfitian survival is a *necessary and sufficient condition* for psychological continuity. The third premiss is similarly uncontroversial; one can consist in an R-relation to a past person-stage to varying degree. P4 allows for the greatest room for debate, but, as we have argued at length, the premise is indeed valid. We conclude that psychological continuity, under the conditions of intransitive connectedness, is the sole constituent of personal identity.

<sup>16</sup>I acknowledge that Parfit has advanced the claim that this sentiment is false, but I also entertain the possibility that Parfit used the term *survive* in a manner contrary to his usual usage. What matters for survival is psychological continuity (by Parfit’s second criterion); indeed, Parfit agrees that both such persons would be psychologically continuous. The statement here, while seemingly contrary to Parfit’s text, is, I claim, ultimately Parfitian.

<sup>17</sup>The precise technicalities of both Lewis’ original argument and of the convincing rebuttals offered by Sider [13] and Belzer [1] are perhaps beyond the scope of this paper, and thus I omit an analysis of any significant degree of depth.

<sup>18</sup>In addition, the Lewisian theory of personal identity leaves unclear precisely what would occur in cases of *fusion*.

<sup>19</sup>And, by the same token, any similar case of fission.

'*Personal identity*' is somewhat of an unfortunate misnomer;<sup>20</sup> what we really mean, and what is in our self-interested concern, does not concern *identity* at all, for identity as it *classically defined* is indefensible as constitutive of *personal* identity. It is comfortable to think of identity as continuous and fundamentally unchanging, and because of this the problems arise. It is more tenable to reject this comfortable fiction. Personal identity comprises an artificial notion, the delineation of which is rather arbitrary.

I have shown that the most superficially convincing accounts of personal identity as constituted by non-Reductionist conceptions of identity are ultimately untenable. Certainly, otherwise non-psychological theories of personal identity are patently indefensible. Conceptually, I submit that personal identity is wholly within the explanatory capacity of survival, and, *a fortiori*, one need not concern oneself with *identity* in *any strict sense*; further, based upon these conclusions, I regard the Cartesian Ego in any of its metaphysical conceptualisations as a comfortable fiction, *at best* redundant, that collapses under scrutiny.

## References

- [1] BELZER, M. Notes on Relation R. *Analysis* 56, 1 (1996), 56–62.
- [2] BUTLER, J. *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*. 1736.
- [3] LEIBNIZ, G. W. *Discourse on Metaphysics*. 1686.
- [4] LEWIS, D. Survival and Identity. In *The Identities of Persons*, A. Rorty, Ed., vol. 3. The University of California Press, 1976.
- [5] LOCKE, J. *An Essay Concerning Human Understanding*. 1700.
- [6] PARFIT, D. Personal Identity. *The Philosophical Review* (1971), 3–27.
- [7] PARFIT, D. *Reasons and Persons*. Oxford University Press, 1984.
- [8] PARFIT, D. Divided Minds and the Nature of Persons. In *Mindwaves*, C. Blakemore and S. Greenfield, Eds. Blackwell Publishing, 1987, pp. 19–25.
- [9] PARFIT, D. The Unimportance of Identity. In *Identity: Essays based on Herbert Spencer lectures given in the University of Oxford*, H. Harris, Ed. Clarendon Press, 1995, pp. 13–45.
- [10] REID, T. Of Mr. Locke's Account of our Personal Identity. In *Essays on the Intellectual Powers of Man*. 1785.
- [11] ROSTALSKA, A., AND URBANIAK, R. Swinburne's Modal Argument for the Existence of a Soul. *Philo* 12, 1 (2009), 73–87.
- [12] SHOEMAKER, S. *Self-knowledge and Self-identity*. Cornell University Press, Ithaca, N.Y., 1963.
- [13] SIDER, T. All the World's a Stage. *Australasian Journal of Philosophy* 74, 3 (1996), 433–453.

---

<sup>20</sup>Although we know intuitively precisely what one means by the term.

- [14] SLORS, M. *The Diachronic Mind: An Essay on Personal Identity, Psychological Continuity and the Mind-Body Problem*, vol. 86 of *Philosophical Studies*. Kluwer Academic Publishers, 2001.
- [15] WIGGINS, D. *Identity and Spatio-Temporal Continuity*. Blackwell Publishers, Oxford, 1967.