The Hybrid Account of Personal Persistence

Abstract

In this paper we argue that persons should be defined as being things that are sometimes capable of first-person thought. We then defend an account (the Hybrid Account) of their persistence conditions. According to it psychological continuity and biological continuity are each sufficient for the persistence of persons, and their disjunction is necessary. We also discuss a recent paper by Olson and Witt (2020) and show that their arguments in that paper fail. Doing so enables us to clarify how the Hybrid Account relates to psychological and animalist accounts of personal persistence, and so helps to clarify the Hybrid Account further.

Introduction

What kinds of changes can persons survive? In particular:

Q1: Can they survive the destruction of their bodies?

Q2: Can they survive an irreversible loss of consciousness?

The philosophical literature contains excellent reasons for answering each of these questions “yes”. The main reason in favour of answering “yes” to Q1 is that we react to thought experiments involving the transfer of a person’s psychology into another body as being cases in which the person goes where their psychology does. Call this reason ‘the transplant intuition’. It has been persuasively developed by, e.g.

Shoemaker (1963). And the main reason in favour of answering “yes” to Q2 is that there are compelling reasons to think that persons would survive in a permanent vegetative state, e.g. in a coma in hospital. Call this ‘the coma intuition’. It has been persuasively developed by, e.g., Olson (1997).

So, it would be good if we had a theory that enabled us to answer both of these questions “yes”. And yet, in the philosophical literature, such theories are thin on the ground. Instead, the vast majority fall either into the Yes-No camp or the No-Yes camp.

In the Yes-No camp are, for example, Locke (1975), Shoemaker (1963, 1970, 1997, 2011, 2016), Parfit (1971), Perry (1972), and Noonan (2003). They all say that, in some sense or other, psychological continuity is both necessary and sufficient for the persistence of persons. Thus, a person could survive the destruction of their body so long as their psychology continued elsewhere (e.g. in another body). So, their answer to Q1 is “yes”. But a person cannot survive if their psychology comes to an end. So, their answer to Q2 is “no”.

In the No-Yes camp are, for example, Williams (1956-7), Thompson (1997), Olson (1997), and Snowdon (1991, 1995). They all say that, in some sense or other, bodily or biological continuity are both necessary and sufficient for the persistence of persons. Thus, a person could not survive the destruction of their living body. So, their answer to Q1 is “no”. But a person could survive an irreversible loss of consciousness if their living body is preserved (e.g. in an irreversible coma in a hospital). So, their answer to Q2 is “yes”.

In the first part of this paper we reject the theories from both camps. Instead, we defend the theory that psychological continuity and biological continuity are *each* sufficient for the persistence of persons, and that their disjunction is necessary. We

argue, that is, for a Yes-Yes theory. Because the 'Yes-Yes' theory is not a very memorable name, however, we call it the 'Hybrid Account' of personal persistence.

In the second part of this paper we then discuss a recent paper by Olson and Witt (2020). Discussing the arguments they present there, and showing why they fail, will enable us to clarify how the Hybrid Account relates to accounts from the other two camps, and so help to clarify the Hybrid Account further.

1. The Hybrid Account of Personal Persistence

As a preliminary, we first make it explicit that we think that how we should answer Q1 and Q2 is largely determined by our understanding and use of the concept of a person. We take this concept to be given by our understanding and use of the first-person reflexive pronoun "I". That is, we take the concept of a person that is relevant to debates about personal persistence to be the following concept: the object of first-personal thought. And we take the content of this concept, and thus its extension, to be determined by its use within our linguistic community. This is why we take the transplant and coma intuition so seriously. Considered from the first-person point of view each of us finds the thought "*I* would survive" to be highly plausible in both transplant and coma cases. And this, we take it, gives us strong evidence about how we apply the concept of a person, and specifically about what persistence conditions we associate with it. And so, we take it that this gives us strong evidence for our Yes-Yes Hybrid account.

So, on our view persons are objects of first-personal thought. But now, if something is an object of first-personal thought then it is a thing that is capable of thinking about itself. It is not required, however, that it is capable of thinking about itself

at *all* times at which it exists. All that is required is that it is capable of thinking about itself at *some* time at which it exists. So the term 'object of first-personal thought' can be expressed in an extensionally equivalent way as 'thing that is at some time capable of thinking about itself in the first-person' or, more neatly, 'thing that is sometimes capable of first-personal thought'. That is, we endorse the following equivalence:

- X is a person iff X is an object of first-personal thought iff X is sometimes capable of first-person thoughts.

Now, because we respect the coma intuition, we think there *are* persons that are *not* capable of thinking about themselves at the ends of their existences, i.e. those who fall into unthinking comatose states. So, on our view, persons are things that can lose the capacity for thought altogether. So, if a human person falls into an irreversible unthinking comatose state, it is no longer capable of thinking about itself in the first-person. But, before it fell into that state its first-personal thoughts had as their object a thing (i.e. itself) that includes those later unthinking stages as parts.[1] Similarly, on our view, persons are things that can start off as unthinking things, and only later gain the capacity for self-conscious thought. This is because as well as respecting the coma intuition, we also respect the foetus intuition, i.e. the fact that each of us thinks that it is overwhelmingly plausible that we were once unthinking foetuses that only later developed consciousness, then self-consciousness. So, at the beginning of its

[1] We do not mean to commit ourselves here to perdurantism by this talk of parts. We wish, rather, to stay neutral between endurantism and perdurantism, and we think that our view can be formulated adequately on either view (given a generous ontology). But we do sometimes help ourselves to the language of perdurantism for ease of expression.

existence a human person cannot think about itself (or indeed, about anything else). But, it can do so later when it develops and gains certain high-level psychological properties, and what it later thinks about when entertaining first-person thoughts is a thing (i.e. itself) that includes unthinking foetus stages as parts.

With the above understanding of personhood in place, then, we state the problem of personal persistence as follows:

> *The Problem of Personal Persistence*: What changes can things that are sometimes capable of first-person thought survive, and what changes bring about their destruction? (Equivalently: What changes can objects of first-person thought survive, and what changes bring about their destruction?)

In fact, it seems unlikely to us that there is any univocal answer to this question, for as we said above, it is one about the content of first-personal thoughts themselves and is thus determined, as the contents of all thoughts are, by complicated facts about how their constituent concepts are put to use by a linguistic community. And the point here is that there may well be communities of beings (i.e. on other planets) capable of first-personal thoughts who use first-personal concepts in different ways from how we use them here on Earth, and so who associate different persistence conditions with themselves. And in fact, it is difficult to know what to say about non-human animals like dolphins and great apes who appear to be capable of first-personal thought, but who may not associate any clear persistence conditions with themselves at all. For as we have said, our only insight into which persistence conditions we (i.e. human persons) associate with ourselves is given by our intuitions regarding which changes

we can and cannot survive, and it is unclear whether *our* intuitions about what changes dolphins and great apes can survive should be taken to determine what changes *they* can in fact survive. That is to say, *our* concept of what changes *we* can and cannot survive determines *our* persistence conditions, but we should not presume that *our* concept of what changes we can survive determines the persistence conditions of *all* persons.

Be the above as it may, we *do* have intuitions about what changes *we* can and cannot survive, and that is enough to be getting along with.

As already mentioned, our intuitions regarding what changes we can survive include the coma intuition, the foetus intuition, and the transplant intuition. It is because of the first of these that we hold that persons can survive an irreversible loss of consciousness. In effect, we have the intuition that we can survive whatever changes our living bodies can survive, and this gives us good evidence that we can indeed survive those changes. But we also have the intuition that we can survive our consciousness being transplanted into another body.[2] In effect, we have the intuition that we can survive the loss of our particular living bodies, so long as our consciousness continues elsewhere, and this gives us good evidence that we can in fact survive such a change.

It might be objected at this point that there are possible cases in which both of the following occur: our living bodies lose consciousness but nonetheless survive, *and* our consciousness continues in some other living body. And in this case, it might be said, we cannot survive twice over, for then each of us would be two things and not one. But the structure of this problem is identical to the familiar problem of branching

[2] Or, perhaps, even into a non-living computer system, or some such. But we ignore this here.

that everybody faces, no matter which theory of personal persistence they adopt. Those, such as animalists, who think that human persons survive if and only if their living bodies survive face the problem that it is possible for human bodies to split amoeba-like such that two bodily continuers result. And those, such as psychological theorists, who think that persons survive if and only if their consciousness continues elsewhere face the problem that a person's consciousness can continue in two separate bodies. One standard response to this problem, and one that we adopt, is to endorse the multiple occupancy thesis, viz. the thesis that before the branching takes place there are two persons present. (See, e.g. Robinson 1985)

We have said our use of the first-personal concept "I" determines the persistence conditions of human persons (i.e. us). This requires that those persistence conditions can be formulated as conceptual truths, i.e. as true *de dicto* necessities. And indeed, they can be so formulated. The problem of personal persistence is often put as if it is a problem about identity, i.e. as follows:

What is the relation R that satisfies the following schema:

Person x at t1 = person y at t2 iff person x at t1 bears R to person y at t2

But this problem is not really one about identity at all, but rather one about what changes persons can and must survive. The schema can be split into its two component necessary (right-to-left) and sufficient (left-to-right) conditions and the problem restated in two English questions as follows:

1. Under what conditions can a person at one time be identical with a person at another?
2. Under what conditions must a person at one time be identical with a person at another?

And these questions are equivalent to the following questions, now put in terms of the truth of *de dicto* necessities (remembering that a person is thing that at some time thinks about itself in the first-person):

1*. Which relations R satisfy the following schema: Necessarily, if x is a person, then if x exists at t1 and t2, then R(x,t1,t2) ?

2*. Which relations R satisfy the following schema: Necessarily, if x is a person, then if x exists at t1, then if some person y exists at t2, and R(x,t1,y,t2), then x=y ?

Specifying a relation that satisfies 1* gives us a sentence that expresses a *passing-away* condition for persons, i.e. one that places a constraint on how persons can vary across time. In ordinary English, it tells us that persons cannot survive certain changes. Specifying a relation that satisfies 2* gives us a sentence that expresses a *preservation* condition for persons, i.e. one that specifies how persons must vary across their temporal extents. In ordinary English, it tells us that persons must survive any changes so long as certain continuities nonetheless obtain. So, if we exhaustively specify every relation R that satisfies 1* and 2* we thereby specify precisely how persons cannot and can vary across their temporal extents. In ordinary English, we say precisely which changes persons can and cannot survive. And as all of this is specified in terms of *de dicto* necessities, all of this is laid down as being a matter of

conceptual truth. Put in such terms, it is simply part of our concept of a person that persons can and cannot survive such changes.

Now, those who endorse the animalist view and its ilk (i.e. that human persons survive if and only if their living bodies do) endorse (very roughly) the following view about how to specify relation R:

- Necessarily, if x is a person, then if x exists at t1 and t2, then x has a body at t1 and a body at t2 that are linked by chains of spatiotemporal/biological continuity and/or connectedness.
- Necessarily, if x is a person, then if x exists at t1, then if some person y exists at t2, and x at t1 has a body that is linked by chains of spatiotemporally/biologically continuity and/or connectedness to the body of y at t2, then x=y.[3]

And those who endorse the psychological account (i.e. that persons survive if and only if their psychology continues somewhere) endorse (very roughly) the following view about how to specify relation R:

- Necessarily, if x is a person, then if x exists at t1 and t2, then x has a psychology at t1 and a psychology at t2 that are linked by chains of psychological continuity and/or connectedness.

[3] In fact, this is not quite right, as a tweak is needed to allow for branching. But this can be done quite easily, and we ignore this complication here and in what follows. For more detail on this way of spelling out the debate about personal identity see Noonan and Curtis (2018).

- Necessarily, if x is a person, then if x exists at t1, then if some person y exists at t2, and x at t1 has a psychology that is linked by chains of psychological continuity and/or connectedness to the psychology of y at t2, then x=y.

Our point is simply that our concept of a person, as illuminated by the transplant, foetus, and coma intuitions, strongly suggests that neither of these accounts is correct, and that instead we ought to adopt the following view (again, very roughly stated):

- Necessarily, if x is a person, then if x exists at t1 and t2, then *either* (i) x has a body at t1 and a body at t2 that are linked by chains of spatiotemporal/biological continuity and/or connectedness *or* (ii) x has a psychology at t1 and a psychology at t2 that are linked by chains of psychological continuity and/or connectedness.
- Necessarily, if x is a person, then if x exists at t1, then if some person y exists at t2, and either (i) x at t1 has a body that is linked by chains of spatiotemporally/biologically continuity and/or connectedness to the body of y at t2, *or* (ii) x at t1 has a psychology that is linked by chains of psychological continuity and/or connectedness to the psychology of y at t2, then x=y.

This is the Hybrid Account. It is more complicated than animalism and the psychological account, for sure. But it is perfectly coherent. And given the strength of the transplant and coma intuitions, this does appear to correctly capture the persistence conditions associated with our concept of a person, and so give the correct answer to the problem of personal persistence. So, we endorse it.

We finish this section by considering two theses that have recently been discussed by Olson and Witt in their (2020). This will lead into the next section where we discuss the arguments contained within that paper in more detail. The two theses are these:

> *Weak person essentialism*: necessarily, if something is a person at a time, then there is no time at which it exists but is not a person
>
> *Strong person essentialism*: necessarily, if something is a person at a time, then necessarily there is no time when it exists but is not a person.

On our understanding of the concept of a person, the weak thesis is a trivial truth. It says that, necessarily, if something is sometimes capable of first-personal thought, there is no time at which it exists but is not sometimes capable of first-personal thought. To see that this is trivial take F to be the property of *being a thing that is sometimes capable of first-personal thought*. Then this thesis has the obviously valid form: If something is sometimes an F, then it is always sometimes an F.

As for the strong thesis, we reject it outright. It entails that nothing that is a person in the actual world could be a non-person in some other world. But note that on our view each of us was once an unthinking fetus. But, as seems clear, each of us could have died in the womb before developing consciousness or self-consciousness. As such, each of us could have died before we developed the capacity for first-personal thought. In such a case we would never have thought about ourselves in the first-person, and so would have failed to be persons. So, each of us is a thing that is a person, but each of us could have failed to be a person. Thus, strong person essentialism is false.

So, on our view weak person essentialism is trivial, and strong person essentialism false. And this is perfectly consistent with our Hybrid account of the persistence conditions of persons. But it is important to note that this is not a feature that is unique to our view. In fact, Shoemaker's well-known psychological account of personal persistence also has this feature. He too accepts that weak person essentialism is trivially true and rejects strong person essentialism, and this is consistent with the psychological account of the persistence conditions of persons given above. That we say these things would no doubt come as a surprise to Olson and Witt themselves, because they claim that Shoemaker (amongst other psychological theorists) holds a particular version of *both* weak and strong person essentialism called '*Lockean* person essentialism', and that this is *inconsistent* with the psychological account of personal persistence. However, these claims of Olson and Witt are false. It will prove instructive to see why, as it will enable us to spell out more precisely how our Hybrid account compares with both animalism and the psychological account.

2. Olson and Witt's Arguments

In 'Against person essentialism' (2020) Olson and Witt identify a view they call 'Lockean person essentialism'. They make three claims about it: First, that it is a false unquestioned dogma. Second, that it is inconsistent with the standard psychological account of personal persistence, even though the two are almost always held together. And third, that there is no easy 'intermediate' position between the inconsistent combination of Lockean person essentialism with the psychological

account on the one hand, and animalism on the other (the two ‘extremes’ (Olson and Witt 2020: 17)).

In this section we argue that:

1. Lockean person essentialism *is* false. But it is not an unquestioned dogma. Most contributors to the debate about personal identity over time, self-described psychological theorists included, need not be understood as accepting it. The evidence for its general acceptance is rather explained by the fact psychological theorists accept the definition of a person we have given above.
2. Lockean person essentialism *is* inconsistent with the psychological account, or, as Olson and Witt cautiously say, almost every possible version of that view, and every version that has ever actually been held by anyone. But:
3. An intermediate view between the two ‘extremes’ which denies Lockean person essentialism is not difficult to come by. The standard psychological account as espoused by Shoemaker is exactly such a view.

As should be clear from section 1, however, we are not defending the psychological account; we think it false for the reasons already given. But it is useful to work through Olson and Witt’s arguments in order to see why the Hybrid account we offer is the *real* intermediate view, rather than the view Olson and Witt identify as such (which is really just the standard psychological account).

Now, though we think that what Olson and Witt call ‘Lockean person essentialism’ is false, as already indicated in section 1, we think that weak person essentialism is true (in fact, trivially true) and consistent with the standard

psychological account of personal persistence. And it is, we think, merely weak person essentialism that the typical psychological theorist (including Shoemaker) adheres to. Why is this important? First, it means that there is not, as Olson and Witt claim, a fundamental and long-unnoticed incoherence in the standard psychological account. But secondly, and more importantly, as we will explain in more detail later, getting clear about this enables us to see that Olson and Witt set up the debate between the psychological theorists and animalists in a way that excludes what may be the most promising option, viz. our very own Hybrid view. Olson and Witt think the only options are: (i) the inconsistent conjunction of Lockean person essentialism with the psychological account, (ii) animalism, and (iii) an 'intermediate view' that they think is unsatisfactory. But, as we have said, we will argue that the 'intermediate view' they describe is just the standard, consistent, psychological account itself. Because of the way they set up the debate they miss the possibility of a position which rejects both the standard psychological account and animalism. This middle way, the Hybrid view that we have already outlined, stands between *their* 'intermediate view' and animalism, and is the *real* intermediate view. It retains the psychological theorists thought that psychological continuity suffices for personal identity but also the animalist view that it is not necessary.

So, now we turn to what Olson and Witt say about all of this. Olson and Witt derive Lockean person essentialism from person essentialism, the definitions of which we reproduce for the sake of clarity:

> *Weak person essentialism*: necessarily, if something is a person at a time, then there is no time at which it exists but is not a person

> *Strong person essentialism*: necessarily, if something is a person at a time, then necessarily there is no time when it exists but is not a person.

What these theses amount to depends upon what is meant by 'person'. Olson and Witt suppose that psychological theorists typically define 'person' much as Locke does, i.e. as 'a thinking, intelligent being with reason and reflection that can consider itself as itself, the same thinking thing, in different times and places'. Indeed, we do too, but we think that this is best captured by the notion of a thing that is *sometimes* capable of first-personal thought. However, Olson and Witt think that it is best captured without the temporal qualification, i.e. simply by the notion of a thing that is capable of first-personal thought. So, by plugging this definition of a person into the two person essentialism theses they define a weak and strong version of Lockean person essentialism thus:

> *Weak Lockean person essentialism*: necessarily, if something is capable of first-personal thought at some time, then there is no time at which it exists but is not capable of first-personal thought.
>
> *Strong Lockean person essentialism*: necessarily, if something is capable of first-personal thought at some time, then necessarily there is no time when it exists but is not capable of first-personal thought.

It is these two theses that Olson and Witt claim are standardly held by psychological theorists, despite the fact they are inconsistent with the standard psychological persistence conditions for persons.

Now, with regard to strong Lockean person essentialism, we have little to say. We do not know why Olson and Witt think that this thesis is standardly held by psychological theorists. Maybe there are some psychological theorists who hold it to

be true, but we know of no evidence to suggest that, e.g. Shoemaker, holds it to be true, and his account is a paradigm psychological account if any is. Certainly, Olson and Witt provide no evidence for the claim. But, at any rate, we need not quibble about that, because it is really weak Lockean person essentialism that does the work for Olson and Witt, for most of their arguments are aimed at showing that even the weak version is false and inconsistent with the psychological account. And, in addition, they do provide evidence that the weak version is held by paradigm psychological theorists like Shoemaker. So, in what follows, we focus, as Olson and Witt do, on the weak thesis.

Psychological theorists are, of course, committed to saying that you (if you are a person) were never an early stage (unthinking) foetus and will never fall into a permanent unthinking comatose state. But Lockean person essentialism implies that you were never a late stage conscious foetus either, and will never suffer from severe (stage 3) dementia where you retain consciousness but lose the capacity for first-personal thought. And, Olson and Witt argue, this is in fact inconsistent with the psychological account (or almost every possible version of that view and every version anyone has actually held). For you, as you are now, fully self-conscious and capable of first-personal thought, are psychologically continuous with the late stage foetus which, though conscious, did not possess the capacity to think about itself in the first-person (any more than a dog does). And if severe dementia lies in what anyone would call *your* future, you are now psychologically continuous with a human being which will be, though conscious, no more capable of first-personal thought than a dog.

In more detail, Olson and Witt argue for this as follows. Psychological continuity is defined via psychological connectedness. Psychological connectedness

is explained as causal dependence between later and earlier psychological states. But when you *first* became self-conscious you were psychologically connected in this sense with the foetus (or infant) as it was just previously. You inherited memories, preferences, and other mental states. Similarly, the devastation of severe dementia is not complete. The resultant being will be in mental states causally dependent on those you were in before your loss of higher consciousness. There will still be connections and hence continuity.

Of course, Olson and Witt point out, we could define a notion of connectedness that avoided this result. We might define connectedness, for example, only in terms of causal dependence between mental states that require self-consciousness and hence ensure that no self -conscious being is psychological continuous with a late stage conscious, but not self-conscious, foetus, or with someone suffering from severe dementia, simply by ignoring any causal dependencies between any non-self-conscious mental states. But no one has done this, and Olson and Witt explain, doing so would have a consequence that 'only a tough-minded metaphysician could seriously believe' (2020: 4): When your friend Sally is affected by severe dementia, though still having many psychological states causally dependent in the right sort of way on psychological states she had before, when she was self-conscious, in fact, on this story, she is dead. She has the same status she would have had if she had died and been cremated – the nursing home resident is a being you have never seen before.

This argument is entirely convincing. Weak Lockean person essentialism is inconsistent with the psychological account. But all that this shows is that psychological theorists, to be consistent, should reject it. And we think it is clear how they should do so. They should simply reject the definition of a person that Olson

and Witt attempt to foist upon them and instead accept the definition of a person we have given, and thus accept weak person essentialism in the form that we accept it, i.e. as a trivial truth.

In fact, we think that this is precisely Shoemaker's position. To see this, consider that Shoemaker notes that the transplant intuition is also plausible if we suppose it is dogs that have their psychologies transferred into new bodies rather humans. As Shoemaker says (2011: 370):

> There is no word for dogs that has a definition similar to Locke's definition of person, one that would lead one to expect a psychological account of their persistence conditions. Still brain transplants involving such creatures are imaginable and the conditional proposition that [psychological theorists] are committed to in the case of persons, namely that if such a transplant resulted in full psychological continuity between donor and recipient then transfer would be mental-subject-preserving, seems plausible here as well … if the recipient … recognizes the owner of the original dog … fawns on them … knows its way around their house … digs for bones where the donor buried them … it would be hard to deny that it is the old dog in a new body.

But how can we say this whilst acknowledging that dogs are animals, as we must? Shoemaker explains:

> … we can say that dogs are animals in the same sense as persons are but can deny that their persistence conditions are biological rather than psychological, and so can deny that they are biological animals. Dogs, like persons, will coincide with biological animals, but will not be identical with them. The same will be true of chimpanzees …

What this overwhelmingly suggests is that Shoemaker thinks of psychological continuers as a general sort of thing under which persons (e.g. the minded beings embodied in human bodies) fall, but under which also fall other minded beings that stand to their bodies in the same way that persons stand to their bodies, but for which we have no specific term (i.e. the beings embodied in dog bodies, chimpanzee bodies etc). For all of which, he strongly suggests, psychological continuity suffices for survival. And so, on this suggestion, we should think of our lives as possibly extending backwards and forwards in time to periods when we lack the capacity for first-personal thought, which other individuals of the sort *never* possess. And thus, it seems, he thinks of persons in a very similar way to the way in which define them, viz. as being things that *sometimes* possesses the capacity for first-personal thought.

Nevertheless, Olson and Witt argue there is evidence in the writings of Shoemaker, and indeed in basic textbook level presentations of the problem of personal identity, that Lockean person essentialism is presupposed, though inconsistent, with the psychological account. But this is not so; the evidence they point to is consistent with the assumption that what the writers in these debates are concerned with is only the persistence conditions of persons as *we* have defined them, and that the only essentialist thesis being presupposed, if any is, is the trivial version of weak person essentialism that we have outlined and that we ourselves are committed to.

One passage from Shoemaker (2011: 360) that they point to, and in which they say he explicitly endorses Lockean person essentialism, reads as follows. After giving Locke's definition of a person, Shoemaker writes:

> Certainly, we conceive of persons as creatures that have, and in some sense necessarily have, mental or psychological properties. Of course, an animalist can hold a version of this view; when the person becomes a human vegetable, entirely devoid of mentality, the animalist can say that it ceases to be a person but does not cease to exist. On such a view Locke's definition only gives us the nominal essence of persons. One might think that the nominal essence reading is recommended by Locke's claims about the unknowability of real essences. But the rest of what Locke says in this chapter [Chapter xxvii of Book II of the *Essay*] goes better with the view that it is necessary *de re* of persons that they are beings that satisfy his definition …. And that, suitably qualified, seems to me an intuitively plausible view.

But does Shoemaker endorse Lockean person essentialism here? Hardly. His comments are intended simply to signal that, for the psychological theorist, the *thing* that is a person ceases to exist when it loses consciousness, whilst for the animalist the *thing* that is a person does not cease to exist. This is the difference Shoemaker wants to stress by contrasting real with nominal definitions, and the *de re* with the *de dicto*. But what he says here is perfectly consistent with his holding the view that persons are things that are sometimes capable of first-person thought. To say that his words contain an explicit commitment to Lockean person essentialism, as Olson and Witt define it, is to push too hard.

Olson and Witt also cite a standard textbook statement of the problem of person identity as evidence that Lockean person essentialism (they actually just write 'person essentialism' but they cannot mean our trivial weak person essentialism) is commonly presupposed:

> The problem of personal identity over time is the problem of giving an account of the logically necessary and sufficient conditions for a person identified at one time being the same person as a person identified at another. (Noonan 2003: 2)

This, of course, presupposes that it makes no sense to talk of a thing's being a person at one time and not at another; it presupposes that personhood need not be temporally qualified (Olson and Witt 2020: 2). But it does not presuppose Lockean person essentialism. Again, it is consistent with 'person' meaning 'thing that is sometimes capable of first-personal thought'.

Olson and Witt quote another passage from Noonan (2003) as evidence of Lockean person essentialism being presupposed:

> [W]hen philosophers speak of the problem of personal identity they do not use "person" in the sense of "human being". Rather they use it in the sense introduced by Locke. (Noonan 2003: 8)

But once more here, the wording, though vague, does not necessarily presuppose what Olson and Witt claim. The thought expressed here is this: when philosophers use the word 'person' they do not use it as a mere synonym of 'human being', for they allow the conceivability of non-human persons (e.g., intelligent extra-terrestrials and angels); rather, they follow Locke in thinking that the various capacities he mentions lie at the core of personhood. But this too is compatible with 'person' meaning 'thing that is sometimes capable of first-personal thoughts'. For, of course this definition does contain essential reference to the Lockean capacities.

In relation to the above, Olson and Witt also consider how to reformulate the traditional statement of the problem of personal identity *without* the presupposition of Lockean person essentialism. But, in fact, no reformulation is needed if 'person' is read as we recommend. Similarly, if we read in this way the typical formulation of the psychological-continuity account they give, i.e.,

> Necessarily, if a person *x* exists at one time and a person *y* exists at another, *x* is *y* iff *x* is in some way psychologically continuous, at the first time, with *y* as it is at the other time, (Olson and Witt 2020: 4),

this needs no reformulation either, as it gives us exactly what we want from an account of personal identity over time.[4] It tells us what changes a thing that is sometimes capable of first-personal thought can survive and what changes it cannot survive. It tells us, in fact, two things – precisely those two things that we have said psychological theorists maintain, i.e. that:

- Necessarily, if x is a person, then if x exists at t1 and t2, then x has a psychology at t1 and a psychology at t2 that are linked by chains of psychological continuity and/or connectedness.

[4] Setting aside, as Olson and Witt note (2020: 4), 'complications due to branching'.

- Necessarily, if x is a person, then if x exists at t1, then if some person y exists at t2, and x at t1 has a psychology that is linked by chains of psychological continuity and/or connectedness to the psychology of y at t2, then x=y.

So far, then, we have argued that the standard psychological account, as defended in the literature, is one that does not presuppose Lockean person essentialism and that is, rather, entirely consistent with our definition of a person (and thus the triviality of weak person essentialism). Moreover, we have argued that this is the position Shoemaker, a paradigm psychological theorist, holds. However, Olson and Witt do consider what the psychological account *minus* Lockean person essentialism looks like. It is this view that they call 'the intermediate view'.

According to the so-called 'intermediate view', we are fundamentally psychological beings, but in a weaker sense than Lockean person essentialism implies, i.e. in the sense that even if no particular mental powers such as intelligence or self-consciousness are essential to us, *having some mental power or other* is. This is indeed entailed by the standard psychological account, for on that view psychological continuity is necessary for persistence, and so persons cease to exist when their psychology comes to an end. And so, on this view, we cannot exist without some mental state or other. The so-called 'intermediate view' therefore entails a version of weak *psychological* essentialism. This can be spelled out as:

> Weak psychological essentialism: necessarily, if something is a person at a time, then there is no time at which it exists but is not conscious.

As we have been at pains to emphasise, *contra* Olson and Witt, a commitment to this principle is not part of a novel 'intermediate' view that lies between the two 'extremes' of the standard psychological account and animalism. Rather, it just is the standard psychological account to which reflection on transplant cases naturally leads us if it leads us to a psychological continuity account at all. Nonetheless, Olson and Witt do provide reasons to reject it, and so perhaps they have a point to make here after all. Whether they do or not depends on their reasons. So, we now consider them. There are four.

First, they point out that this view implies that no biological organisms have mental powers, not even ones physically indistinguishable from us. This *is* a consequence of the standard view that we persons (i.e. beings sometimes capable of first-person thought) are psychological continuers. The biological organism coincident with me was once a mindless foetus. So, it is not a psychological continuer. So, it is not a *person*. So, it is not even now, when physically indistinguishable from me, a being that is capable of first-personal thoughts. And if it does not now possess that mental power it cannot possess any. But if the complex biological organism coincident with me now lacks mental powers, no biological organism (or any that we know of – Olson and Witt 2020: 17) ever has mental powers. However, this is not a new problem for psychological theorists. And, in fact, this consequence is a *feature* of the standard view, not, its defenders like Shoemaker claim, a flaw. The view just *is* that we and other animals are psychological continuers, but we coincide temporarily (and sometimes permanently) with mindless biological organisms, which we might call, using a term of art, 'biological animals' (Shoemaker 2011: 370), just as statues coincide temporarily (and

sometimes permanently) with hunks of clay, which though coincident are numerically distinct because differing in some properties.

Second, Olson and Witt point out that according to the so-called 'intermediate view', acquiring the higher level powers that makes something a person creates no new entity, and losing them destroys none; people as such have no 'ontological significance'. Again, however, this is not a problem but a *feature* of the view, not a flaw. Olson and Witt cite Baker (2000) as thinking that it *is* a problem, but this doesn't make it one. And anyway, Olson and Witt themselves effectively respond to Baker:

> "ontological significance" is a technical term. Baker's statement that Fs as such have no ontological significance in fact means nothing more than that Fs are only contingently F. (For lumberjacks as such to have no ontological significance is for them to be only contingently lumberjacks.) What she presents as an unattractive consequence of denying person essentialism is only a restatement of that denial in dyslogistic terms. Its force is purely rhetorical. (Olson and Witt, 2020: 14)

Third, Olson and Witt point out that the so-called 'intermediate view' rules out any account of personal identity as such, i.e. any account that applies *only* to persons. According to the view, we are psychological continuers and so share our persistence conditions with things that are never self-conscious beings, like dogs. Again, a feature not a flaw, unsurprising in the light of the transplant intuition, even if some (Baker 2000, Gert 1971, Johnston 2010) say that this makes the view not even a contender for an account of personal identity over time. In other words, these

people are simply wrong. From the standard psychological theorists viewpoint saying this is comparable to insisting that there must be an account of the persistence conditions of *musical geniuses* as such.

Fourth and finally, Olson and Witt point out that the intermediate view is incompatible with Parfit's influential account of personal identity, that a person persists for a day only if she exhibits a certain number of direct psychological connections. They are right, but the particular view of Parfit's they explain is that 'you persist just if you have at least half as many direct psychological connections … over a day as there actually are during a day in the life of nearly every person' (Parfit 1984: 206). But a psychological continuity theorist should surely not be regarded as committed to this (charmingly precise) condition for persistence.

So, to summarise, we do not think Olson and Witt's 'intermediate view', i.e., the standard psychological account according to which all persons are psychological continuers, has been shown by Olson and Witt in their (2020) to have any problems which were not already familiar.

However, as we have said, we do not accept the psychological account, for we respect the coma intuition. Each of us may one day fall into a permanent unthinking state. Psychological theorists may be able to explain senses in which this statement is true, but the difficulty is that they have also to acknowledge senses in which they are false. But there are no such senses. It is just true that each of us may fall into an unconscious unthinking state. Psychological continuity is not necessary for personal identity.

So, should we be animalists? No. The transplant intuition is compelling. Even if psychological continuity is not necessary for persistence, some form of it suffices. We need a middle way.

Participants in the personal identity debate need not assume that there is an account of the persistence conditions of persons as such. But they must assume that there is an account which is true of *all* persons, even if it is true of other things as well. But, when we think about all the kinds of persons which seem conceivable (human, dolphin, extra-terrestrial, robotic, angelic, divine) is this plausible?[5] Boethius defined a person as an individual substance of rational nature. Perhaps the persistence conditions of a person depend upon the kind of individual substance it is.

The response of the psychological continuity theorist will be that what makes their account compelling is the transplant intuition and this also shows that there *must* be a common sort to which all persons belong for which persistence conditions can be given. However, what the transplant intuition makes compelling is only that some form of psychological continuity *suffices* for personal persistence. It is consistent with it that the common sort is one for which it is not necessary as well. We think that is what must be correct, given the coma intuition. Olson and Witt think that the only options are: (i) the inconsistent conjunction of the psychological account with Lockean person essentialism, (ii) animalism and (iii) their 'intermediate view' (which is really just the standard psychological account). Because of the way they have set up the debate however, they have missed another possibility: a position which rejects both the psychological account (because it conflicts with the coma intuition) and the animalist view (because of its conflicts with the transplant intuition).

[5] See Noonan (1978: 351)

This middle way – the Hybrid view we outlined in section 1 - is the *real* intermediate view. It is intermediate between two consistent views. It retains the thought of the psychological account that psychological continuity suffices for personal identity, which is all that the transplant intuition supports, but adds that physical continuity also suffices, which is all that the coma intuition supports. It is, we think, the best of both worlds.[6]

References

- Baker, Lynne Rudder 2000, *Persons and Bodies* (Cambridge: Cambridge University Press)
- Gert, Bernard 1971, 'Personal Identity and the Body', *Dialogue* 10: 458-478
- Johnston, Mark 2010, *Surviving Death* (Princeton: Princeton University Press)
- Locke, John 1975, *An Essay Concerning Human Understanding* (Oxford: Oxford University Press) (2nd ed. first published 1694)
- Noonan, Harold & Curtis, Benjamin L. (2018). 'The Simple and Complex Views of Personal Identity Distinguished'. In Valerio Buonomo (ed.), *The Persistence of Persons Studies in the Metaphysics of Personal Identity Over Time*. Germany: Editiones Scholasticae, 21-40.
- Noonan, Harold 1978, 'Locke on Personal Identity', *Philosophy* 53: 343-351
- Noonan, Harold 2003, *Personal Identity*, 2nd ed. (London: Routledge)
- Noonan, Harold 2021, 'Personal Identity and the Hybrid View: A Middle Way' *Metaphysica*, volume 22, (2)
- Olson, Eric 1997, *The Human Animal* (New York: Oxford University Press)

[6] See also Noonan (2021) where some of the details of the Hybrid Account have also been worked out.

- Olson, Eric. and Witt, Karsten 2020, 'Against person essentialism', *Mind* https://doi.org/10.1093/mind/fzaa016
- Parfit, Derek. (1971), 'Personal Identity' *The Philosophical Review*, Vol. 80, No. 1, 3-27
- Parfit, Derek 1984, *Reasons and Persons* (Oxford: Oxford University Press)
- Perry, John (1972), 'Can the Self Divide?' *The Journal of Philosophy*, Vol. 69, No. 16, 463-88
- Robinson, Denis, 1985 'Can amoebae divide without multiplying'. *Australasian Journal of Philosophy* 63: 299-319
- Shoemaker, Sydney, 1963, *Self Knowledge and Self-Identity (*Ithaca NY: Cornell University Press)
- Shoemaker, Sydney (1970) 'Persons and Their Past' *American Philosophical Quarterly*, Vol. 7, No. 4, 269-85
- Shoemaker, Sydney (1997), 'Parfit on Identity' in Dancy (ed.) *Reading Parfit*, (Oxford: Blackwell) 135-48
- Shoemaker, Sydney, 2011, 'On What We Are', in Shaun Gallagher (ed.), *The Oxford Handbook of the Self* (Oxford: Oxford University Press), 352-371
- Shoemaker, S. 2016. 'Thinking animals without animalism'. In S. Blatti and P.Snowdon (eds) *Animalism: New Essays on Persons, Animals, and Identity.* Oxford: Oxford University Press
- Thompson, Judith J. (1997), 'People and Their Bodies,' in Dancy (ed.) *Reading Parfit*, (Oxford: Blackwell), 202-29

- Snowdon, P. F. (1991). ‘Personal Identity and Brain Transplants’. *Royal Institute of Philosophy Supplement* 29, 109-126.
- Snowdon, Paul F (1995), ‘Persons, Animals, and Bodies’ in Bermúdez, Marcel & Eilan (eds.) *The Body and the Self, Cambridge*, MA: MIT Press, 71-86
- Williams, Bernard A. O. (1956-7), ‘Personal Identity and Individuation’ *Proceedings to the Aristotelian Society*, Vol. 57, 229-52