

Lecture 3 **Psychological Continuity and Survival**

1. Introduction

Last week we considered Locke's version of the psychological continuity view, and then turned our attention to Williams' observations on transplant cases. Recall that transplant cases involved testing our intuitions concerning our *survival*. Do we survive? And if we do, how do we do so? (e.g. as the same body with a different consciousness, or as the same consciousness with a different body?) Williams argued that our intuitions seem to be malleable; they seem to be sensitive not just to the facts of the case, but to the manner in which the case is described. Williams also argued that, if such cases can be described in stages, it is not clear at which stage we should declare the person in question to have ceased to exist, or to have moved into another body.

Today, we'll look at Parfit on the question of survival and personal identity. Parfit famously argued that personal identity is not "what matters". We'll examine how he arrives at this conclusion by carefully reconstructing his arguments.

2. Arguments from Survival

To understand Parfit's arguments, we need to understand the general structure of survival-based arguments. Many thought experiments in the debate about the persistence of persons attempt to test the limits of the **relation**

'X is **the same person as** Y'

They purport to do this by testing our intuitions about X's **survival**. But in order for intuitions about survival to tell us anything about the *same person as* relation, we need the following assumption:

Necessity: If X is the same person as Y then X survives as Y.

In order for us to conclude from the **failure** of survival that X does not persist, it must be the case that survival is a necessary condition on the persistence of persons. If this were not the case, then the failure of survival would be consistent with a person's persisting.

But notice that some survival-based arguments also assume sufficiency holds, i.e.:

Sufficiency: If X survives as Y then X is the same person as Y.

With this in mind, we can now understand Parfit's arguments as, in part, challenging the **sufficiency of survival** for the persistence of persons.

3. Assumptions about the Question of Identity

At the beginning of his paper, Parfit explains that he will challenge two assumptions made in the personal identity literature.

1. In all cases, the question about identity must have an answer.
2. The answer to questions about identity are necessary for answering questions about **survival**, moral responsibility, memory and other related concepts.

You might think **1** because identity does not seem like a vague concept. It seems plausible to say that for any a and b , ' $a = b$ ' is either true or false.

On **2**, this also seems plausible when you consider our intuitions about these other concepts. We've already thought a little bit about survival. But now consider first-person memory: it doesn't seem like you could have first-person memory of anyone's experiences but your own. Similarly, it seems wrong to hold a person who is non-identical with you responsible for your actions.

Nevertheless, Parfit argues against each of these claims.

Against **1**, Parfit motivates the idea that we should reject this assumption by considering identity questions about clubs and nation-states. Consider the following:

*Suppose that a certain club exists for several years, holding regular meetings. The meetings then cease. Some years later, some of the members of this club form a club with the same name, and the same rules. We ask: 'Have these people reconvened the very same club? Or have they merely started up another club, which is exactly similar?' There might be an answer to this question. The original club might have had a rule explaining how, after such a period of non-existence, it could be reconvened. Or it might have had a rule preventing this. But suppose that there is no such rule, and no legal facts, supporting either answer to our question. And suppose that the people involved, if they asked our question, would not give it an answer. There would then be no answer to our question. The claim 'This is the same club' would be **neither true nor false**. (Parfit 1984: 213, original emphasis)*

The idea is that, once you've stated all of the facts of the matter, there no further fact about whether or not the club is the same. Thus, insofar as there is at least one case for which the identity question does not have a definite answer (if you accept the club case), then this undermines **1** (which was a universally quantified claim).

Parfit then turns his attention to the second claim – that *personal identity* (and the persistence of persons) is what matters with respect to prudential (i.e. pertaining to practical self-interest) and moral considerations.

4. **Fission Cases**

Here we return to the familiar territory of transplant-like thought experiments. This time, instead of supposing that your brain (or consciousness) is transferred in its entirety, we are to imagine that your brain is severed, and each *hemisphere* is transplanted into a distinct body. Call the recipients Lefty and Righty (according to which hemisphere they've received). In this case, one of three things might happen:

- (i) You do not survive fission.
- (ii) You survive as **one of** Lefty or Righty.
- (iii) You survive as both.

He takes each of these in turn.

On (i), he argues that this is implausible because it is possible to survive the transplant of a *single* hemisphere. People have, in fact, survived with only a single hemisphere of their brain. As such, even though hemisphere transplants have not occurred we have reason to think this is possible. If we can suppose that we would survive a single hemisphere transplant, then we cannot deny that we would survive a *double* hemisphere transplant without denying that identity is an **intrinsic relation**

Intrinsicness: Whether *X* is (numerically) identical to *Y* depends solely on features internal to *X* and to *Y*. (e.g. Whether *X* is the same sandwich as *Y* only depends on whether *X* and *Y* are made of the same slices of bread, filled with the same filling, etc.)

Now, if it is the case that you survive in the case of a single transplant, but do not in the case of a double transplant, then your survival depends on an *external factor*—namely, on the *number* of individuals with your psychological make-up.

On (ii), Parfit argues that this is also implausible. There is no non-ad hoc way of deciding which of Lefty or Righty you survive as. So (ii) suffers from arbitrariness.

Finally, against (iii), Parfit argues that this leads to unacceptable results based on other beliefs we hold about the relation of identity. First, we were interested in **numerical identity**. If (iii) claims that you are numerically identical to Lefty and Righty, then this will lead to contradiction.

- P1.** Lefty is not numerically identical to Righty. (By hypothesis)
- P2.** Post-surgery, you are numerically identical to Lefty and to Righty. (By (iii))
- C1.** Lefty is identical to Righty. (By the transitivity and symmetry of identity)

Now, perhaps the advocate of (iii) could deny **P1** and say that Lefty and Righty *are* identical after all. But this is implausible, Parfit argues, because Lefty and Righty seem to have all the attributes required to be *distinct* persons. Indeed, after the surgery, their experiences and memories would become increasingly divergent. This would have the result that Lefty is

identical to Righty, and yet does not have first-person memory of any of Righty's experiences and vice versa. Worse still, this option will still lead to contradiction. Suppose that, at some point post-surgery, Lefty feels thirsty but Righty does not:

- P3.** Lefty and Righty are identical. (By hypothesis)
- P4.** For all x and y , x is identical to y iff (for all p , p is a property of x iff p is a property of y). (Leibniz's Law)
- P5.** Lefty is thirsty. (By hypothesis)
- P6.** Righty is not thirsty. (By hypothesis)
- C2.** Lefty and Righty are not identical. (By P4, P5, P6)

Perhaps instead, the advocate of (iii) could accept **P1**, but deny that personal identity is transitive. One issue with this proposal is that it seems to change the subject. The relation we were interested in in the first place was the numerical identity of persons. It is fine to say that there is *some* intransitive relation that holds between Lefty, Righty, and you. But to call it 'personal identity' is effectively to say that the relation we originally sought isn't present. And that is not really to accept (iii) so much as to accept something like (iii)*:

(iii)* You bear an important, intransitive relation to Lefty and Righty.

But what should we say about the fission case then? It seems there's no good option!

Given what we've just said, we can understand Parfit's solution to the problem as simply accepting (iii)*. He does eventually give an answer to the identity question. But he argues that it is (iii)* and not the answer on identity that is more important. It is this distinct relation that matters.

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. (1971: 10)

5. Psychological Continuity and What Matters

Parfit begins claiming the fission case shows that "what matters for survival need not be one-one" (10) the way that the identity relation is. It is just that, in most cases that we ordinarily encounter (i.e. not sci-fi thought experiments) the relation that matters for survival *coincides with* identity. And this misleads us into thinking that identity is necessary for survival (or, if you like, that survival suffices for identity).

Parfit proposes to give an account of what matters in terms of psychological continuity. First, though, he defines the concept of **q-memory** and other related notions. The idea is to define a concept that is very like our concept of memory, but one that does not beg the identity question.

Q-remembering: Say that “I am q-remembering an experience if

- (1) I have a belief about a past experience which seems in itself like a memory belief,
- (2) someone did have such an experience, and
- (3) my belief is dependent upon this experience in the same way (whatever that is) in which a memory of an experience is dependent upon it.” (15)

From here, if I q-remember something, the thing that I q-remember is a **q-memory**. Now, I can have a q-memory of something that someone distinct from me did. But I can also have q-memories of things that I did; all memories are q-memories.

In a similar way, we can define **q-intentions**, **q-beliefs**, **q-responsibility**. If you like, we can define all of the elements we thought were relevant to psychological continuity, but in ways that do not assume identity.

Given this, he defines **psychological connectedness** as follows:

Psychological Connectedness: *X* is psychologically connected to *Y* to the extent that *X* and *Y* share q-states. So, for instance, to the extent that *Y* q-remembers *X*'s actions, *X* q-intends that *Y* perform certain actions, *Y* is q-responsible for *X*'s actions, etc.

Now, psychological connectedness is **intransitive**. It does not follow from *X*'s q-remembering *Y*'s actions, and *Y*'s q-remembering *Z*'s actions that *X* q-remembers *Z*'s actions. But we can define a transitive relation on this one by taking its **ancestral**.

Psychological Continuity: *X* is psychologically continuous with *Y* just in case *X* and *Y* are related by a chain of psychological connectedness.

Parfit proposes the following definition of personal identity:

Personal Identity: *X* is the same person as *Y* iff *X* is psychologically continuous with *Y*, and for all *Z* such that *Z* is psychologically continuous with *X*, *Z* = *Y*.

In other words, psychological continuity must be **non-branching** in order to suffice for personal identity. Of course, it follows from this that Intrinsicness is false. But Parfit is happy to accept that consequence. He argues that this is not a problem because the relation that matters for survival, for prudential concern, and other questions *is* intrinsic.

What matters for survival is **psychological connectedness**. Why connectedness rather than continuity? This is because, Parfit argues, what matters for survival must admit of degrees. Consider the fission case again. It seems plausible to say that you survive the surgery that results in Lefty and Righty. But what should we think about your survival at a point 50 years after the surgery? It now seems less plausible that you continue to survive. Lefty and Righty would be markedly different from how you are now. Most of their (q-)memories would be

post-surgery memories. They may have entirely different personalities than you do now. In taking psychological connectedness to be the relation that matters for survival, Parfit can accommodate this intuition.

So, what should we think about survival's relationship to identity? On Parfit's view, identity is good *evidence* for survival. In practice, psychological continuity doesn't come apart from identity, since in practice, people do not 'branch'. But notice that identity isn't sufficient for survival either. Consider: *X* can be psychologically continuous with someone that is not psychologically connected to *X*. (e.g. 15-y.o. you, you now, and some future person who q-remembers your actions, but has forgotten everything you did at age 15)