Do I Still Exist?

Ben White

18.02.2020
The Problem of Personal Identity

What makes a person remain the same person over time?
The Problem of Personal Identity

Who cares?!?!
The Problem of Personal Identity

The same question can be asked about things that are not persons, e.g. rivers, trees, or ships.
The Problem of Personal Identity

In the case of *persons*, though, the question takes on a special interest for us, since

(a) we are persons, so we might naturally be interested to know under what conditions *we* continue to exist, and

(b) a person is typically held *responsible* only for things that *they themselves* did, so we might want to know under what conditions a person at one time is the *same* person as a person at another time.
A Few Preliminaries...

While we sometimes say that someone has become a “different person” after undergoing a radical change in their lifestyle, character, or beliefs and values, we should be careful to distinguish

(a) cases wherein one person actually ceases to exist and is replaced by another, different person (perhaps inhabiting the same body),

from

(a) cases wherein a person becomes a different sort of person than they were before while nevertheless remaining the same person.
Qualitative vs. Numerical Identity

Here it can be helpful to distinguish *qualitative* identity from *numerical* identity.

Qualitatively identical, but numerically distinct.
Qualitative vs. Numerical Identity

Just as material objects might undergo various changes in their color, shape, or size without ceasing to be the same object that they were before, so too a person might undergo various changes in their “accidental features” without thereby ceasing to be the same person.
Qualitative vs. Numerical Identity

In other words, a person at one time may be *qualitatively distinct* from a person at another time while nevertheless being *numerically identical* to that person.
We should also distinguish the question “What makes a person remain the same person over time?” from the question “Under what conditions is a person at one time identical to something that exists at another time?”

If persons are essentially persons, meaning that a person cannot cease to be a person without thereby ceasing to exist, then these two questions will have the same answer.
Are Persons Essentially Persons?

We may, however, want to allow for the possibility that a person can exist at times when it is not a person.
If, for instance, we think that to be a person requires having certain mental features (e.g. thoughts, experiences, beliefs, and desires), we may still want to say that I existed as a zygote, before I came to possess such features (and thus before I became a person), and perhaps that I could also continue to exist as a comatose patient who had lost all mental capacities (and was thus no longer a person).
In that case, being a **person** might be like being a professional basketball player.
Just as one can *become* and later *cease* to be a professional basketball player without thereby *ceasing to exist*, so too it may be that one can become and later cease to be a *person* while remaining the same individual throughout.
Are Persons Essentially Persons?

If persons aren’t essentially persons (just as professional basketball players aren’t essentially professional basketball players), then the question “What makes a person remain the same person over time?” may thus have a different answer than the question “Under what conditions is a person at one time identical to something that exists at another time?”

With that point in mind, let’s focus on the first question.
Personal Identity as Psychological Continuity

Most of us take our mental states (e.g. our memories, thoughts, beliefs, desires, etc.) to be central to our identity as persons.

This lends support to the idea that what makes a person at one time the same person as a person at another time is for the mental states of the one to be appropriately related to the mental states of the other.

In slogan form:
Personal identity consists in psychological continuity.
On this view, if two people Bob and Jane were to undergo a brain transfer operation, so that the inhabitant of Bob’s body woke up after the operation with Jane’s beliefs, memories, etc. and *vice versa*, Jane would then *inhabit* Bob’s body (or rather, what was once Bob’s body would now be *hers*), and *vice versa*. 
One of the most notable advocates of the psychological continuity theory is John Locke.

Locke distinguishes *human beings* from *persons*.

Each of these has a different “criterion of identity”; i.e. what is required to remain the same *person* over time is different from what is required to remain the same *human* over time.

John Locke (1632 - 1704)
Locke’s definition of a **human being**: “[T]he idea... [of a hu]man... is nothing else but of an animal of such a certain form” (II.27.viii.).
Locke’s definition of a person:

“[P]erson stands for... a thinking intelligent being that has reason and reflection, and can consider itself as itself, the same thinking thing in different times and places, which it does only by that consciousness which is inseparable from thinking” (II.27.ix.).

John Locke
(1632 - 1704)
“[S]ince consciousness… is that which makes everyone to be what he calls self,... in this alone consists personal identity... And as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person; it is the same self now it was then, and it is by the same self with this present one that now reflects on it that that action was done” (II.27.ix.).

The criterion of identity for persons is thus, roughly, continuity of consciousness, with special importance placed on memory.

John Locke
(1632 - 1704)
“[S]uppose I wholly lose the memory of some parts of my life, beyond a possibility of retrieving them, so that perhaps I shall never be conscious of them again; yet am I not the same person that did those actions, had those thoughts that I once was conscious of, though I have now forgot them? To which I answer, that we must here take notice what the word I is applied to; which, in this case, is the [hu]man only... But if it be possible for the same [hu]man to have distinct incommunicable consciousness at different times, it is past doubt the same [hu]man would at different times make different persons” (II.27.xx.).
Reid’s Objection

Suppose an *old general* is reflecting back on his life and can remember a time when, as a *young officer*, he won an important battle.

Suppose that when he was a *young officer*, he could still remember a time when he was whipped at school as a *boy*.

The *old general* has, however, lost all memory of this incident.
Reid’s Objection

According to Locke’s theory of personal identity, the old general = the young officer, and the young officer = the boy at school. This entails that the old general = the boy at school.
Reid’s Objection

Yet, on Locke’s view, the old general ≠ the boy at school (since the old general has no memory of that person’s actions or experiences)!

Doesn’t remember!
Reid’s Objection

“[T]herefore, according to Mr LOCKE’s doctrine,... the general is, and at the same time is not the same person as him who was flogged at school” (Essays, 276).
Psychological Continuity as Causal Continuity

Hume offers a way of avoiding Reid’s objection by making the identity of a person dependent on the causal relations that connect their current mental states with their future and past mental states.

“[T]he true idea of the human mind is to consider it as a system of different perceptions or different existences which are linked together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other” (Treatise, I.iv.6).
Psychological Continuity as Causal Continuity

“Whatever changes [a person] endures, [their] several parts are still connected by the relation of causation” (*Treatise*, I.iv.6).

David Hume (1711 - 1776)
A Problem for Psychological Continuity Theory?

What if the person I am now is *equally* psychologically continuous with *two different* people in the future?

Which person, if any, would I be?
Can a Person Split in Two?

People with severe neurological disorders sometimes undergo an operation wherein the left and right hemispheres of their brain are severed from one another.

Each hemisphere thereafter functions more or less independently of the other, yet we typically treat the person *after* the operation as (numerically) identical to the person *before* the operation.
What if, however, I were to undergo such an operation, and then have *one* half of my brain transferred into *one* body, and the *other* half into *another* body?
The people inhabiting these two bodies would then seem to be *equally* psychologically continuous with me. Yet the person I am *now* couldn’t be identical to *both* of them, for the people inhabiting these bodies wouldn’t be identical to one another...
Can a Person Split in Two?

Some Psychological Continuity Theorists claim that if I were to undergo such an operation, I would \textit{cease to exist}, and two \textit{new} persons would come into being (one inhabiting the body to which my left brain was transferred, and the other inhabiting the body to which my right brain was transferred).
Can a Person Split in Two?

However, what if only one half of my brain were transferred to another body, while the other half was destroyed?

*Then* it seems like I would *continue to exist*, as the person inhabiting the body to which the remaining half of my brain had been transferred...
Can a Person Split in Two?

This seems weird though...

Why should *destroying* one half of my brain enable me to *continue* existing as the same person, whereas *preserving both* halves of my brain *annihilates* me?
Identity vs. Survival

Some philosophers, e.g. Derek Parfit, have argued that even though you would not be identical to either of the two people brought into existence after the brain transfer operation, you would nevertheless survive as the two of them.

Both of the resulting persons would qualify as “future selves” of you, by virtue of the psychological continuity between your current self and them (again, even though you would not be identical to either of them).

Derek Parfit (1942 – 2017)
Identity vs. Survival

Parfit further argues that in reflecting on our continued existence, what matters to us is not the continued existence of a person identical to us, but rather a person (or persons) who are psychologically continuous with our current selves (i.e. who share our memories, beliefs, intentions, values, character traits, etc.).

Derek Parfit
(1942 – 2017)
Thus, although you would in a sense *cease* to exist if both halves of your brain were preserved and *continue* to exist if one half were destroyed (since in the former case there would no longer be any person *identical* to you), you nevertheless ought to prefer the *former* option, as this will ensure greater psychological continuity between your current self and certain persons in the future (although those persons would not be *you*).
Other philosophers, e.g. David Lewis, argue that the possibility of such cases show that persons can overlap, by sharing certain temporal parts or “person stages.”
Persons and Person-Stages

On this view, a person is a *temporally extended* object, with *temporal* as well as *spatial* parts.
Persons and Person-Stages

Different person-stages qualify as parts of the same person by virtue of bearing certain relations to one another. For Lewis, the key relation is (you guessed it) psychological continuity.
Persons and Person-Stages

Sounds plausible enough...

But how could different persons share the same temporal parts?
Persons and Person-Stages

Consider a clay statue and the lump of clay that it’s made up of.

Even though the statue and the lump of clay share the same location, one might argue that they can’t be identical, because they have different persistence conditions.

If, e.g., you were to smush the statue into a ball, you would destroy the statue, but the lump of clay would still exist.
Persons and Person-Stages

Just as a statue and the lump of clay that it’s made up of can share the same location without being identical, so too different people might share some of the same temporal parts without being identical.
Persons and Person-Stages

If Lewis is right, then a person doesn’t *cease* to exist when they undergo fission. Rather, they have been 2 different persons *all along* (which happened to overlap by sharing certain person-stages in common *prior* to undergoing fission).

David Lewis
(1941 – 2001)
Lewis consequently rejects Parfit’s claim that the question of whether one survives in the future is distinct from the question of whether the person one is now is identical to any person(s) in the future.

For Lewis, a person would survive fission precisely because the person-stages existing prior to fission and certain person-stages exiting after fission would be temporal parts of the same person.
Persons and Person-Stages

According to Lewis, the question of whether my body currently belongs to 1 person or 2 people depends on whether I in fact undergo fission in the future.

This seems a bit odd...

David Lewis
(1941 – 2001)
Persons and Person-Stages

This also generates a further puzzle...

If at some point in the future, I do in fact split into two different people (so that my current person stages are also stages of them), when I now use the word “I”, which person does it refer to?

David Lewis
(1941 – 2001)
"Here I am!"
Persons as Indeterminate

Perhaps in these unusual cases, there simply is no definite answer to the question whether some person in the future is the same person as we are now.
Persons as Indeterminate

Imagine a range of cases wherein varying amounts of your body are replaced by duplicate matter.
Persons as Indeterminate

Some might argue that even if your entire body was replaced, the resulting person would still be you (so long as you retain the same memories, beliefs, habits, etc.).
Persons as Indeterminate

But imagine that a perfect duplicate of your body is assembled **without** your current body being destroyed. Surely we wouldn’t say that the duplicate is **you**!

Why should we treat the case where your current body **is** destroyed any differently?

Why wouldn’t this just be a case wherein you are **annihilated** and **replaced** by a duplicate?
Persons as Indeterminate

If asked whether the person emerging from such a replacement operation would be you, Parfit suggests that while the answer will clearly be “Yes” in cases where relatively few of your cells are replaced, and clearly “No” in cases where all or nearly all of your cells are replaced, there will still be a range of intermediate cases where the answer is unclear.
Persons as Indeterminate

Parfit’s argues that, in these intermediate cases, there simply is no answer to the question of whether the person coming out of the operation would be you.

It is neither true nor false to say that the resulting person is you!

Derek Parfit
(1942 – 2017)
Persons as Indeterminate

Importantly, this is not because there are additional facts that we need to know in order to answer this question, but rather because our concept of “person” or “self” is itself vague and indeterminate.

Derek Parfit
(1942 – 2017)
The Self as Indeterminate
Persons as Indeterminate

The question of whether the person in such intermediate cases is the same or not is consequently merely a linguistic question about how we want to use the terms “person” and “self.”

We may decide to answer the question one way or another, and thereby render the boundaries of our concept of “self” or “person” more determinate by stipulation, but this is merely a matter of convention.
Persons as Indeterminate

Accepting this idea could enable us to treat personal identity as a matter of degree.

Perhaps there is a sense in which the person you are now is more of the same person that you were 5 minutes ago than the person you were at some point in the distant past that you now have no memory of (even if there is still a causal connection between your current mental states and the mental states of that past self).
Do Persons Exist?

The Buddhist doctrine of *anatta* (“non-self”) can be interpreted as holding that persons do not exist.

“Buddha has spoken thus: ‘O brethren, actions do exist, and also their consequences, but the person that acts does not... There exists no Individual, it is only a conventional name given to a set of elements.” (Cila Mara)
Do Persons Exist?

In his provocatively entitled paper “I Do Not Exist,” Peter Unger likewise argues that there are in fact no such things as persons.

Peter Unger
(1942 – ?)
Do Persons Exist?

Unger arrives at this shocking claim by way of a modified version of the famous sorites paradox.

1. A pile of 10,000 grains of sand constitutes a heap.
2. Removing one grain of sand from a heap will not turn it into a non-heap.
3. Therefore, a single grain of sand constitutes a heap. (!?!?!)

sorites paradox
Do Persons Exist?

Unger arrives at this shocking claim by way of a modified version of the famous *sorites paradox*.

1. If I exist, then I consist of a finite number of cells.
2. If I exist, then the removal of one cell from me will not make a difference as to whether I exist.
3. Therefore, I could exist without any cells at all. (!?!?!!)
The conclusion Unger thinks we should draw from this reasoning is that there never were any persons (or heaps) to begin with.

The only things that exist, on his view, are absolutely simple material things (if there are such), and things (e.g. molecules, perhaps) that cannot lose any of their parts without thereby ceasing to exist.
So...

Do I still exist?

Sure... but in some cases there may be no determinate answer.

Anyway, that’s not what matters. What matters is whether someone psychologically continuous with you still exists (whether or not that person is you).

Sure you do! But there is currently no way of telling how many persons you are...

You never did!