

GENERAL PHILOSOPHY

WEEK 6: PERSONAL IDENTITY

JONNY MCINTOSH

INTRODUCTION

This week, we look at questions surrounding **PERSONAL IDENTITY**, questions concerning our status as **PERSONS**.

- What sorts of things are we? What can we be?
- When do we begin (and cease) to exist?
- What sorts of changes can we undergo?

In particular, we'll look at questions surrounding the **PERSISTENCE OVER TIME** of a person:

- what it is for a person at one time to be identical to a person at another time, and why it matters.

The plan:

- The Questions
- Psychological Continuity
- Animalism

THE QUESTIONS

Our main question is: what is it for a person at one time to be identical to a person at another time? But what exactly do we mean by **IDENTICAL**?

X and Y are **QUALITATIVELY** identical IFF X and Y have the same intrinsic properties (qualities)

Examples:

- Each of a pair of socks
- Identical twins

X and Y are **NUMERICALLY** identical IFF X and Y are the very same individual

Examples:

- Clark Kent and Superman
- Jay Z and Shawn Carter

Our question: what is it for a person at one time to be **NUMERICALLY** identical to a person at another time?

Questions of numerical identity are often non-trivial:

- "Is Clark Kent, the journalist for *The Daily Planet*, Superman, the superhero?"
- "Is Robert Galbraith, author of *The Cuckoo's Calling*, J. K. Rowling, the author of the *Harry Potter* books?"

We're particularly interested in questions of numerical identity **OVER TIME**:

- "Am I the same person as the person who gave the mind & body lecture last week?"
- "Is this person in the photo me or my brother?"

Our question:

- What is it for person X at t to be numerically identical to person Y at t'?

Or, more generally:

- What is it for person X at t to be numerically identical to object Y at t'?

(The second formulation allows for the possibility that a person is not always a person, and so allows for the possibility that they will one day be a corpse or were once a foetus **WITHOUT** assuming corpses and foetuses are persons.)

The question is not just intrinsically interesting; it also seems to be particularly significant for us.

It seems that each of us has a special concern (an apprehension or anticipation) for what happens to **OURSELF**.

The question is also central to many other questions of personal identity:

- When do we start to exist? Can a person be identical to an earlier foetus?
- When do we cease to exist? Can a person be identical to a body in a vegetative state, or a corpse?

- What sorts of things can we be? What sorts of changes can we survive?

Can a person survive destruction of their body, for example?

Does a person survive destruction of their body if their thoughts are somehow transferred to a robot?

After all, we seem to survive at least some changes:

- I have slightly longer hair than I did last week
- I remember some things I didn't previously remember
- I've also forgotten some of what I used to remember

PSYCHOLOGICAL CONTINUITY

LOCKE'S VIEW

Suppose the soul of a prince is transferred to the body of a cobbler, and vice-versa, on Monday night. On Tuesday morning, which person has which body? Intuitively, the prince has the body of the cobbler and the cobbler has the body of the prince.

These intuitions seem to show:

- What is required for X at t to be identical to Y at t' is **NOT** bodily continuity.
- Rather, what is required is some form of **PSYCHOLOGICAL** continuity.

But what form of psychological continuity, exactly?

The example of the prince and the cobbler is Locke's. He is often read as offering a **MEMORY CRITERION** for personal identity.

- X at t is numerically identical to Y at t', where t is later than t', IFF X remembers Y's experiences.

- The person X in the cobbler's body on Tuesday is numerically identical to the person Y in the prince's body on Monday since X remembers doing things Y did.
- I am numerically identical to the person who gave last week's lecture because I remember giving it.

REID'S OBJECTION

- The old general remembers the exploits of the brave soldier
- The brave soldier remembers the experiences of the young boy
- But the old general does not remember the experiences of the young boy

According to Locke's Memory Criterion:

- The old general = the brave soldier
- The brave soldier = the young boy
- But the old general ≠ the young boy!

Numerical identity is **TRANSITIVE**: if X is numerically identical to Y and Y is numerically identical to Z then X is also numerically identical to Z.

But the memory criterion is **NOT** transitive: if X remembers Y's experiences and Y in turn remembers Z's experiences, it does not follow that X remembers Z's experiences.

Reid's objection shows that Locke's Memory Criterion won't do as it stands. But it is easy to come up with something better in the same spirit.

- X at t is numerically identical to Y at t', where t is later than t', IFF there is a chain of memory relations linking X to Y.

- The old general is numerically identical to the young boy because the old general remembers the exploits of the brave soldier, who in turn remembers the experiences of the young boy.

(Something for the maths and philosophy students: This revised criterion takes personal identity to consist in what is known as the **ANCESTRAL** or **TRANSITIVE CLOSURE** of the memory relation. It plays an important role in Frege's definition of finite cardinals.)

BUTLER'S OBJECTION

Any account of personal identity in terms of **MEMORY** is problematically circular.

To understand either version of the Memory Criterion, we need to know what it is for a person X to remember another person Y's experiences. But, by definition, X remembers Y's experiences only if X is numerically identical to Y. So in order to know what it is for X to remember Y's experiences, we need to **ALREADY KNOW** what it is for X to be numerically identical to Y!

In response, we might try to use the notion of **QUASI-MEMORY**: X quasi-remembers Y's experiences IFF what X seems to remember are Y's experiences. (This is Sydney Shoemaker's response to the problem.)

(A worry: won't an account made out in terms of quasi-memory overgenerate? Mad Heimson thinks that he is Hume. But he isn't Hume!

But it is very unlikely that what mad Heimson seems to remember are in fact Hume's experiences! And if they are...)

OTHER OBJECTIONS

Can a person survive irreparable memory loss? Even the revised account says "no".

But perhaps a more sophisticated psychological criterion could handle this.

Other worries concern the **MOTIVATION** for psychological views.

Locke's case of the prince and the cobbler seems **CONCEIVABLE**. But so what?

Recall Descartes's epistemological argument: why think conceivability is a guide to possibility?

And Bernard Williams suggested that intuitions about such cases depend on how they are told.

CASE 1

Suppose your memories and psychological and behavioural dispositions are somehow transferred to the body of person X, and vice-versa. Beforehand, you are told that, after the procedure, whoever inhabits one body will be given a reward, and whoever inhabits the other will be punished. You are allowed to choose which body gets the reward and which gets the punishment. Which body will you choose?

Intuition: we'd choose that body X gets the reward.

CASE 2

You are told that you will be tortured.

Intuition: this sounds bad!

CASE 3

You are told that you will be tortured, but only after your memory has been removed.

Intuition: this sounds bad too!

CASE 4

You are told that you will be tortured, but only after your memory has been removed, and replaced with those of someone else.

Intuition: this sounds even worse!

CASE 5

You are told that you will be tortured, but only after your memory has been removed, and replaced with those of someone else. That other person will be given a reward.

Intuition: now you are just rubbing my nose in it.

CASE 6

You are told that you will be tortured, but only after your memory has been removed, and replaced with those of someone else. That other person will be given a reward, but only after having **THEIR** memories removed and replaced with yours.

Intuition: this doesn't make things any better!

Williams's thought: but **CASE 6** is just **CASE 1**!

If Williams is right, our intuitions about these cases are not at all clear.

FISSION CASES

Derek Parfit's fission cases provide a major challenge to (all versions of) the psychological view.

TRANSPLANT

The left hemisphere of a person's brain is transplanted into another body. Call the original person Origin and the new person Lefty.

Lefty is psychological continuous with Origin in whatever way the psychological view requires, and so is, on that view, the same person as Origin.

FISSION

Both hemispheres of the brain are transplanted into new bodies. Call the new persons Lefty and Righty.

Both Lefty and Righty are psychologically continuous with Origin in whatever way is required, and so are both, on that view, the same person as Origin.

But now we're in trouble!

1. Lefty = Origin
2. Righty = Origin
3. So Lefty = Righty

But Lefty \neq Righty. Contradiction!

This shows psychological continuity between X and Y is not **SUFFICIENT** for X and Y being the same person.

Could it nevertheless be **NECESSARY**? X at t is numerically identical to Y at t', where t is later than t', IFF X is psychologically continuous with Y **AND** ...

Whatever the extra condition is, it will have to be a condition that at most **ONE** of Lefty and Righty satisfy.

But any condition that **JUST** one satisfies will be arbitrary. So it will have to be a condition that neither satisfies.

One option: the extra condition is a "no competitors" condition.

X at t is numerically identical to Y at t', where t is later than t', IFF X is psychologically continuous with Y **AND** there is no other object Z at t which is psychologically continuous with Y.

This nicely handles Parfit's cases:

- In **TRANSPLANT**, Lefty is the same person as Origin.
- In **FISSION**, neither Lefty nor Righty is the same person as Origin.

But, on this view, whether X and Y are the same person may depend on factors extrinsic to both. This may seem odd.

Parfit's own conclusion: "identity is not what matters".

What matters to us is what happens to beings psychologically continuous to us, regardless of whether or not they count as the same person.

ANIMALISM

According to **ANIMALISTS** (Paul Snowdon, Eric Olson), Person X at t is numerically identical to Y at t' IFF X and Y are the same animal

Animalists can allow that:

- You were once a foetus
- You weren't always a person
- You can be identical to something with which you are not psychologically identical

Animalism is not obviously the same as the view that what matters for personal identity is **BODILY CONTINUITY**.

- On that view, you will one day be a corpse.
- But the animalist can deny that the corpse is an animal.

Why be an animalist? The **TOO MANY THINKERS** argument:

1. I am standing here right now
2. There is a (human) animal standing here right now
3. There is only one thing standing here right now
4. So I am numerically identical to the animal that is standing here right now

Could we deny 3, and say that there are two (or more) things standing here right now?

- I am thinking (perceiving etc.) here right now
- The human animal standing here is thinking here right now
- Unless I am that animal, there are two things thinking here right now

And if there are two things thinking here right now, why is only one of them a **PERSON**?

What does the animalist say about the various cases?

In **TRANSPLANT** and **FISSION**, Origin is a distinct animal, so a distinct person to Lefty and Righty.

This might seem OK in the case of **FISSION**, but not so clear this is the right result in **TRANSPLANT**.

In Locke's case, the animal that has the prince's body on Monday is the same as the animal that has the prince's body on Tuesday.

Similarly, in Williams's cases, the animal that has your body before the procedure is the same animal as the one that has your body afterwards.

SUMMARY

We've looked at various issues surrounding personal identity - in particular, at answers to the question:

- What is it for person X at t to be numerically identical to Y at t'?

Locke attempts to motivate a simple **MEMORY CRITERION** by means of a transfer case, the prince and the cobbler.

Objections from Reid and Butler suggest Locke needs a more sophisticated version of the view, but this still has problems, particularly with Parfit's **FISSION** case.

There are also questions about the sorts of cases Locke appeals to. They seem to be **CONCEIVABLE**, but are they genuinely **POSSIBLE**?

And Williams argues that our intuitions about Locke-style transfer cases are highly sensitive to the way that the cases are presented.

Parfit suggests that what his fission cases show is that identity is not what matters to us.

ANIMALISTS suggest an alternative story altogether, according to which personal identity is a matter of being the same animal, and offer the **TOO MANY THINKERS** argument for their view.

Next week: free will.