

# Lewis' Personal Identity

**Paul So**

University of Maryland

## **ABSTRACT**

In their debate about personal identity, Derek Parfit (1971) thinks that our common sense notion of personal identity cannot be preserved, since formal identity fails to preserve it in the fission paradox scenario; David Lewis (1976) thinks that it can be preserved with tense-identity, according to which two continuant persons C1 and C2 are tense identical at T1 iff they share the same person stages at T1. I argue that Lewis' notion of tense identity falls short of its intended goal to capture our common sense notion of personal identity, which is committed to numerical identity, not mere qualitative identity. Whereas for Lewis, two continuant persons C1 and C2 are *prima facie* numerically identical insofar as they are indiscernible, sharing the same person-stages, this indiscernibility relation between C1 and C2 only makes them qualitatively identical, rather than actually numerically identical.

Keywords: David Lewis, Derek Parfit, Personal Identity, Tense Identity, Numerical Identity, Qualitative Identity, Common Sense Notion of Personal Identity, Formal Identity, Survival, and Fusion Paradox

In this essay, I will survey the debate between Derek Parfit and David Lewis on personal identity and then identify the source of disagreement between Parfit and Lewis, which turns on whether or not our common sense notion of personal identity over time can be preserved given Parfit's argument against it. Parfit thinks that our common sense notion of personal identity cannot be preserved, since formal identity fails to preserve it, but on the other hand, Lewis thinks that it can be preserved with tense-identity (I will explain this term later). I will argue that Lewis' notion of tense identity falls short of its intended goal to preserve our common sense notion of tense identity. While Lewis' notion of tense identity does preserve some of our common sense notion of personal identity over time, it is insufficient in capturing numerical identity. I will not state any opinions or arguments about

## compos mentis

Parfit's argument against personal identity overtime; his argument only serves as a context for understanding Lewis' argument for tense identity.

### **PARFIT VS. LEWIS**

Parfit points out an incompatibility between the formal nature of identity and the criterion of psychological continuity: the formal nature of identity requires transitivity, but psychological continuity can allow intransitive relations among psychological states, which obtain in the fission paradox (Parfit 1971, 8). For example, if my body is badly damaged in T1 and my right and left hemispheres are transplanted to two separate bodies respectively in T2, then two versions of me in T2 are psychologically continuous with me in T1, I am psychologically continuous with them, but they are not in psychological continuity with each other. Parfit argues that fission cases show that identity, which requires transitivity, does not matter, instead what matters is survival. Parfit writes: "The relation of the original person to each of the resulting people contains all that interests us-all that matters-in any ordinary case of survival. This is why we need a sense in which one person can survive as two." (Parfit 1971, 10). In other words, what matters is that two persons in T2 survive me from T1: they survive me insofar as they have my memories, personalities, beliefs, preferences, and values. In this sense, their existence are my legacy.

Lewis, on the other hand, disagrees with Parfit. While Lewis agrees that there is an incompatibility between identity and psychological continuity, identity does not need to be completely rejected. Lewis writes:

We can agree with Parfit (and I think we should) that what matters in questions of personal identity is mental continuity or connectedness, and this might be one-many or many-one, and admits of degree. At the same time we can consistently agree with common sense (and I think we should) that what matters in question of personal identity- even in the problem cases-is identity. (Lewis 1976, 19)

Instead, Lewis provides what seems like an another version of identity that is compatible with psychological continuity and it could handle intransitive cases such as the fission paradox. This notion of identity is that persons consist of temporal instances called person-stages. Person-stages are basically temporal moments or instances of a single person's life. These person-stages are psychologically related

So

to each other by mental-connectedness (i.e. causal dependence of mental states) and mental-continuity (i.e. connections among mental states over time) (Lewis 1976, 20). Lewis thinks that a person is essentially an entire interrelation of person-stages, which he calls *continuant person* (Lewis 1976, 21). This is where Lewis introduces a distinction between the R-relation and the I-relation.

According to Lewis, the R-relation is the relation among person-stages and this relation consists of mental connectedness and continuity among person-stages (Lewis 1976, 20). For example, in T1 I am concerned whether or not I will survive. What this means is whether or not there is some future person-stage that is R-related to a person stage of me at T1. The I-relation, on the other hand, is a relation between *continuant persons*. C1 (*continuant person*) and C2 are I-related if and only if all their person-stages are R-related to each other. What this means is that C1 and C2 are identical to each other if they share all the same person-stages (Lewis 1976, 21). Lewis goes as far to claim that he “cannot tolerate any discrepancy in formal character between the I-relation and the R-relation” (Lewis 1976, 21). He explicitly states that “the I-relation is the R-relation ” (Lewis 1976, 22). The kind of discrepancy Lewis has in mind is that it is possible that the I-relation remains transitive whereas the R-relations are not transitive.

Lewis thinks that this notion of identity, which is based on the distinction between the I-relation and the R-relation, can handle the fission paradox. The fission paradox is that in a hypothetical scenario, a single *continuant person* in T1 can split into two *continuant persons* in T2, both of which are psychologically continuous with that single person in T1 and vice versa, but they are not psychologically continuous with each other. In terms of the R-relation, a single person’s person-stage in T1 is R-related to two person’s person-stages in T2 and vice versa. However, two persons’ person-stages in T2 are not R-related to each other. In terms of I-relation, two *continuant persons* in T2 are I-related to a *continuant person* in T1 insofar as both had shared a person-stage with a *continuant person* in T1 prior to fission. However, two *continuant persons* in T2 are not I-related to each other. Lewis points out that there is intransitivity in both the I-relation and the R-relation, which shows that Lewis’ notion of identity handles the fission case pretty well insofar as there is no discrepancy between the I-relation and the R-relation (Lewis 1976, 24–26).

However, there is a question about how we are suppose to count persons in fission cases. One problematic possibility is that there are in fact two *continuant persons* all along whose person-stages happen to overlap at a certain point until fission at T2. In fact, this description is quite compatible with Lewis’ distinction

between the I-relation and the R-relation, so his distinction cannot rule out the description that some continuant person in fact happens to be two continuant persons who happen to share all the same person-stages. Furthermore, this description leads to what Lewis calls overpopulation: there are more continuant persons than there are person-stages. In other words, prior to fission (T2), there are two continuant persons C1 and C2 whose person-stages happen to overlap (Lewis 1971, 25).

Lewis finds a way out of this problem by presenting a new term: tense-identity. Whereas formal identity requires strict identity across time, tense-identity allows identity at a certain time (Lewis 1976, 26). For example, in the fission paradox, two persons in T2 are not formally identical to a single person in T1, since there is an intransitive relation among them. Formal identity requires transitive relations among person-stages across time. However, tense identity allows that two continuant persons are identical (or I-related) to each other at a certain time if they share all the same person-stages that are R-related to each other. However, in T2, there is no longer tense identity between continuant persons.

According to Lewis, there are three conditions for tense-identity: "This is the relation that holds between continuants C1 and C2 if and only if (1) they both exist at some time no later than *t*, (2) at any time no later than *t*, either both exist or neither does, and (3) at any time no later than *t* when both exist, they have exactly the same stages" (Lewis 1976, 28). The relation that Lewis mentions is tense-identity and "*t*" refers to a moment when fission transpires. In the pre-fission scenario, C1 and C2 both exist no later than *t* and both share exactly the same person-stages prior to *t*. Given that two continuant persons in the pre-fission scenario satisfy (1)-(3), both C1 and C2 are identical (insofar as they are identical at a certain time) at any time prior to *t*, but not identical after *t*. Both C1 and C2 became two separate persons at *t*.

In the light of the debate between Lewis and Parfit, it is important to ask if there really is a genuine disagreement between them. After all, when Lewis decides to provide tense-identity as an alternative to formal identity, it seems that he is agreeing with Parfit that formal identity does not matter and survival matters. If one tries to be sympathetic to Parfit's view, it is difficult to see how Parfit would disagree with Lewis' tense identity. Parfit may agree with Lewis that two continuant persons C1 and C2 share all the same person-stages prior to fission, but Parfit might reply that tense-identity fails to capture our common sense view of personal identity.

The real disagreement between Parfit and Lewis is whether or not our common sense view of personal identity can be preserved at all given Parfit's argument.

So

On one hand, Lewis thinks that our common sense view of personal identity can be preserved by his account of tense-identity. On the other hand, Parfit thinks that only formal identity is an appropriate means to preserve personal identity over time, which fails to do so. Lewis would disagree with Parfit that the failure of formal identity to account for our common sense intuition of personal identity over time is the failure of common sense intuition of identity. Parfit would probably think that tense-identity cannot preserve our common sense intuition of identity. Parfit might think that tense-identity merely redescribes survival.

In fact, in Parfit's response to Lewis is that he fails to establish his case that identity matters. Parfit agrees with Lewis that what matters is that C1 has a present person stage that is R-related to a future person stage. However, Parfit argues that in order to show that identity matters Lewis has to show that a future person-stage belongs to C1 just as a present person-stage belongs to C1. However, in the fission scenario, it happens to be the case that a future person-stage does not belong to C1, but rather it belongs to some other continuant person who merely survives C1. A future continuant person has a person-stage that is R-related to that of C1, but his person-stage does not belong to C1 in the same way C1's person stage belongs to him (Parfit 1976, 94-95). In other words, Parfit thinks that a future person-stage should not only be R-related to me but it also has to belong to me in the same way my present person stage belongs to me. However, in the fission case, a future person stage cannot belong to me the same way my present stage belongs to me.

### **COMMON-SENSE SAVED?**

While I will not defend Parfit's position that identity does not matter, I do think that Lewis gives an unconvincing argument to support his case. In this sense, I agree with Parfit that Lewis does not succeed in supporting his claim that identity matters. My main argument is that Lewis' tense identity does not succeed in preserving our common sense intuition of identity. I will begin with identifying our common sense intuition of personal identity.

The common sense intuition of personal identity is that we are a single unified person who persists over time. The person at T2 is the same person as me at T1, this sameness is *numerical identity* rather than qualitative identity. Some person at T2 may resemble me at T1 in terms of physical appearance, intelligence, personality, beliefs, memories, and such, but it is logically conceivable that such a person is not me but rather my clone. My clone would be qualitatively identical to me, but not

numerically identical to me. If we consider the fission case, my cerebral hemispheres could be divided and then transplanted in different bodies, which leads to two separate persons who are not numerically identical to each other or to myself prior to split-brain transplant. These separate persons may have my memories, beliefs, personalities, and such, but, so far, they are only (partially) qualitatively identical to me prior to fission. They share my psychological qualities, but they are still not numerically identical to me.

Parfit argues that as long as these persons survive me insofar as they continue to preserve essential qualities from me, that is all that matters. Parfit seems to be content with qualitative identity in cases of survival and thinks that numerical identity does not matter. Lewis agrees with Parfit that, strictly speaking, survival matters and that formal characteristics inherent in numerical identity do not matter, but he seems to think that we can have another notion of identity called tense-identity which functions similarly to that of numerical identity under certain conditions. As long as C1 and C2 exist prior to fission and share all the same person-stages that are R-related to each other, then they are I-related to each other.

While Lewis' tense identity does allow us to count C1 and C2 as a single person prior to the moment of fission, it also counts C1 and C2 as two persons at the moment of fission. This seems like a virtue of tense-identity, Lewis tests the virtue of tense identity by applying it to the case of Methuselah. Methuselah is a biblical character who lived for 969 years, but by the time he lived to 969 he probably wouldn't remember what he was like at the age of 100. If one divides Methuselah's age in 100 years each in terms of person-stages, then there are continuum ways to divide Methuselah's age in 100 years. One can start from Methuselah's birth to his 100th year or start from his 50th year to 150th. If this is the case, then there are infinite persons named Methuselah. Given that Methuselah at the age of 969 forgets what he was like at the age of 100, his person-stages are not transitively R-related and I-related. Lewis says that his tense-identity can solve this problem, since all continuant persons share the same person-stages at a certain time, so there is in fact one Methuselah rather than an infinity of them (Lewis 1976, 30–31).

Tense identity seems to give us an answer we want to hear in both the case of fission and Methuselah. However, there are some objections to consider:

First, while tense-identity might generate answers that we want to hear in both fission cases and the case of Methuselah, it does not capture the essence of our common sense intuition of personal identity. The common sense intuition of our personal identity is that my present self is the same as my future self insofar as

So

both are numerically identical to each other. For many personal identity theorists, what underpins this numerical identity between our present and future selves is psychological continuity. If my future self's mental states are an extension of my present self's mental states, then both are the same person.

On the other hand, Lewis' tense identity amounts to something quite different. It does not give us numerical identity between continuant persons, but rather qualitative identity between continuant persons. Strictly speaking, C1 and C2 are not numerically identical in the formal sense (i.e. transitivity, symmetry, and reflexivity), but they are qualitatively identical insofar as they share the same temporal properties (such as person stages) that renders them *prima facie* indiscernible. The reason why I use the term "prima facie indiscernible" is that C1 and C2 appear to be indiscernible because they share the same temporal parts, but they are still logically distinct from each other. Metaphysically, C1 and C2 are inseparable by virtue of sharing the same temporal parts, but logically they are still distinct from each other since logically they can exist independently of each other in the fission case. Just because C1 and C2 are inseparable from each other, it does not follow that we should logically treat them as a single entity.

Consider the following analogy: It is the case that programs are inseparable from hardwares, but logically they are distinct from each other. Softwares consist of the series of binary codes that consists the number 0 and 1. The physical hardware, on the other hand, consists of materials such as silicon microchips, motherboard, hard drive, and others. In other words, Softwares are analyzed primarily in terms of what binary codes they have, whereas hardwares are analyzed in terms of what relevant parts constitute them. Both software and hardwares can be analyzed independently of each other, so both of them are logically distinct from each other. Likewise, two continuant persons are inseparable from each other, but they are logically distinct from each other.

Second, Lewis might face the problem of the ambiguity of identity. Lewis already mentions this problem by using an example of someone named "Ned". The name "Ned" might suffer from an ambiguity, since if the person "Ned" is in fact two indiscernible continuant persons who will split up at the moment fission, then the name "Ned" is ambiguous between two continuant persons. Lewis responds to this objection by pointing out that this is not a problem as long as "Ned" is essentially about someone before fission. In other words, we count two continuant persons as "Ned" until they split up into two separate continuant persons. Furthermore, Lewis said that facts about "Ned" such as "Ned is tall", "Ned is frightened", and "Ned is

waiting to be duplicated" are true about "Ned" but not about two persons who split up at the moment of fission (Lewis 1976, 29).

What if we replace "Ned" with an indexical "I"? Suppose that someone named Ned uttered "I am shy" at T1 prior to the moment of fission (T2) and at the moment of fission there are two continuants who are also shy. If both continuants are indiscernible at T1, then both uttered "I am shy" at the same time, place, and other contexts. However, the indexical "I" from "I am shy" is not a single indexical that rigidly designates a single person, but in fact there are two indexicals that rigidly designate two persons. After all, the indexical from "I am shy" can only designate someone who utters that statement and it happens to be the case that there are two continuant persons at T1 who uttered that statement. How can we count both continuants at T1 as one person when an indexical from the statement "I am shy" is in fact two indexical rigid designators? It would seem unusual to count both continuants as a single person given the above argument.

With regards to the first objection, Lewis might respond that we do not have to preserve everything about our common sense intuition of personal identity, preserving some of it should be enough. As long as C1 and C2 are temporally indiscernible by sharing the same person-stage, then we should count C1 and C2 as the same person at a certain time. When C1 and C2 cease to be temporally indiscernible at the moment of fission, then we should count both as two persons. This way of counting is intuitive because we tend to count spatiotemporal discrete objects the same way. If X and Y occupy the same location at the same time, then we count X and Y as one thing. Likewise, if C1 and C2 share the same person-stages at the same time, then we should count them as one person.

While it is the case that we count object X and Y as one object if X and Y occupy the same spatial and temporal location (same exact place at the same time), it may also be the case that X and Y are logically distinct from each other even though both occupy the same location at the same time. In principle, X and Y can superpose each other. For example, in physics there is a phenomenon called interference or wave propagation. Two waves can superpose each other when their corresponding crests meet each other at the same frequency, such superposition leads to indiscernibility between two waves (i.e. constructive interference). While we might count these waves as one wave for practical purposes, logically both waves are distinct from each other. Likewise, one could say that both continuant persons C1 and C2 superpose each other by sharing the same person-stages, but they are still logically distinct from each other. Technically, we should count C1 and

So

C2 as two continuant persons rather than one.

It also seems as though Lewis' method of counting could be forensic. For clarification, Locke believes that the term "person" is a forensic term insofar as it ascribes moral agency to someone who may not necessarily be a single unified agent. As long as some person is a moral agent who can be held culpable or praiseworthy for *his* (or *her*) action by simply remembering or being aware of that action, then we ought to treat it as though it is a single person. We may treat a split-brain patient as a single person even though he or she has two minds inside one body, since a split brain patient can remember or be aware of his or her action. How is this relevant to Lewis' tense identity? Lewis appears to think that as long as two continuants act as a single unified agent, then we should count them as a single person.

However, it may not be enough that we count both continuants as a single person just because they appear like a single person. We may really want to count someone as a single person because that person is authentically a single person. To count two continuants as a single person on the basis of indiscernibility comes as close to pretending that they are a single person. It would seem that we count someone as a single person, rather than two persons, simply because it is convenient to our need to ascribe moral agency to someone. This hardly corresponds with our common sense intuition about personal identity, so it seems that Lewis does not succeed in preserving our common sense intuitions about personal identity.

With regards to the second objection, Lewis might respond that given that two continuants are indiscernible, we should treat "I" from "I am shy" as a pseudo-indexical. In this case, "I" looks like an indexical that refers to a single speaker, but in fact it refers to two continuant persons who are engaged in the same speech activity in an indiscernible manner. If two continuant persons engage in the same utterance in an indiscernible manner, then we should count both as a single speaker until the moment of fission. However, at the moment of fission, when two continuants utter the same statement "I am shy" there is a genuine indexical that rigidly designates the speaker.

There are at least two concerns with this potential response. First, one concern is that all indexical statements with an "I" might be a pseudo-indexical, since nobody is a single continuant person under Lewis' view. Furthermore, If two continuants split up at the moment of fission, why assume that they will not go through another moment of fission again? If there is another fission moment, then both continuants would utter a pseudo-indexical. Logically, this could go into a potential infinite

## compos mentis

regress. Second, suppose that I have a clone who is qualitatively identical to me and both of us uttered the statement "I am shy". In this situation, me and my clone are indiscernible, but the statement "I am shy" are two statements with separate indexicals that designate two speakers. This situation is analogous with the case of two continuant persons insofar as both cases are cases of indiscernibility. If we should treat two continuants as a single speaker due to indiscernibility, why not treat me and my clone as a single speaker? The point of this analogy is to show that indiscernibility is not an excuse to treat two persons as one speaker.

In conclusion, Lewis and Parfit agree that formal identity does not matter, but survival matters. However, there seems to be a real disagreement between Lewis and Parfit. The disagreement is whether or not common sense intuition of personal identity matters. Parfit does not think it matters, but Lewis does. Lewis attempts to preserve common sense intuition of personal identity with his notion of tense identity. However, I argued that Lewis' tense identity does not succeed in preserving our commonsense intuition of personal identity. I argued that our common sense intuition of personal identity inclines toward numerical identity. While Lewis' tense identity seems to give us something very close to numerical identity, it is essentially qualitative identity.

## REFERENCES

- Lewis, David. 1976. "Survival and Identity." In *The Identities of Persons*, edited by Amelie Oksenberg Rorty, 17–40. Berkeley: University of California Press.
- Parfit, Derek. 1971. "Personal Identity." *Philosophical Review* 80: 3–27.
- Parfit, Derek. 1976. "Lewis, Perry, and what matters." In *The Identities of Persons*, edited by Amelie Oksenberg Rorty, 91–108. Berkeley: University of California Press.