Do Lefty and Righty Matter More Than Lefty Alone?

Johan E. Gustafsson[1,2,3,4] · **Petra Kosonen**[3,5]

Abstract

Derek Parfit argues that fission is prudentially better for you than ordinary death. But is having more fission products with good lives prudentially better for you than having just one? In this paper, we argue that it is. We argue that, if your brain is split and the halves are transplanted into two recipients (who both have good lives), then it is prudentially better for you if both transplants succeed than if only one of them does (other things being equal). This upshot rules out, among other things, that the prudential value of standing in the relation that matters in survival to multiple people is equal to their average well-being.

Your brain is divided, and each half is transplanted into the body of one of two people whose brains have been fatally injured. Each of the resulting people would, if the transplant succeeds, be psychologically (and phenomenally) continuous with you as you are now.[1] Call the person that would have the left half of your brain 'Lefty' and the person that would have the right half of your brain 'Righty'.[2] Let's assume that each year in Lefty's and Righty's lives would be of equal positive well-being.

---

[1] Wiggins (1967, p. 53) and Parfit (1984, pp. 254–255).

[2] Following Strawson (1970, p. 186).

✉ Johan E. Gustafsson
johan.eric.gustafsson@gmail.com
https://johanegustafsson.net

Petra Kosonen
kosonenpetra@gmail.com
https://petrakosonen.com

1 Department of Philosophy, University of York, York YO10 5DD, UK

2 Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, Gothenburg 405 30, Sweden

3 Population Wellbeing Initiative, University of Texas at Austin, Austin TX 78712, USA

4 Institute for Futures Studies, Stockholm, Holländargatan 13, 101 31, Sweden

5 Department of Philosophy, University of Oxford, Radcliffe Humanities, Woodstock Road, Oxford OX2 6GG, UK

Springer

Consider, first, an outcome where both transplants succeed:

Double Success   Lefty lives for 40 years, and Righty lives for 40 years.

Given the transitivity of identity, you can't be identical with both Lefty and Righty, who are plausibly distinct people.[3] And, since your relation to Lefty is much the same as your relation to Righty, you can't plausibly be identical with one of them but not the other. So you are neither identical with Lefty nor identical with Righty. Even so, Derek Parfit argues that this outcome is prudentially better for you than an outcome where both transplants fail:

Double Failure   Neither transplant succeeds.

Parfit claims that Double Failure is prudentially worse for you than an outcome where one transplant succeeds:

Single Success   Lefty lives for 40 years, and the transplant to Righty does not succeed.

In this outcome, there's only one survivor. And it seems that, in principle, you could survive a brain transplant and that you could survive with just half of your brain. Hence it's plausible that you survive in Single Success and that surviving in Single Success is prudentially better for you than dying in Double Failure.[4] But, if Single Success is prudentially better for you than Double Failure, then your relation to Lefty in Single Success must contain what matters in survival.

Next, note that your relation to Lefty is the same in Double Success as in Single Success—the only difference being the lack of uniqueness. Since the relation that matters is plausibly intrinsic, it shouldn't depend on uniqueness. The addition of Righty should not make it the case that your relation to Lefty no longer contains what matters in survival. And, as we argued earlier, your relation to Lefty in Single Success does contain what matters in survival. So your relation to Lefty in Double Success must also contain what matters in survival. And, by symmetrical reasoning, your relation to Righty in Double Success must contain what matters in survival. Accordingly, Parfit concludes, Double Success should be prudentially better for you than Double Failure.[5]

But how does Double Success compare to Single Success? As we have seen, your relation to Lefty in Double Success contains what matters in survival. And your equivalent relation to Righty in that outcome also contains what matters.[6] Nevertheless, even though both your relation to Lefty and your relation to Righty contain what matters, we may still doubt that Double Success is prudentially better for you

---

[3]   The transitivity of identity is the principle that, if $X = Y$ and $Y = Z$, then $X = Z$.

[4]   Parfit (1984, p. 254, 1995, p. 42).

[5]   Parfit (1971, p. 5, 1984, pp. 261–262, 1993, pp. 24–25, 1995, pp. 42–43).

[6]   Parfit (1971, p. 10, 1984, p. 262, 1993, p. 25).

than Single Success. For example, consider the following (at first glance compelling) view:[7]

> *The Prudential Average View*   Let the *chancy equivalent* of an outcome be a prospect where all fission events are replaced by an even-chance lottery of being any of the fission products. The prudential value of an outcome is equal to the expected prudential value of its chancy equivalent.[8]

In other words, the Prudential Average View regards the prudential value of splitting into Lefty and Right as equal to the expected prudential value of an even lottery between becoming Lefty or becoming Righty—in other words, the average between Lefty's and Righty's well-being.

The Prudential Average View entails that Double Success is prudentially equally as good for you as Single Success, since Lefty and Righty have equally good lives. We will argue, however, that Double Success is prudentially *better* for you than Single Success. Thereby, we show that the Prudential Average View is false.[9]

Consider a case where only the transplant to Righty succeeds but there is a slight mishap in this transplant so that Righty only lives on for 20 years:

> *Half Success*   Righty lives for 20 years, and the transplant to Lefty does not succeed.

The earlier argument that Single Success is prudentially better for you than Double Failure also shows, changing what needs to be changed, that Half Success is prudentially better for you than Double Failure.

Now, consider the following variation, where both transplants succeed overall but (as in Half Success) there is a slight mishap in the transplant to Righty so that Righty only lives on for 20 years:

> *One-and-a-Half Success*   Lefty lives for 40 years, and Righty lives for 20 years.

Your relation to Lefty is the same in One-and-a-Half Success as in Single Success. And it must be prudentially good for you to stand in that relation to Lefty, since Single Success is prudentially better for you than Double Failure. Likewise, your relation to Righty is the same in One-and-a-Half Success as in Half Success. And

---

[7] See Tappenden (2011, p. 302) for a similar view. Since the relation that matters plausibly comes in degrees, one might prefer a weighted average (fission nodes having weighted chancy equivalents). This complication, however, won't matter for our argument.

[8] For a precise account of what counts as fission event, we may use the account of branching in Gustafsson (2019, p. 2322).

[9] Ross (2014, pp. 255–256) presents a sequential fission case where everyone ends up prudentially worse off than they could have been given some compelling assumptions about prudential aggregation. That problem won't be decisive here, however. It seems that, on any plausible view of what matters in survival, the relations that Lefty and Righty stand to each others' futures do not contain what matters. And, if so, any plausible view of prudential aggregation will lead to everyone being prudentially worse off than they could have been in case Lefty and Righty end up facing each other in a Prisoner's Dilemma; see Luce and Raiffa (1957, pp. 94–95) and Tucker (1980).

it must be prudentially good for you to stand in this relation to Righty, since Half Success is prudentially better for you than Double Failure. So, from a prudential perspective, One-and-a-Half Success differs from Single Success only in that you also stand in a relation to Righty which contains what matters in survival—and this, as we have seen, is prudentially good for you. Since the relation that matters in survival is plausibly intrinsic, the prudential value of standing in that relation to a future person should not be diminished by your standing in that relation to some other person.[10] Accordingly, One-and-a-Half Success has everything that is prudentially good for you in Single Success and, in addition, you stand in a relation to Righty that is prudentially good for you. Therefore, One-and-a-Half Success should be prudentially at least as good for you as Single Success.[11]

---

[10] More generally, we are committed to the claim that the (final) prudential value of standing in the relation that matters in survival to a future person depends only on the intrinsic features of that relation, your life, and that person's future life—and not on what other future people you stand in that relation to. This invariance claim may seem to commit us to a prudential analogue of the Repugnant Conclusion:

> *The Prudential Repugnant Conclusion*   For any person with a life of high quality, there is some number of people who have lives that are barely worth living such that it would be prudentially better to stand in relation R (the relation that matters in survival) to each of them than to stand in relation R to the person with the high quality life.

See Holtug (2001, p. 60, 2010, p. 122) who states a version where—instead of the relation that matters holding to people with very low well-being—the relation that matters holds to a very low degree to a large number of people with high well-being; and, for the standard Repugnant Conclusion, see Parfit (1984, p. 388). Note that the invariance claim does not commit us to the additivity of prudential value—that is, it doesn't commit us to the following view:

> *The Prudential Total View*   The prudential value of standing in relation R (the relation that matters in survival) to some people is equal to the sum of the well-being of the future part of their lives.

See Holtug (2001, p. 55, 2010, p. 118); and, for a person-stage based variation, see Williams (2014, p. 406). Even if A and B both have intrinsic value, we need not conclude that the intrinsic value of the combination of A and B is the sum of their intrinsic values. See Moore (1903, p. 28). Hence, given that the invariance claim does not commit us to the additivity of prudential value, it doesn't commit us to the Prudential Repugnant Conclusion. Consider, for instance, a kind of perfectionist who thinks that some very good lives are lexically better than any number of mediocre lives (compare Parfit 1986, pp. 161–164). Even if such views are combined with the invariance claim, they can still avoid the Prudential Repugnant Conclusion. Those views can consistently agree that the prudential value of each person is invariant but reject that the prudential values for all people should be added up to get the prudential value of the whole. Moreover, even if we were to get the Prudential Repugnant Conclusion (given some further plausible assumptions), this needn't be a problem: Even though the Prudential Repugnant Conclusion is counter-intuitive, it may still be true. See Zuber et al. (2021).

[11] Note that this claim by itself rules out the Prudential Average View. So, in order to not assume the point at issue against that view, we didn't merely assume this claim, we argued for it. The argument can be generalized to support the following principle:

> *The Prudential Mere-Addition Principle*   If you stand in relation R (the relation that matters in survival) to the same people in outcomes X and X^+ and each of these people have at least as good lives in X^+ as in X except that, in X^+, you also stand in relation R to an additional person who has a good life, then X^+ is prudentially at least as good for you as X.

Basically, this principle states that the mere addition of an extra person to which you stand in the relation that matters should be prudentially at least as good for you as not adding that person. Ross (2014, p. 257) proposes an analogous principle for the addition of bad lives.

Next, compare One-and-a-Half Success to Double Success. From a prudential perspective, the only difference between these outcomes is that one of the people to which you stand in the relation that matters (namely, Righty) has a better life in Double Success than in One-and-a-Half Success. Since the relation that matters in survival is plausibly intrinsic, the prudential value for you of standing in the relation that matters to Lefty shouldn't be affected by the change in Righty's life. And it is prudentially better for you to stand in the relation that matters to Righty in Double Success than to Righty in One-and-a-Half Success, because Double Success is better for Righty than One-and-a-Half Success. Consequently, Double Success should be prudentially better for you than One-and-a-Half Success.[12]

Now, since One-and-a-Half Success is prudentially at least as good for you as Single Success and, moreover, Double Success is prudentially better for you than One-and-a-Half Success, it follows by the transitivity of *prudentially at least as good as* that Double Success is prudentially better for you than Single Success.[13]

Among other things, this result rules out the Prudential Average View. We can show that, other things being equal, having more fission products with good lives is prudentially better for you.[14] Given that each person involved has a good life, fission is not only prudentially better for you than death, it is (other things being equal) prudentially better for you than survival without fission.[15]

---

[12] This argument can be generalized to support the following principle:

The Strong Prudential Pareto Principle If (i) you stand in relation *R* (the relation that matters in survival) to the same people in outcomes *X* and *Y*, (ii) and, for each of these people, *X* is at least as good as *Y*, and (iii) *X* is better than *Y* for at least one of them, then *X* is prudentially better for you than *Y*.

[13] The transitivity of *prudentially at least as good as* is the principle that, if *X* is prudentially at least as good as *Y* and *Y* is prudentially at least as good as *Z*, then *X* is prudentially at least as good as *Z*.

[14] With the help of the Prudential Mere-Addition Principle (see note 11) and the Strong Prudential Pareto Principle (see note 12), we can generalize the argument to show that having more fission products with good lives is always prudentially better for you, other things being equal. To show that it's prudentially better for you (other things being equal) to stand in relation *R* (the relation that matters in survival) to an additional person with a good life at well-being level *l*, suppose that (other things being equal) that additional person had a slightly worse but still good life at well-being level l^-. By the Prudential Mere-Addition Principle, we find that it's prudentially better for you to stand in relation *R* to the additional person at well-being level l^- than not to stand in relation *R* to that person other things being equal. Then, by the Strong Prudential Pareto Principle, we find that it's prudentially better for you to stand in relation *R* to that person if their life is at well-being level *l* rather than l^- other things being equal. Then, by the transitivity of *prudentially at least as good as*, we find that it's prudentially better for you to stand in relation *R* to the additional person at well-being level *l* than not to stand in relation *R* to that person, other things being equal.

[15] This makes the prospect of uploading (that is, the prospect of becoming a digital person) more attractive, since being digital allows you to easily split into many (and for the resulting digital people to split further, and so on). And, if we accept the Prudential Total View, the prospect of uploading followed by an explosion of splits is prudentially good in the extreme.

References

Gustafsson, J. E. (2019). Non-branching personal persistence. *Philosophical Studies, 176*(9), 2307–2329.

Holtug, N. (2001). The repugnant conclusion about self-interest. *Danish Yearbook of Philosophy, 36*(1), 49–68.

Holtug, N. (2010). *Persons, interests, and justice*. Oxford: Oxford University Press.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.

Moore, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.

Parfit, D. (1971). Personal identity. *The Philosophical Review, 80*(1), 3–27.

Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.

Parfit, D. (1986). Overpopulation and the quality of life. In P. Singer (Ed.), *Applied Ethics* (pp. 145–164). Oxford: Oxford University Press.

Parfit, D. (1993). The indeterminacy of identity: A reply to Brueckner. *Philosophical Studies, 70*(1), 23–33.

Parfit, D. (1995). The unimportance of identity. In H. Harris (Ed.), *Identity: Essays based on Herbert Spencer lectures given in the University of Oxford* (pp. 13–45). Oxford: Clarendon Press.

Ross, J. (2014). Divided we fall: Fission and the failure of self-interest. *Philosophical Perspectives, 28*(1), 222–262.

Strawson, P. F. (1970). Chisholm on identity through time (a response). In H. E. Kiefer & M. K. Munitz (Eds.), *Language, belief, and metaphysics* (pp. 183–186). Albany, NY: State University of New York Press.

Tappenden, P. (2011). Expectancy and rational action prior to personal fission. *Philosophical Studies, 153*(2), 299–306.

Tucker, A. W. (1980). A two-person dilemma. *The UMAP Journal, 1*(1), 101.

Wiggins, D. (1967). *Identity and spatio-temporal continuity*. Oxford: Blackwell.

Williams, J. R. (2014). Nonclassical minds and indeterminate survival. *The Philosophical Review, 123*(4), 379–428.

Zuber, S., et al. (2021). What should we agree on about the repugnant conclusion? *Utilitas, 33*(4), 379–383.