

10 The ship of Theseus

Boy, did Theo have problems with that car he bought at Joe's! It started off with little things – a door lock needed replacing, some fiddly bits in the rear suspension fell off, the usual. Then bigger stuff started to go wrong – first the clutch, then the gearbox, finally the whole transmission. And there were plenty of knocks along the way, so the car was rarely out of the body shop. And so it went on – and on and on . . . Unbelievable. 'But not as unbelievable,' Theo ruefully thought, 'as the fact that the car's just two years old and every single bit has now been replaced. Hey, look on the bright side – maybe I've got a new car!'

Is Theo right? Or is it still the same car? The tale of the car of Theo – or, more usually, the ship of Theseus – is one of many puzzles used by philosophers to test intuitions about the identity of things or persons over time. It seems our intuitions in this area are often strong but conflicting. The story of Theseus' ship was told by the English philosopher Thomas Hobbes, who then elaborated further. To pick up Theo's version . . .

Honest Joe didn't live up to his name. Most of the bits he'd replaced in Theo's car were working fine, and he'd mended any that weren't. He'd saved the old parts and had been fitting them together. After two years, he'd assembled an exact copy of Theo's car. He thought it was a copy. Maybe it was Theo's car?

Identity crises Which is the original? The car Theo has, now built entirely of new parts, or Joe's version, built entirely of the original parts?

timeline

AD 1637

The mind-body problem

1644

Cogito ergo sum

1655

The ship of Theseus

It probably depends who you ask. Whichever, the identity of the car over time isn't nearly as neat and tidy as we might wish.

It isn't just a problem with cars and ships. People change enormously over a lifetime. Physically and psychologically, there may be very little in common between a 2-year-old toddler and the doddering 90-year-old who's taken his place 88 years later. So are they the same person? If they are, what makes them so? It matters – is it just to punish the 90-year-old for something he did 70 years earlier? What if he doesn't remember it? Should a doctor allow the 90-year-old to die because that wish was expressed 40 years earlier by a (supposedly) previous version of himself?

This is the problem of personal identity, which has exercised philosophers for hundreds of years. So just what are the necessary and sufficient conditions for a person at one time being the same person at a later time?

Animals and brain transplants The commonsense view is probably that personal identity is a matter of biology: I am now who I was in the past because I am the same living organism, the same human animal; I am linked to a particular body which is a single and continuous organic entity. But imagine for a moment a brain transplant – an operation we can envisage being within reach of future technology – in which *your* brain is transferred into my body. Our intuition is surely that *you* have a new body, not that my body has a new brain; if so, it seems that having a particular body is *not* a necessary condition of personal survival.

This consideration has led some philosophers to retreat from body to brain – to claim identity is linked not to the whole body but to the brain. This move fits our intuition regarding the brain transplant case but still does not quite do the job. Our concern is with what we suppose emanates from the brain, not with the physical organ itself. While we may be uncertain how brain activity gives rise to consciousness or mental activity, few doubt the brain somehow underlies that activity. In considering what makes me

1950

The Turing test

1981

The brain in a vat

The brave officer

Thomas Reid attempted to undermine Locke's position through the following story: *A brave officer was once flogged for robbing an orchard while a schoolboy; in his first military campaign he captured an enemy standard; late in life he was made general. Suppose he still remembered his flogging when he took the standard, but when made*

general he remembered capturing the standard but not being flogged.

Locke could accept the implication of Reid's objection: his thesis involves a clear distinction between human being (organism) and person (subject of consciousness), so the old general would in a real sense be a different person from the young boy.

me, it is the 'software' of experiences, memories, beliefs, etc. that concerns me, not the 'hardware' of a particular lump of grey matter. My sense of being me would not be much shaken if the total sum of those experiences, memories, etc. were copied onto a synthetic brain, or indeed if someone else's brain could be reconfigured to hold all my memories, beliefs, etc. I am my mind; I go where my mind goes. Based on this view, my identity isn't linked to my physical body, including my brain, at all.

Psychological continuity Taking a psychological approach to the question of personal identity, rather than a biological or physical one, let's suppose that each part of my psychological history is joined to earlier parts by strands of enduring memories, beliefs, etc. Not all (and perhaps none) of these need extend from start to finish; provided there is a single, overlapping lattice of such elements, then it remains my history. I remain me. The idea of psychological continuity as the main criterion of personal identity over time comes from John Locke. It is the dominant theory among contemporary philosophers, but is not without problems of its own.

Imagine, for instance, a *Star Trek*-style teleportation system. Suppose this records your physical composition down to the last atom and then transfers this data to some remote location (say from London, Earth, to Moonbase 1), where your body is exactly replicated (from new matter) at the precise moment your body in London is annihilated. All is well – provided you adhere to the psychological continuity thesis: there is an uninterrupted stream of memories, etc. flowing from the individual in London to the one

on the Moon, so psychological continuity and hence personal identity is preserved. You are in Moonbase 1. But suppose the transporter failed and neglected to carry out the annihilation in London. Now there are two of 'you' – one on Earth and one on the Moon. According to the continuity account, because the psychological stream is preserved in both cases, they are both you. In this case, we have little hesitation in saying that you are the individual in London while the one on the Moon is a copy. But if this intuition is right, we seem to be forced back from the psychological to the biological/animal account: it appears to matter that you are the old meat in London rather than the new meat on the Moon.

Getting your self straight Such mixed intuitions may come from asking the wrong questions, or applying the wrong concepts in answering them. David Hume drew attention to the elusiveness of the self, claiming that, however hard you look in on yourself, you can only ever detect individual thoughts, memories, experiences. While it is natural to imagine a substantial self that is the subject of these thoughts, he argues this is wrong – the self is no more than the point of view that makes sense of our thoughts and experiences, but cannot itself be given in them.

This idea of the self as a substantial 'thing', which we take to be our essence, causes confusion when we imagine ourselves undergoing brain transplants or being annihilated and reconstituted elsewhere. We assume our personal survival in such thought experiments somehow depends on finding a place for this self. But if we stop thinking in terms of this substantial self, things become clearer. Suppose, for instance, that the teleporter functions correctly in annihilating your body in London but produces two copies on the Moon. Asking which one is you (equivalent to 'where has my self ended up?') is simply asking the wrong question. The outcome is that there are now two human beings, each starting off with exactly the same fund of thoughts, experiences and memories; they will go their own way and their psychological histories will diverge. You (essentially the fund of thoughts, experiences and memories) have survived in the two new individuals – an interesting form of personal survival, achieved at the cost of your personal identity!

the condensed idea
What makes you you?