

Personal identity: An introduction

The syllabus asks two questions about personal identity: what does it take to be a person? and what does it take to persist as the same person over time? Our focus shall be primarily on the second, and even then, we will follow just one thread of thought – there are many others.

A FIRST DISTINCTION

In order to understand what is at issue in personal identity, it is important to distinguish between *numerical* identity and *qualitative* identity. Throughout life, we change what we are like as people, e.g. what I want may change in light of experience, some of my most fundamental beliefs, e.g. whether God exists or not, may change. In common speech, we often say things like ‘he was a different person after the cancer scare’. This use of ‘different person’ picks out a qualitative change. But it presupposes that there is just one person here, before and after the cancer scare – otherwise, who does ‘he’ refer to? Qualitative change presupposes numerical identity. Persons can persist through qualitative change. What it is for a person to persist is the question of numerical identity – what does it take for someone to be the *same* person in this sense?

THREE THEORIES OF PERSONAL IDENTITY

Dualism

Dualism claims that I am a mental substance, and the persistence of this mental substance is what personal identity consists in. I do not depend on any body in order to be and continue being me.

Psychological continuity

Materialism is the theory that the only substance is matter. A substance is something that can exist independently of anything else. Materialism denies that we have souls that can exist independently of our bodies. Because matter is the only substance, everything that exists must exist as a material object. However, even if I can only exist if I have a body, this doesn’t mean that I can only exist in *this* body. Just as a piece of computer software can be copied from one computer to another, perhaps I can exist in different bodies. On this theory, what makes me me is not what body I am in, but what is distinctive about me as a person, in particular, what psychological properties I have. So my memories, my desires, my emotions are all important to my personal identity. This is called the *psychological theory of personal identity*.

A famous historical example of the psychological theory was held by John Locke, who argued that it was memory that was central: if I remember doing something some years ago, then I am that same person. But I can’t remember everything I’ve ever done in my life. So it is not current memories alone that Locke relies on. I can now (t3) remember events from a previous time (t2). At t2, I was able to remember things from a still earlier time (t1). So I was at the same person at t2 that I was at t1; and I am the same person now, at t3, as I was at t2. So I am the same person now as I was at t1 even if I can’t now

remember anything that happened at t_1 . It is *overlapping chains* of memory that comprise personal identity. Locke used this criterion to criticise the dualist view. If somehow all the memories that your mental substance has were swapped with all the memories that my mental substance has, then surely I would 'go' with my memories, not the persisting substance.

But what, we may ask, of times that I don't have any memory of at all, e.g. when I'm asleep? Surely I'm still the same person I was when I was asleep last night. Sidney Shoemaker revises Locke's criterion by appealing to causal dependence: I have the memories I do today not only because I was the person who had the experiences these are memories of, but also because of my persistence through sleep. When I wake up, I inherit my psychological properties from the man who was asleep. We don't need to rely on just memories here, either, but can invoke the many types of psychological state that persist through time and have a causal influence on our future psychological states.

Physical continuity

The *bodily theory of personal identity* says that, although psychological properties are important, there must also be *physical continuity*. As with psychological continuity, at any point in time, my body is a little bit different from how it was before. Over a long period of time, it is made of completely different matter. But it is still the same body, because during each change in matter, most of my body stayed exactly the same. This is material continuity. This theory of personal identity says that I need to have material continuity in order to be the same person.

Two famous versions of the bodily theory approach the question of material continuity differently. One version, developed by Bernard Williams, singles out the brain as the matter that needs to continue, since it is the matter that supports psychological continuity. In order to persist, more than half of my brain needs to persist. The second version says that personal identity consists in being the same human animal. My body as a whole needs to satisfy the condition of material continuity, therefore, needs to persist for me to persist. This doesn't mean that I can't lose a limb or even several; but it needs to be fundamentally the *same organism*.

AN OBJECTION TO THE PSYCHOLOGICAL THEORY

How can we decide between these theories? Contemporary discussions of personal identity often start with a particular problem with the psychological theory that can be illustrated by a thought experiment. In *Star Trek*, people 'teletransport' from the *USS Enterprise* onto the surface of a planet and back again. Suppose the teletransporter malfunctioned, and instead of 'erasing' Captain Kirk onboard ship before recreating him on the planet's surface, it didn't erase him, but it also recreated him on the planet's surface. Which one of these two identical Captain Kirk's would be the 'real' one? Both seem to have a causal dependency on the psychological states of the original Captain Kirk. If psychological continuity is *all* that personal identity consists in, they are both Captain Kirk. But this is logically impossible – one person cannot become two persons, even if the two persons are qualitatively identical with the one person. This is because identity consists in numerical identity – and one thing is never two things!

Even if the teletransporter doesn't malfunction, the Captain Kirk on the planet's surface can't be the same person as the one that was onboard ship before teletransporting. We have just said that he wouldn't be that person if another Captain Kirk was created by the

teletransporter malfunctioning. But whether the Captain Kirk on the planet is the person he is cannot depend on someone else existing. We can't say 'he is Captain Kirk if the teletransporter didn't malfunction' but 'he isn't Captain Kirk if the teletransporter did malfunction'. Either he is or he isn't Captain Kirk. Identity must depend just on the relationship between the Captain Kirk before teletransportation and the Captain Kirk afterwards; it can't depend on something external, e.g. the creation of another person by a malfunction. Put simply: identity does not logically allow for duplication; psychological continuity does logically allow for duplication; therefore psychological continuity cannot be identity. If we think the Captain Kirk after teletransportation is the same person as the Captain Kirk beforehand, we are confusing *qualitative* identity with *numerical* identity.

Dualism can solve the problem by claiming that Captain Kirk remains Captain Kirk only if he retains Captain Kirk's soul. It would therefore most likely reject outright the possibility of teletransportation, unless a 'soul teletransporter' could be created. The bodily theory solves the puzzle in a different way. In the case of teletransporter malfunction, it is impossible for both 'Captain Kirks' to have bodily continuity with the original Captain Kirk, if we define 'bodily continuity' right. For example, if bodily continuity means more than 50% of the original matter continues in the body, then it is not possible for both bodies to have more than 50% of the matter in Captain Kirk's original body. So if we invoke the criterion of bodily continuity, we can avoid the problem of duplication. Teletransportation must involve bodily continuity (e.g. the transformation of matter into energy and then beaming the energy and transforming back into matter) in order to preserve personal identity.

REVISING THE BODILY CRITERION

However, we probably don't want to leave the criterion of bodily continuity as simply 'more than 50% of the matter in the original body'. Consider a case in which your brain is transplanted into another body, and the brain from that body is transplanted into your body. Which body are you now 'in'? If all your memories, beliefs, desires, etc. depend primarily on your brain, then our intuition will be that you 'go' with your brain – in other words, this is a case of a 'body transplant' rather than a 'brain transplant'. So the correct criterion of bodily continuity is actually 'more than 50% of your brain' persists.

However, there is a problem if we think we are essentially minds but we are invoking bodily continuity just to avoid the problem of duplication. For consider another thought experiment: your brain and mine are both erased of all psychological properties and then 'reprogrammed' so that your brain has all the psychological properties that mine had, and mine has all the psychological properties that yours had. If I 'go' with my psychological properties, I now have what was your brain (and body), while you have mine. So the criterion that identity requires more than 50% of the brain to continue can't be right. We might now think that a combined criterion is necessary – identity requires both psychological and material continuity. This means that in this case of reprogramming brains, identity is not preserved.

There is another problem facing any theory that invokes psychological continuity. If psychological continuity is necessary for personal identity, then I am not identical with the new-born baby whose body became my body, because that baby did not have a mind that is psychologically continuous with me. Another way of putting this is that the psychological theory denies that we are human animals. So what is my relationship to the animal that is thinking about this puzzle now? Are there two things here, me (not

continuous with the baby) and a thinking animal (continuous with the baby), thinking the same thoughts? That seems impossible. We might solve this puzzle by saying that we are essentially human animals, and that being a person is only a temporary property that we have (it starts to exist after we do); or we can say that we are persons, but part of what it is to be a person is to be a thinking animal; or we can say that there is no thinking animal here, because animals can't think, only persons think – so there is a person (me) and an animal (my body), but no thinking animal. You pays your money, you takes your pick...