Your Name:

Personal Identity: cases, some views, and your view

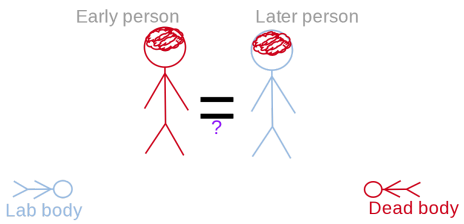
Below, you will find four sections. The sections contain:

1) a list of possible **situations** about which the question 'Does that person survive?' is especially interesting. There, I ask you to respond *whether* you think the person would survive the given event and, importantly, *why* you think so.
2) some possible **views** on survival and personal identity through time.
3) a **table** where you can record:
    a) what your thoughts were about the potentially difficult situations in Section (1);
    b) what you think the views on survival in Section (2) entail about those situations in Section (1).
4) a series of **further questions** you can ask yourself in order to further develop your own view on survival and personal identity through time.

1) Potentially interesting situations and questions

- When thinking about the situations and questions below, try to do so as independently as you can. That is, try (the best you can) not to be *too* attached to any of the particular, pre-formulated views or 'theories' (the body view, psychological view, etc.) that are found in Section (2) of this handout. If you can do this to some extent, then your thoughts are what we may call your "pre-theoretical" intuitions. But, at the same time, you should be able to give some reasons for your answers. You should reflect on those intuitions. These are supposed to help you decide which view seems most plausible to you.

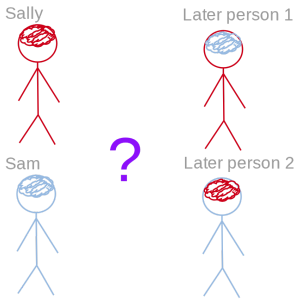
Case 1: **Brain transplant (or 'body transplant'?)**: Suppose Bob got in an accident. His brain remains intact, but the rest of his body is way too damaged. The year is 2224. The technology to create brainless human bodies in labs is available. Doctors create a brainless body in the lab and move Bob's brain to it. At the end of the procedure, there is a complete human body. The brain is Bob's old brain, and it is well-connected to the rest of the lab body. Thus, the body functions just like any other person's body. There is a person there! The body/person wakes up. Doctors ask 'how do you feel?' The person says: 'I feel ok. I remember the accident. It was horrible. But now I feel ok. Have Gordon, Alex, and my other friends come to visit me?'



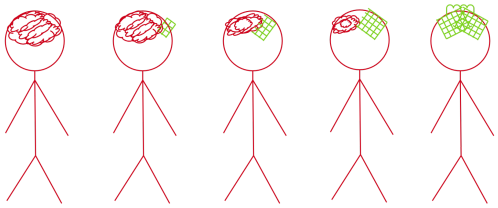
Question: Did Bob survive the accident? Does he still *exist*? More specifically, is the person with Bob's old brain and with the new 'lab' body, *him*? (Is that person numerically identical to Bob?)

Circle one of these: Yes – No – Unclear
Whatever your answer is, explain why:

Variation of Case 1: Brain swap: Suppose doctors take two persons, Sally and Sam, and swap their brains across their bodies. Now Sally's original brain is in Sam's original body, and Sam's original brain is in Sally's original body. Where are them? Is Sally in the body that has Sally's original brain and Sam's original body, or is she in the body that has Sam's original brain and Sally's original body?



Case 2: **Gradual brain replacement**: The year is 4024. Doctors can replace your heart with a prosthetic heart that functions as well as the original, biological heart. They typically do so under any sign that your heart may fail in the future. But that's not all. More impressively, doctors can also replace brain cells in your brain with prosthetic cells that work just as well as the original, biological brain cells. Suppose doctors find that Gordon's entire brain will "fail" in 20 years unless he starts a gradual replacement of his current brain. Doctors can replace a few thousand brain cells per second. Because the artificial cells work as well as the biological ones, no memories or personality traits whatsoever are lost in virtue of the replacement of brain cells. Gordon decides to go through the 'treatment.' Doctors do a few hours of cell replacement every day at night when Gordon is asleep. In fifteen years, by the end of the 'treatment,' the brain Gordon started with is completely gone. But there was no loss of memory or personality in virtue of cell replacement.



Question: Did Gordon survive the gradual replacement of his brain cells? Does he still exist? More specifically, is the person with Gordon's memory and personality (but without his brain) Gordon?

Yes – No – Unclear?
Why?

Variation of Case 2: What if the new cells are human brain cells, just like the old ones, except that they are new—say, grown in labs?

Case 3: **Amnesia**: Alex suffered an accident yesterday. She went unconscious (in the same sense as we go unconscious every night when we go to sleep). Her body was brought to the hospital. Today, at the hospital, 'the person in her body' (we don't want to prejudge whether that is her or not) doesn't have any of the memories that Alex used to have. The person now in the hospital doesn't remember Alex's friends or family, she doesn't remember her day yesterday, she doesn't remember being a teenager. She has no memory of her past. *Yet*, she has the same preferences as Alex used to have. When we play to her some songs that Alex used to love, she loves them too. When we show her songs Alex used to dislike, she dislikes them too. Likewise, her preferences for food, movies, places, people, and everything else, are exactly the same as Alex's. Her desires and plans for her future are also the same as Alex's previous desires and plans. She wants to become a lawyer, just like Alex wanted. Her moral character is the same as Alex's. She thinks that war is never justifiable but that lying is sometimes justifiable (namely, in exactly the same circumstances Alex used to think lying is justifiable). None of the memories are the same as Alex's, but the remaining aspects of Alex's psychology and personality are all still in that person in the hospital.

Question: Did Alex survive the accident? Is the person at the hospital *her*?

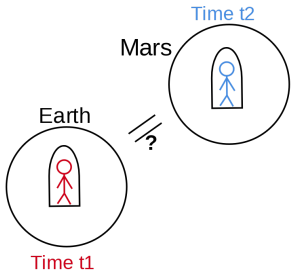
Yes – No – Unclear?
Why?

Case 4: **Complete loss of personality**: Suppose Peter suffers an accident. He goes unconscious (in the same sense as we go unconscious every night when we go to sleep). His body is brought to the hospital. Today, at the hospital, 'the person in his body' (we don't want to prejudge whether that is him or not) doesn't have any of the memories that Peter used to have. The person now in the hospital doesn't remember Peter's friends or family, he doesn't remember his day yesterday, and he doesn't remember being a teenager. He has no memory of his past. *Moreover* (different from Alex's amnesia case), the person in Peter's body has completely different preferences, desires, plans, and overall personality.

Question: Did Peter survive the accident? Is the person at the hospital him, or did he stop existing after the accident? Are Peter and the person in the hospital numerically the same person?

Yes – No – Unclear?
Why?

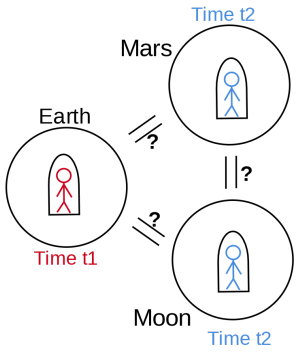
Case 5: **Teleportation**: The year is 302024. The teleportation technology that we used to see only in science fiction (like *Star Trek*) is now available. Ted can enter a box, which is located on Earth, that scans his entire body, atom by atom. As the scanner does so, it destroys Ted's body and immediately sends the information (maybe through a sort of light signal) to a second box, which is located on Mars. This second box takes that information and arranges atoms inside of it in exactly the same way as atoms in Ted's body were arranged on Earth. By the end of the process, there is an atom-by-atom perfect replica of the body that was previously on Earth. Since the replica is, atom by atom, just like the previous body, the replica's brain is just like the one previously on Earth. The body/person on Mars is conscious, and his memories, desires, plans, and overall personality are just like Ted's.



Question: Did Ted survive teleportation? Is the person on Mars really him?

Yes – No – Unclear?
Why?

Case 6: **Malfunctioning teleportation**: The year is 302024. The teleportation technology described above is available. Ted enters the teleporter, but it malfunctions. It creates two replicas of Ted's body, one on Mars and the other on the Moon. Both—the body/person on Mars and the one on the Moon—are conscious. They both have Ted's previous memories, desires, plans, and overall personality.



Questions: Did Ted survive teleportation? If so, is he the person on Mars or the one on the Moon? Also, is the person on Mars numerically identical to the person on the Moon?

Yes – No – Unclear?
Why?

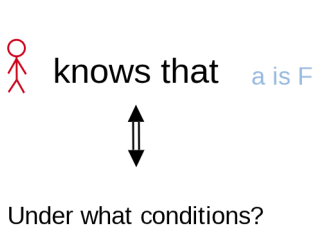
2) Some views on personal identity

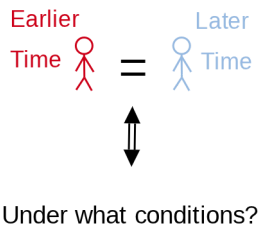
Views on personal identity over time attempt to answer the question: 'Under what conditions does a person survive from an earlier time to a later time?'. In other words, the question is: 'Given a person E, at an earlier time, and a person L, at a later time, under what conditions are E and L numerically identical?' In yet, what words, these views attempt to give us a general criterion of personal identity, which can be formulated by completion of the following schema:

A person at an earlier time, 'E', is numerically identical to a person at a later time, 'L',
when, and only when, …

…… ?

The task of finding a general criterion of personal identity isn't that different from the task of finding an analysis of knowledge. In the latter case, we tried to find the necessary and sufficient conditions for a person to know a proposition. In the case of personal identity, we want to find the necessary and sufficient conditions under which E and L are numerically identical. The figures below represent each of these tasks:





Body View: The body view roughly says that you survive when (and only when) your body survives.

More precisely: Suppose that we have a person at an earlier time—let's call her 'E'—and a person at a later time—call her 'L'. The view says:

E is numerically identical to L
when, and only when,
E's body is numerically identical to L's body.

Example: I survive a haircut because my body survives it. That person who entered the barbershop is numerically identical to that person who left the barbershop because 'they' have the same body. Even though it changed a little bit (because of the haircut), the body that entered the barbershop is still the same as (meaning, 'numerically identical to') the body that exited the barbershop.
Question: what is it for an early body to be identical to a later body?

Brain View: The brain view roughly says that you survive when (and only when) your brain survives.
More precisely:

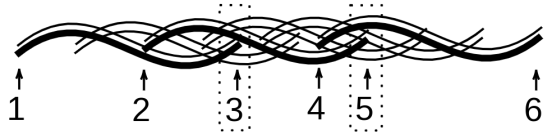
E is numerically identical to L
when, and only when,
E's brain is numerically identical to L's brain.

Example: I survive a haircut because my brain survives it. That person who entered the barbershop is numerically identical to that person who left the barbershop because 'they' got the same brain; the brain that 'entered' the barbershop is numerically identical to that brain that 'exited' the barbershop.
Question: what is it for an early brain to be identical to a later brain?

Memory Continuity View: Roughly, the memory continuity view says that you survive when (and only when) your "flow of overlapping memories continues." We constantly create new memories and forget old ones. As long as this is gradual and there is enough overlap between new and old memories, the "flow of overlapping memories continues" and so you survive.

Here are two pictures and a more precise definition that help elucidate this view.

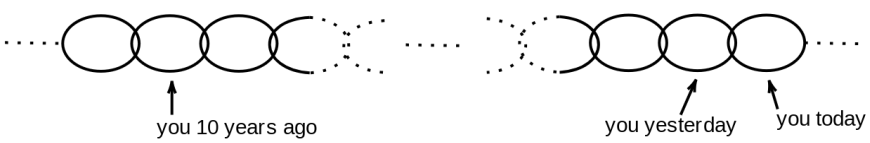
Picture 1: Imagine a long rope extending from left to right (say, from point 1 to point 6, as in this figure).



No fiber of the rope spans the entire distance from the beginning to the end (in the figure above, none connects 1 to 6 directly). But the fibers at the beginning that happen to end halfway (say, the one I highlighted in bold that goes from 1 to 3) overlap with other fibers (like the ones that go from 2 to 5). These fibers don't reach either the beginning or the end of the rope, but they may overlap with still further fibers that extend all the way to the endpoint (like the fiber that goes from point 4 to 6), or maybe they could overlap with fibers that overlap with fibers that reach the endpoint, or maybe they overlap with fibers that overlap … that overlap with fibers that reach the endpoint.

The rope can be seen as representing a person spread across time. Each vertical slice of the rope corresponds to a specific moment in the person's life. (For example, in the figure, slice 3 could represent you on March 1st, and slice 5 could be you on April 1st.) The fibers represent that person's memories. Memories that are created early in a person's life may be forgotten, but they will overlap with other memories that may be preserved all the way until the end of that person's life, or they may overlap with still further memories that are preserved all the way until the end of that person's life, or … … .
(Thickness of fibers doesn't represent anything. This was just to highlight particular fibers in the rope.)

Picture 2: Imagine a chain. Each ring of the chain is directly connected to two other rings—one on its left, the other on its right. Even rings that are not directly connected are still indirectly connected to each other by one or more intermediate rings that are directly connected. Each ring of the chain represents a person at a moment (not a memory, and thus, unlike the rope picture). Let's say that each of the rings represents you on a day. You today have a direct memory connection to you yesterday because you today remember enough of the experiences of you yesterday. This direct memory connection is represented by the direct link between rings that represent you yesterday and you today.



You today don't remember enough of your experiences 10 years ago. Thus, you today are not directly connected to you 10 years ago. Still, you are (numerically) the same person as you were 10 years ago. This is because you today are directly connected to you yesterday, and you yesterday were connected to you two days ago, and so on … up to you 10 years ago.

In this case, we can say that you today are connected to you 10 years ago through
an overlapping chain of direct memory connections.

More precisely:

E is numerically identical to L
when, and only when,
there is an overlapping chain of direct memory connections that connects L back to E.

Example: Suppose there is an old general who right now doesn't remember any of the experiences he had when he was 8 years old. But suppose the old general still remembers most of the experiences of when he was a young soldier at war, and that when he was a young soldier at war, he remembered most of the experiences of when he was 8 years old. The memory view says that the old general is numerically identical to the 8-year-old kid even though he doesn't remember the experiences of the kid because they are connected by the same overlapping chain of direct memory connections. This is in part because the soldier is remembered by the old general and remembers the kid, and so the soldier is directly connected to the kid and to the old general.

Psychological Continuity View: The psychological continuity view is very much like the memory view (see above). Roughly, the psychological continuity view says that you survive when (and only when) your "flow of overlapping psychological and personality traits continues." That is, instead of restricting itself to memories, this view takes on board memories, desires, preferences, intentions, plans, beliefs, and other psychological and 'personality' traits that we have. All the time, our psychology/personality (i.e. memories, desires, preferences…) changes. We gain new ones and lose old ones. As long as this change in our psychology is gradual, as long as there is enough overlap between new and old psychological traits, the "flow of overlapping psychological and personality traits continues," and so you survive.

The pictures above, of the rope and of the chain, help clarify the psychological view, too. Except that now, the fibers of the rope now represent not only memories but other relevant contents of our psychology as well (that is, beliefs, desires, intentions, etc.). In turn, the links between rings in the chain represent direct *psychological* connections. These connections depend not only on memories but also on the other relevant contents of our psychology (beliefs, etc.).

More precisely:

E is numerically identical to L
when, and only when,
there is an overlapping chain of direct *psychological* connections that connects L back to E.

Example: Consider the old general in the 'memory continuity view' example. He is directly connected to the young soldier not only because of his memories but also because of the other relevant psychological contents of his mind.

Contrast: Consider Alex's case of amnesia, where the memory is completely lost, but the 'personality' is kept intact. What do you think the 'memory continuity view' and the 'psychological continuity view' would say about this case?

Soul View: The soul view roughly says that you survive when (and only when) your soul survives. A soul is an immaterial thing that we all have, that somehow animates our bodies and that somehow contains our experiences, memories, and personality. But the soul is its own thing; it is neither the body, brain, personality, or psychology. So it is, in principle, possible to detach the soul from all that.

More precisely:

A person at an earlier time, 'E', is numerically identical to a person at a later time, 'L',
when, and only when,
E's soul is numerically identical to L's soul.

Example: I survive a haircut because my soul survives it. That person who entered the barbershop is numerically identical to that person who left the barbershop because 'they' have the same soul.

Can you think of a different view?

A person at an earlier time, 'E', is numerically identical to a person at a later time, 'L',
when, and only when,

_____.

3) Tracking some of your thoughts and other people's thoughts

You can do a few things in the table below. First, you can record your 'intuitions'—that is, your responses to the questions asked in Section (1)—under column 1 below. You can write down 'yes' (the person survives), 'no', or 'unclear'. Second, you are invited to think about how the different views described in Section (2) would respond to the questions asked in Section (1). For example, what do you think the 'Psychological View' (column 5 below) would say about Alex's Amnesia (row 3 below)? What about the Soul View, the Memory View, etc? You can record your answers to these questions below. (Suggestion: annotate why you think so on the next page.) Third, you can come up with your own cases, describe them on the next page, add them to the table (rows 7 or 8), and record what you think each view should say about your new cases. Finally, you can come up with a 'New View' or a new version of a view, describe it, add it to the table (column 7), and complete its column.

	1 Your intuitions	2 Body View	3 Brain View	4 Memory Continuity	5 Psychol. View	6 Soul View	7
1 Brain transplant							
2 Brain replacemnt.							
3 Amnesia							
4 Loss of person.							
5 Teleport							
6 Malfunct. Teleport							
7							
8							

4) Developing and defending your views

Here are a few things that may help you figure out and defend your own views on the topic:

- Assessing views and raising objections: The views in Section (2) are stated as follows 'E (any person at an earlier time) is numerically identical to L (any person at a later time) when, and only when, … [Criterion X is true] … ' Different views offer different criteria for personal identity, 'X'. For example, the brain view says that E is (numerically) identical to L when, and only when, 'they have the same brain'. About each view, you should ask:
    - Are there possible situations where E is *identical* to L even though [Criterion X] is false? (Example: Is there a logically coherent situation where E and L are *identical*, but they *fail* to have the same brain?)
    - Are there possible situations where E and L satisfy [Criterion X], but they fail to be identical? (Is there a logically coherent situation where E and L have the *same brain*, but they *fail* to be identical?)

  (Exercise: Try to think of how situations in Section (1) provide positive responses to these questions regarding the views in Section (2).)

  Moreover, you can ask:
    - Is it plausible that *one* is in a position to *know* that E is (numerically identical to) L whenever she is in a position to *know* that Criterion X is true? How? (For example, given the soul view, ask: 'When I am in a position to know that a person at a later time is (numerically) the same as that person at an earlier time, am I also in a position to know that E's soul is the same as L's?'.) If not, is the view in trouble? Why?
- Still raising objections: If you think you can defend well an intuition you have about a case—say, you can defend well the thought that Alex survives the amnesia in case 3—and if you think a certain view is committed to the opposite result—say, that Alex doesn't survive the amnesia case—then you can turn your justification for that intuition into an objection against that view. And thus, such a view is probably not one you want to endorse.
- Thinking about further cases to raise further objections: Likewise, you can try to come up with new cases, new possible situations about which you and a certain view disagree.
- Rethinking intuitions and entailments: If what a view entails about a case disagrees with your reflected intuitions about that case, try to reconsider both your intuitions—'Do I really think that?'—and what you think that view entails—'Does the view really entail that?'
- Rethinking views: The list of views is not at all exhaustive. For example, there are many possible variations of the views above. Try to think about how relatively small modifications to the views above could change what they entail. Could those views be modified in a way that you would find them overall better? (Suggestion: try to think about different versions/modifications of the psychological view that give different responses to the teleportation and malfunctioning teleportation cases.)
- Formulating new views: By reflecting more and more on the reasons you tend to give for saying that someone survives or does not survive an event, you may be able to work out a different theory that is not a mere 'modification' of the ones in Section (2).