

Parfit on Personal Identity in Lewis'
Metaphysics:

How a broader conception of 'being-the-same-person-as' can help respond to Humphrey-style objections

Sheridan Cunningham, University of Toronto

David Lewis is famous among philosophers for proposing a metaphysical picture of reality as made up of discrete, causally closed spatio-temporal worlds. His account treats individuals as worldbound, meaning they do not have identity across possible worlds. Its plausibility therefore relies on the success of Lewis' Counterpart Theory, which is a method to analyze the truth of counterfactuals about people in the actual world. Alvin Plantinga, in "Transworld Identity or Worldbound Individuals?", presents two objections to Counterpart Theory containing the implicit premise that personal identity is equivalent to 'being-the-same-person-as'. One reason to suspect this equivalence does not hold is Parfit's distinction between our intuitions about survival and identity. Drawing on Parfit's distinction, I suggest a way of conceptualizing of counterparthood as 'being-the-same-person-as', while distinguishing this relation from identity. Finally, I show how this conceptualization can neutralize Saul Kripke's famous 'Humphrey Objection' to Counterpart Theory in *Naming and Necessity*.

Sophia XIV

Lewis defends a view known as modal realism, which states that possible worlds are maximally spatiotemporally interrelated wholes, with the same ontological status as the actual world. Possible worlds are metaphysical constructs, which represent ways the actual world could be. Most philosophers argue possible worlds are abstract, and are such things as sets of consistent propositions. Lewis' account instead holds that possible worlds are just as real as the world in which you are reading this paper. The term 'actual' is a merely indexical term, referring to the world in which the speaker happens to be talking. These worlds are defined by spatial and temporal relations. Anything that is spatially temporally related to anything else is a part of the same world as it. Consequently, individuals can only exist at a single possible world, or else they would stand in spatiotemporal relations to objects at other possible worlds, which would violate the maximal definition of a world. For instance, if I exist in the actual world, but also exist (by being identical with a thing that exists) in a possible world where the Allies lost WW2, then I would stand in temporal relations to events in that world, and spatial relations to people in that world. That would violate the definition of a possible world, because then those two worlds would be spatiotemporally related, and would be the same world. In other terms, on Lewis' account, identity is worldbound¹.

Lewis argues counterfactuals can be analyzed by examining the nearest possible world to the actual world in which the antecedent holds, and asking whether the consequent holds. Counterpart Theory explains how we can analyze counterfactuals about individuals that exist at our world, given that they do not exist in other possible worlds. An individual's

counterpart at another possible world is the individual at that world which is most relevantly similar to the individual, if a relevantly similar individual exists. This relation is not transitive or symmetric, unlike the identity relation. It is reflexive: every individual is their own counterpart in their possible world². What all of an individual's counterparts have in common is that individual's essence, meaning that essence and counterparthood are interdefinable³. When analyzing counterfactuals about what would be the case for an individual, we examine what is the case for their counterpart in that relevant possible world.

Alvin Plantinga has offered two objections to the theory of worldbound identity, as supplemented by counterpart theory, which target the consequences of defining what is possible for a person in terms of a non-identical person.

Firstly, according to counterpart theory, an object's essence is what is common across its counterparts⁴. Consider two properties I might have: the property of being self-identical, and the property of being identical with me (meaning, the person I am actually am)⁵. Since all of my counterparts are self-identical, meaning they are all identical with themselves, I am essentially self-identical⁶. However, my counterparts do not have the property of being identical with me (meaning, the person I actually am)⁷. What Plantinga considers "genuinely paradoxical" is the consequence that, since I am not essentially identical with myself, "I could have been someone else [...] distinct from me"⁸. The force of Plantinga's objection comes from this being a deeply unintuitive consequence.

Sophia XIV

Plantinga then argues that worldbound identity entails that objects are intensely modally fragile, meaning they would be different objects given even the smallest change in what obtains⁹. Consider a possible world that is exactly like the actual world, except one more raindrop falls during a storm. My counterpart in that world is not identical to me, by definition. Plantinga infers that, therefore, my counterpart in that world is not the same person as the actual me¹⁰. Thus, the statement “if another raindrop had landed, I would be a different person” is true¹¹. Put generally, if anything had been other than it actually is, I would have been a different person, or I would not exist (if I do not have a counterpart at that possible world)¹². Plantinga suggests, plausibly, that this contradicts our basic intuitions about existence¹³. After all, what does the number of raindrops that fall have to do with the person I am?

Both of Plantinga’s objections contain the implicit assumption that if a person is not identical to someone, they are not the same person as them. Plantinga’s first objection moves from the premise that “I am not essentially identical with the person I actually am” to the premise that “I could be a distinct person”. This only follows on the assumption that identity is a necessary condition for being-the-same-person-as, which takes the form of the hidden premise “if I am not essentially identical with the person I actually am, I am not essentially the same person as the thing I actually am”. The second objection can be reconstructed to recognize the implicit premise as such:

Sophia XIV

P1 If another raindrop had fallen, the person I would be would not be identical with the person I actually am

P2 If the person I would be would not be identical with the person I actually am, I would be a different person

C If another raindrop had fallen, I would be a different person

These premises hard to articulate, because on Lewis' view, what would be true of me if another possible world had obtained is what would be true of my counterpart, who is not identical with me. On counterpart theory, P1 would hold because in the world where the raindrop falls, my counterpart would not be identical with me, defined as the person I am in the actual world. P2, on the other hand, only holds if personal identity is equivalent with 'being-the-same-person-as'. If we assume that P2 holds, then Plantinga's argument follows logically. We might be inclined to assume this equivalence. Lewis himself appears to, by stating that "Your counterparts resemble you closely in content and context in important respects. [...] But they are not really you"¹⁴. However, there are also reasons to distinguish these two relations.

One reason to do so is Parfit's argument that when we question our intuitions about personal identity, we are actually asking what is sufficient for our continued existence, which is a more important question. As both Lewis and Plantinga note, problems in time are often analogous to problems in modality. Parfit gives a possible solution to the question of identity through time in "Personal Identity", by arguing that our intuitions about survival often diverge from what would be the

case if personal identity was our main concern. For instance, we can conceptualise one person surviving as two people. Imagine splitting one brain in half, and connecting each half to a different body. If the person survives at all, which seems plausible, then they must survive as both people, because there is no rational reason to say they survive as one and not the other¹⁵. This implies personal identity is not necessary for survival, as an identity relation between the initial person and the two product-people would require the product-people also be identical to one another, which they would not be¹⁶. Parfit goes on to argue the relations we care about when assessing survival are those of psychological continuity, or the causal continuity between psychological states¹⁷. Psychological continuity is not one-to-one; it can hold between multiple beings. There is nothing essential to psychological relations (such as remembering an experience), that requires they only exist between people who are identical with one another; therefore, psychological continuity could hold between people who are not identical with one another¹⁸. Survival differs also from personal identity in that it is a matter of degree¹⁹. Parfit argues there are cases where a person appears intuitively to partially survive a situation, such as if they merged with another person²⁰. ‘Psychological connectedness’ describes the degree to which two individuals are directly psychologically related, and therefore the extent to which a person survives. Unlike psychological continuity, it is not transitive²¹.

Both Lewis’ definition of counterparts, and Parfit’s definition of survival in terms of psychological connectedness, and intransitive and context-dependent. Lewis argues that counterparthood is a more powerful tool than identity for

describing what could be the case, because it is not transitive or symmetric, unlike identity. Since counterparthood is determined by *relevant* similarity, it is also context dependent on Lewis' view. Parfit's psychological connectedness is also formulated as an alternative to the identity relation, and is intransitive for similar reasons. Since it affords of degree "the drawing of these distinctions can be left to the choice of the speaker and be allowed to vary from context to context."²² This suggests that underlying both cases is the intuition that what matters to personhood is not a strict identity relation, but a more complex, context-sensitive web of relations.

However, psychological connectedness is not directly analogous to counterparthood, for a few reasons. Firstly, psychological connections cannot be used to describe how counterparts are related to one another. Psychological connections are causal, and on Lewis' account there cannot be causal relations between worlds. Secondly, survival is a matter of degree, whereas something cannot be 'more or less' a counterpart of something else. Though an object's counterpart on another world might change depending on the context of analysis, it always is or is not a counterpart, with no middle ground. Thirdly, an individual has at most one counterpart at a possible world, whereas psychological connectedness can hold between one person and multiple other individuals at the same time.

That being said, Parfit's distinction suggests a framework for understanding how and why we might distinguish between 'being-the-same-person-as' and personal identity. It shows there are at least a few conceivable cases where identity is unable to capture our intuitions about a person's continued existence. So,

Sophia XIV

what could ‘being-the-same-person-as’ be, if not being personally identical? In order to prevent the inference of P2, it must be the case that even if we are not identical with our counterparts, we are the same person as them. This means it is a judgement of relative similarity, rather than ontological fact. This is not entirely at odds with our intuitions, as Parfit demonstrates with respect to sameness through time. Moreover, it is not uncommon to hear someone say “I was a different person back then” about their misguided youth, or “they could be the same person” about two people with similar personalities. This suggests that we have a commonplace conception of personhood which is distinct from personal identity. On the analysis I am suggesting, these two comments can be taken seriously. Consider the first claim. If the speaker considers their decision-making processes to be relevant to who they are as a person, and they have changed significantly in them over time, then, they could become sufficiently relevantly dissimilar from their past self as to become a different person. In the second case, the statement could be translated as an assertion that there could be a possible world where two similar people share the same counterpart. In other words, two people are so similar, that there is a close possible world where the most relevantly similar person to each of them is the same.

Let us apply this distinction to Kripke’s famous ‘Humphrey objection’ to counterpart theory, which is often taken to be one of the most damning responses to Lewis. The structure of the objection is similar to Plantinga’s²³. Consider the counterfactual ‘if he had pursued a grassroots voter mobilization strategy, Humphrey would have won the election’. Humphrey’s counterpart on the nearest possible world where

Humphrey pursues this strategy is not Humphrey, but “someone else” we will call HumphreyG²⁴. Kripke argues that Humphrey probably “could not care less” about how successful HumphreyG is; because Humphrey cares about whether he would have won, not about whether some similar person would have won²⁵.

Our analysis accepts that Humphrey is not identical to HumphreyG, but does not accept that HumphreyG is therefore a different person. This takes the sting out of the objection, because if Humphrey is the same person as HumphreyG, Humphrey should care about the results of HumphreyG’s election. After all, what we care about is being the same person, not the more technical matter of identity.

Despite this promising neutralisation, there are two main questions that remain for my suggested modification of Counterpart Theory. Firstly, does the rejection of the equivalence between personal identity and being-the-same-person-as fully deal with Plantinga’s objections, or do the objections arise on the grounds of identity alone? Plantinga suggests his first objection remains, because the modified Counterpart Theory still entails ‘x is essentially self identical’ is not coextensive with ‘x is essentially identical with x’. Alone, it is hard to judge what damage this does to Lewis’ theory. On one hand, this consequence appears to have an argumentative status similar to the ‘incredulous stare’, since it merely articulates a fairly unavoidably odd consequence of worldbound identity and Lewis’ definition of essence. After being robbed of its consequences for personhood, the amount of unintuitive true counterfactuals it produces is seriously restricted. Moreover, Counterpart Theory is a necessary part of Lewis’ overall

metaphysical picture; meaning that on Lewis' way of thinking, giving up some plausibility for the problem-solving benefits it provides is a valid trade-off. Secondly, the question remains of whether this modification to Lewis' theory has unwanted consequences for his overall metaphysical picture. To this question, I am more confident that any consequences would be limited. The conception of 'being-the-same-person-as' retains all the useful features of counterparthood; it merely adds a psychological dimension to the picture, by recognising the ways in which it aligns with our intuitive, day-to-day conception of personhood.

Notes

1. Plantinga, *Essays in the Metaphysics of Modality*, 2-3.
2. Lewis, "Counterpart Theory and Quantified Modal Logic," 114.
3. Ibid.
4. Plantinga, *Essays in the Metaphysics of Modality*, 11.
5. Ibid, 13.
6. Ibid.
7. Ibid.
8. Ibid, 14.
9. Ibid.
10. Ibid, 15.
11. Ibid.
12. Plantinga also suggests a second, and more existentially frightening result: if another world had obtain, the person that I am would not exist. "I" might exist, in that I would have a counterpart who would exist, leading to it being true that 'if another raindrop had fallen, I would exist'. However, since that counterpart would not be identical with the me who actually exists, it is also true that 'if another raindrop had fallen, the person who I actually am would not exist'. This extension of the objection does not apply to Lewis' modal realist ontology. The person who I actually am would still continue to exist, because all possible worlds exist and have the same status as the actual world. The

person I actually am would not be actual, and extra-raindrop-Sheridan would be actual; however, actuality is merely indexical, making it far less existentially worrying.

13. Ibid.
14. Lewis, "Counterpart Theory and Quantified Modal Logic," 114.
15. Parfit, "Personal Identity," 7-9.
16. Ibid.
17. Ibid, 12.
18. Ibid, 15.
19. Ibid, 18-19
20. Ibid.
21. Ibid, 21.
22. Parfit, "Personal Identity," 25.
23. Kripke, *Naming and Necessity*, 45.
24. Ibid.
25. Ibid.

References

Kripke, Saul. *Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press, 1972.

Lewis, David K., and The Journal of Philosophy, Inc. "Counterpart Theory and Quantified Modal Logic." *Journal of Philosophy* 65, no. 5 (1968): 113–26. <https://doi.org/10.2307/2024555>.

Parfit, Derek. "Personal Identity." *The Philosophical Review* 80, no. 1 (1971): 3–27. <https://doi.org/10.2307/2184309>.

Plantinga, Alvin. *Essays in the Metaphysics of Modality*. Edited by Matthew Davidson. Oxford University Press, 2003. <https://doi.org/10.1093/0195103769.001.0001>.