



This is a repository copy of *Against Person Essentialism*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/159376/>

Version: Accepted Version

---

**Article:**

Olson, E. and Witt, K. (2020) *Against Person Essentialism*. *Mind*, 129 (515). pp. 715-735.  
ISSN 0026-4423

<https://doi.org/10.1093/mind/fzaa016>

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Against Person Essentialism

ERIC T. OLSON\*  
*olson.et1066@gmail.com*

KARSTEN WITT\*  
*University of Applied Sciences for Police and Public Administration, North Rhine-Westphalia*  
*karsten.witt@hspv.nrw.de*

It is widely held that every person is a person essentially, where being a person is having special mental properties such as intelligence and self-consciousness. It follows that nothing can acquire or lose these properties. The paper argues that this rules out all familiar psychological-continuity views of personal identity over time. It also faces grave difficulties in accounting for the mental powers of human beings who are not intelligent and self-conscious, such as foetuses and those with dementia.

## 1. Person essentialism

‘Person essentialism’ is the claim that every person is a person essentially. It comes in a weaker and a stronger form. The weaker says that nothing can be a person at one time and a nonperson at another: no person can cease to be a person without ceasing to exist, and no nonperson can become a person. The stronger says that nothing can be a person in one circumstance and a nonperson in another: no person could have been a nonperson and no nonperson could have been a person. We can state them like this:

Weak: Necessarily, if something is a person at a time, then there is no time when it exists but is not a person.

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

The stronger form entails the weaker but not vice versa. If personhood is a property you could not lack, then you cannot gain or lose it. But it might be a property you cannot gain or lose even though you could either have or lack it throughout your life, like *being at some time a lumberjack*. The two claims are nearly always held together.

\* Olson and Witt are equal co-authors and their names appear in alphabetical order.

Their import depends on what it is to be a person. Most philosophers say that it is to have certain special mental powers: a person is something that is intelligent and self-conscious in a certain way, where self-consciousness is the ability to think about oneself in the first person. You have that ability even when you are asleep and not exercising it. But human embryos don't have it. They have only the capacity to develop it. The tradition derives from Locke, who famously defined 'person' as 'a thinking intelligent being, that has reason and reflection, and can consider itself as itself' (Locke 1975, p. 335).

Our interest is in the combination of person essentialism with Lockean personhood. It says that nothing could have the mental powers that make something a person only temporarily or contingently, so that whatever is a person in the Lockean sense is so essentially. Call this *Lockean person essentialism*.

## 2. Person essentialism in current debate

Lockean person essentialism pervades contemporary discussions of personal identity. Many endorse it explicitly.<sup>1</sup> More often, though, it is tacitly presupposed by the terms of the debate. Nearly all proposed definitions of personhood, for example, have the form:

$x$  is a person =<sub>df</sub> ...,

a typical completion being something like 'x is intelligent and self-conscious'. Taken at face value, this implies that whatever is a person is a person *simpliciter*, without temporal qualification. But your being a person *simpliciter* rules out your being a nonperson at another time: if you had once been a nonperson, your being a person must amount to your being a person *now*. So your being a person without temporal qualification would entail the weaker form of person essentialism. To leave open the possibility that a thing might become or cease to be a person, we should have to say instead what it is to be a person at a time, completing the formula:

$x$  is a person at time  $t$  =<sub>df</sub>....

Yet almost no one defines personhood in this way. Person essentialism makes the two formulations equivalent: if nothing can be a person only temporarily, the temporal qualification adds nothing.

<sup>1</sup> For example, Baker (2000, p. 11), Buchanan and Brock (1990, pp. 159f.), Lowe (2012, p. 146), Rosenberg (1998, p. 49), Sharpe (2015, p. 65), and Shoemaker (2011, p. 360). Nozick's view (1981, pp. 78f.) is similar, and has some of the same troubling consequences.

It may be that authors omit the temporal qualification only to simplify their formulations, without meaning to assume person essentialism. They define ‘person’ as a ‘being that is intelligent and self-conscious’ in the way that we might define ‘lumberjack’ as ‘someone whose job is cutting down trees’. No one would take this to imply that lumberjacks must be lumberjacks without temporal qualification, or to rule out the possibility that someone might, by changing her employment, become or cease to be one. We omit the qualification because everyone knows this, and would add it in a context where precision was important. And it may be the same with definitions of personhood.

Be that as it may, another well-known problem, that of personal identity over time, is almost invariably stated in a way that clearly presupposes person essentialism. Here is a representative statement from a highly regarded introduction:

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, p. 2)

Derek Parfit begins his discussion of the problem by asking ‘What makes a person at two different times one and the same person?’ (1984, p. 202; see also 2012, pp. 6f.). Locke himself took the question to be ‘what makes the same person’ (1975, p. 336). Galen Strawson calls a similar formulation ‘the canonical personal identity question’ (2011, p. 77), and it is enshrined in nearly all philosophical reference works.<sup>2</sup>

The ‘canonical question’ takes cases where a person exists at one time and a person exists at another, and asks what is necessary and sufficient for them to be one person rather than two. Unless we presuppose person essentialism, however, this question will not ask what sort of thing you could survive, but only what you could survive *and remain a person* (Mackie 1999, pp. 224-7). It will not ask what it takes for something that is not a person at another time—a corpse, say, or an embryo—to be you. It will ask under what circumstances you could exist at another time and be a person, but not under what circumstances you could exist at another time without being a person.

<sup>2</sup> For example, Blackburn (2005), Bunnin (2004, p. 512), Flew and Priest (2002, p. 302), Garrett (1998; 2005, p. 459), Penelhum (1967, p. 95), D. Shoemaker (2019), S. Shoemaker (2015), Snowdon (2005), and P. F. Strawson (1995).

But this is not how friends of the canonical question understand it. Here is a typical answer:

Necessarily, if a person  $x$  exists at one time and a person  $y$  exists at another,  $x$  is  $y$  if and only if  $x$  is in some way psychologically continuous, at the first time, with  $y$  as it is at the other time.

(Set aside complications to do with branching.) Those giving answers like this do not intend to say only what it takes for someone to survive and remain a person. They mean that psychological continuity is necessary for us to continue existing at all. If you lose all psychological continuity, that's the end of you: you don't just cease to be a person. Yet if person essentialism were false, the answer would be compatible with your one day being a corpse, even though no corpse could be psychologically continuous with you. Because a corpse is not a person (in the Lockean sense), the answer would say nothing about how you could relate to one. It would be equally consistent with your having been a microscopic embryo, again without psychological continuity. It would tell us nothing about what would happen to you if you ceased to be a person, or what you might have done before becoming one. For all it would say, you could survive complete psychological *discontinuity*, as long as this happened at a time when you were not a person (Olson 1997, p. 26). Without the assumption of person essentialism, answers to the canonical question would be completely silent about such cases. None of those posing the question or giving anything like the typical answer would accept this. They ask what it takes for a past or future person to be you because they assume that *only* a person could be you: you could not exist at a time without being a person then.

To ask about our persistence conditions without presupposing person essentialism, we must replace the canonical question with this neutral one:

If something is a person at one time and something exists at another time (whether or not it is a person then), what is necessary and sufficient for them to be one and the same?

Person essentialism makes the two questions equivalent: if you could not possibly exist without being a person, there is no difference between what you could survive while remaining a person and what you could survive *simpliciter*. The only reason to prefer the canonical question over the neutral one is the presumption of person essentialism. Yet very few philosophers ask the neutral question, showing how widespread this presumption is.

So person essentialism is widely held. And so is the Lockean conception of personhood. The same overview we quoted earlier states that ‘when philosophers speak of the problem of personal identity they do not use “person” as a mere synonym of “human being”. Rather they use it in the sense introduced by Locke’.<sup>3</sup> Person essentialists are almost always Lockean person essentialists.

There are, of course, non-Lockean accounts on which intelligence and self-consciousness are either unnecessary for personhood (Wiggins 1980, p. 171; Oderberg 2005, p. 81; Schechtman 2014, pp. 119-25) or insufficient (Lewis 1976, p. 22; Noonan 1998, p. 316). We have no objection to them or to their combination with person essentialism. They are not our concern. Henceforward we will use ‘person’ to mean ‘person in the Lockean sense’.

While it’s clear that Lockean person essentialism is the dominant view, the reasons for its dominance are less evident. There is nothing obvious about it. It’s a tautology that whatever is a person is a person, but not that whatever is a person *must* be a person if it exists at all. Nothing in the Lockean account of personhood rules out someone’s losing the mental powers that make her a person and continuing to exist as a nonperson, just as a lumberjack might give up felling trees and take up dressmaking instead. And arguments for the view are scarce. What makes it so magnetic?

It may seem to follow from the widely held view that our persistence consists in a sort of psychological continuity: that a past or future being is you just if she is in some way psychologically continuous, then, with you as you are now (in the absence of branching, perhaps). That would rule out your existing without any mental properties at all—as an embryo or a corpse, say—as you cannot be psychologically continuous with such a being. But as it would not rule out your existing with mental properties falling short of personhood, it does not entail person essentialism. We will show in a moment that psychological-continuity views are actually inconsistent with person essentialism.

A better proposal starts by asking what we are ‘most fundamentally’. Whatever exactly this amounts to, the answer would presumably have to be a kind we belong to essentially: no kind that a thing belonged to only contingently would count as fundamental in the relevant sense.

<sup>3</sup> Noonan (2003, p. 8). See also Glover (1988, p. 14), Parfit (1984, p. 202; 2012, p. 6), McMahan (2002, p. 46), G. Strawson (2009, pp. 59f.), and Sharpe (2015, p. 56), in addition to those cited in footnote 1.

And one natural answer is that our fundamental kind (or one of them, if we could have several) is *rational being*: a thing with certain cognitive powers, including the ability to think about itself in the first person—in other words, *person* in the Lockean sense.

But our fundamental kind might be *rational being* in a weaker sense than actually having these powers: in the sense of having the capacity to acquire them, perhaps. Nor is *rational being* the only plausible candidate for our fundamental kind. Why not *human being*—a thing with a certain biological nature? Members of this kind are not necessarily rational, even if that is their normal condition. Neither of these alternatives would support person essentialism.

We suspect that Lockean person essentialism is nothing more than an ancient and unquestioned dogma.

### 3. The psychological-continuity problem

Whatever its appeal, the view faces serious objections. First, it is inconsistent with the view that our persistence consists in a sort of psychological continuity. This is especially surprising as the two views are almost always held together. And both are, of course, associated with Locke. Yet the claim that we are essentially people in the Lockean sense rules out all proposed views of this sort.

Psychological continuity is usually defined in terms of causal dependence (Shoemaker 1984, pp. 89f.). A mental state or event can be causally dependent on an earlier one: the memory of a past experience, for example, will be causally dependent on the experience, and your current ambivalence towards lumberjacks is causally dependent on the attitude to the same effect that you had last month. Then we can say that a thing  $x$  is *psychologically continuous*, at a time  $t$ , with a thing  $y$  as it is at an earlier time  $t^*$  just if the mental states  $x$  is in at  $t$  are causally dependent in the right way on the mental states  $y$  is in at  $t^*$ , or when they are connected by a chain of such causal dependencies. (The qualification ‘in the right way’ is meant to rule out causal dependence of a sort that has nothing to do with a person’s persistence. Many of your current beliefs are causally dependent on beliefs your mother had when you were a child: you have them now because of what she believed then and accordingly taught you. But that is no reason to suppose that you *are* your mother: it’s the wrong sort of causal dependence.)

Now consider those with severe (‘stage-3’) Alzheimer’s dementia. They are deprived of intelligence and self-consciousness to a drastic

extent (Reisberg *et al.* 1982, Reas 2017, WHO 2019). As bioethicists generally acknowledge, this makes them nonpeople in the Lockean sense (for example, Kuhse (1999, p. 356), DeGrazia (2005, pp. 6, 164), and Schechtman (2010, p. 276)). If no person can become a nonperson, then no person could survive in this condition: to get severe dementia is not just to cease to be a person, but to stop existing altogether. The resulting being, though alive and conscious, must be a numerically different entity (DeGrazia 2005, p. 167).

But even if severe dementia destroys most of your mental powers, including self-consciousness, the devastation is not complete. The resulting being will still be in mental states causally dependent on those you were in shortly before the loss. He or she will inherit some of your memories, preferences, and tastes. And this causal dependence will be secured by the continued functioning of your nervous system—the right way by anyone’s lights. This will make her psychologically continuous with you as you were when you were still a person (DeGrazia 2005, p. 168). It follows from almost any psychological-continuity view of personal identity that she will be you, despite not then being a person—contrary to Lockean person essentialism in both its strong and weak forms.

The same goes for the beginning of life. You grew from a certain foetus. At some point during her prenatal development, that foetus acquired her first mental powers. But she did not become intelligent and self-conscious straight away. She was merely conscious first. Because she was not then a person, person essentialism rules out her being you. Yet you were psychologically continuous, when you were first a person, with the foetus (or infant) as she was just previously: again, you inherited some of her memories, preferences, and other mental states. Almost any psychological-continuity view will imply that the foetus (or infant) is you.

Call this the *psychological-continuity problem*. It would no doubt be possible to devise a psychological-continuity view that avoids it by ruling out our being the conscious nonpeople we grew from or that would result from our developing severe dementia. Our persistence might consist in a sort of psychological continuity that holds throughout the time when we are people, but not between a person and the conscious foetus she grew from or the mentally impaired nursing-home resident that would result from her lapsing into stage-3 dementia. For example, we might say that a being existing at another time is you just if it is intelligent and self-conscious then and its having these properties in particular is causally dependent in the right way on your



having them now or vice versa. Or we might say that our persistence consists only in the continuity of those mental properties that require self-consciousness. Continuity of mental properties not required for personhood—the sort we find in fetuses, dementia sufferers, and dogs—would be irrelevant to our persistence.

That would be compatible with Lockean person essentialism. But only a tough-minded metaphysician could seriously believe it. Imagine that your friend Sally gets severe senile dementia. At the institution where she is confined, you encounter a doddering resident of familiar appearance. You may well say ‘That’s not the Sally I knew’. But then you might say the same if she were definitely still a person but dramatically altered—following a stroke, say, or heavily sedated. These are not judgments about numerical identity. You might just as easily say ‘She’s not the person she used to be’, which presupposes that she existed both before and after getting dementia. The current proposal implies that the nursing home resident is a being you have never seen before, and that your friend now has the same status she would have if she had died and been cremated.

In any event, we are not aware of any actual proposal of this sort. Some attribute such a view to Parfit—specifically, that our persistence consists in a degree of psychological continuity that holds only while we are people (McMahan 2002, p. 45). This is a mistake. Parfit did once say (1984, p. 206) that you persist just if you have at least half as many direct psychological connections—a memory caused by an earlier experience, for instance—over a day as there actually are during a day in the life of nearly every person (and there is no branching). But if there are enough such connections during your first day as a person, there will almost certainly be enough during the 24-hour period beginning ten seconds earlier. It would be an extraordinary coincidence if the required number of connections first appeared in even one person at precisely the moment she became a person in the Lockean sense. The two events are completely independent: the definition of personhood implies nothing about numbers of direct psychological connections. To say that they must always occur together would be like saying that all people necessarily acquire the physiological capacity to reproduce at precisely the moment they reach the legal age of sexual consent.

This reveals a fundamental incoherence in orthodox thinking about personal identity (including Parfit’s influential account). The dominant view for the past century has been that some sort of psychological continuity is both necessary and sufficient for us to persist. (Debates

about our persistence are very often about which sort it is.) And as we have seen, most of this discussion presupposes Lockean person essentialism. Yet the two views are incompatible. We are all psychologically continuous—in the right way and with no branching—with beings as they are at times when they are not people in the Lockean sense. It follows that either we were once nonpeople, contrary to person essentialism, or such psychological continuity is insufficient for us to persist, contrary to psychological-continuity views.

#### 4. The thinking-animal problem

A different problem for Lockean person essentialism has to do with the mental powers of human beings that do not count as people in the Lockean sense. One version of it applies to human organisms, another to foetuses and those with dementia. Start with organisms.

Person essentialism implies that no human organism is ever a person. That's because no human organism is a person essentially. Human organisms are not intelligent and self-conscious throughout their existence. They do not come into the world as Lockean people, but as embryos entirely devoid of mental powers. And even if one did somehow come into being as a mature adult—if a capricious deity were to create a duplicate of you *ex nihilo*, say—it would still be *possible* for it to exist without being a person, in a state of severe dementia or the like. So if whatever is a person *must* be a person throughout its existence, no human organism could ever be a person. Given that to be a person is to be intelligent and self-conscious, that makes it absolutely impossible for a human organism to have these mental properties. And if no *human* organism could be intelligent and self-conscious, it's hard to see how any other biological organism could be.

This is surprising. Why should it be impossible for an organism to have these properties? You are intelligent and self-conscious. The organism you see in the mirror has the same brain that you have, the same surroundings, the same education, and the same behaviour in both actual and counterfactual situations. What could prevent it from using that brain to think in the way that you do?

This is a version of the so-called 'thinking-animal problem' brought against psychological-continuity views of personal identity over time (Snowdon 1990, pp. 89f.; Olson 1997, pp. 102-8). Those views say that necessarily all people persist by virtue of psychological continuity. Because no organism, human or otherwise, persists by virtue of psychological continuity (van Inwagen 1990, pp. 182-8; Olson 2015, pp.

102-5) it follows that none could be a person. Given a Lockean account of personhood, this rules out any organism's being intelligent and self-conscious—again leaving us wondering why.

Three answers to this question have been proposed.<sup>4</sup> All three say that organisms cannot have the mental powers sufficient for personhood because they cannot have mental properties at all. Psychology and biology—*mind* and *life*—are metaphysically incompatible.

The first says that organisms cannot have mental properties because they are material things. Thinking or conscious beings, ourselves included, must be immaterial (Foster 1991, pp. 206-12; Hasker 2010, pp. 181-3).

The second says that organisms cannot have mental properties because of their parts. It's not organisms that think, but rather their brains or central nervous systems. An organism thinks only in the sense of having a thinking part, much as a car is powerful by having a powerful engine. Having hands and feet as parts makes it impossible for a thing to have mental properties in anything but this loose and superficial sense (Hudson 2007, pp. 218f.; Parfit 2012, p. 17).

The third answer is that organisms cannot have mental properties because of their persistence conditions. It belongs to the nature of mental properties that nothing can have them unless psychological continuity is sufficient for it to persist. But psychological continuity is not sufficient for an organism to persist. If your brain were transplanted into another head, the recipient may end up psychologically continuous with you. Yet this would not move an organism to another head: it would only move an organ from one organism to another, just as a liver transplant does. *Two* organisms can be psychologically continuous. The claim is that precisely this fact prevents them from having mental properties (Shoemaker 2011).

## 5. The conscious-foetus problem

If one of these three proposals were right, the implication of Lockean person essentialism that no organism could be intelligent and self-conscious would not be especially troubling: we should know already on other grounds that no organism could have *any* mental property. But a further difficulty arises in the case of human beings whose mental properties fall short of personhood.

<sup>4</sup> Or maybe four: see our discussion of Baker in §6.

Suppose you one day lose your power of self-consciousness owing to senile dementia and, being a person essentially, thereby cease to exist. Call the being this leaves behind *D*. *D* presumably exists today, while you are healthy: your later demise could hardly bring her into existence. But according to person essentialism *D* is never a person, as she is not one essentially. Not being a person, she is never intelligent and self-conscious—not merely when suffering from dementia, but even now, in her prime. Why not? Again, you are intelligent and self-conscious, and *D* now has the same brain that you have, the same surroundings, the same education, and the same behaviour. What could prevent her from using that brain to think as you do?

The proposed solutions to the thinking-animal problem are no help here. They purport to explain why organisms cannot have the mental properties required for personhood by denying that organisms can have mental properties at all (because they're material things or have the wrong parts or persistence conditions). But *D* *has* got mental properties. She can think to an extent, and is at least minimally conscious during her waking hours. She can suffer: that's why those with dementia deserve humane treatment. Lockean person essentialism implies that beings with mental properties falling short of personhood in senile dementia cannot be intelligent and self-conscious when healthy, despite then being physically and behaviourally identical to normal adult people. There must be some explanation for this remarkable fact. But it cannot be the metaphysical incompatibility of mind and life.

Our beginnings create a similar puzzle. The foetus you grew from—call her *F*—was once conscious but not self-conscious. At that point she was not a person. Yet she would seem to exist now: the emergence of intelligence and self-consciousness in the normal course of her development could hardly have destroyed her. It follows from Lockean person essentialism that she never became a person, as it is metaphysically impossible for anything that is not self-conscious at some time to acquire the mental properties sufficient for personhood, even when healthy and mature. *F* can never become self-conscious. But why not? She now has the same brain that you have, and the same surroundings, education, and behaviour. What prevents her from using that brain to think as you do? Once more, the answer cannot be that she can have no mental properties at all. She *has* got them—or at least she did when she was a foetus.<sup>5</sup>

<sup>5</sup> David Hershenov (2006, pp. 233f.) gives a brief and insightful discussion of these problems, but does not connect them with person essentialism.

How the two problems are related—the one about dementia and the one about foetuses—depends on the relation between D and F. It may be that the conscious nonperson that would result from your getting stage-3 dementia is the very conscious foetus you grew from, combining the two problems into one. Otherwise there will be two conscious nonpeople and thus two problems. We will assume for the sake of simplicity that D is F.

We can summarize the unified problem like this: the foetus that gave rise to you, and which might one day get severe dementia, has mental properties for a time, but not those sufficient for personhood. It is not then a person. Yet it would appear to exist now, with a mature and healthy brain. And because it is capable of having mental properties, we should expect it to have the same mental powers now that you have, making it now a person. In that case a nonperson could become, and later cease to be, a person, contrary to person essentialism. Call this the *conscious-foetus problem*.

## 6. Solutions

The problem arises from two claims. First, something can be a foetus (and not a person) at one time and a healthy adult at another. It is possible, in other words, for a human foetus to survive birth and childhood in normal conditions. The other is that a being that was once a conscious foetus can have the mental properties sufficient for personhood in its adult state. (We omit the analogous claims to do with dementia sufferers.) Together they entail that something can be a person only temporarily. Person essentialists must reject one of them.

To reject the first is to deny that a foetus can become a healthy adult. The foetus you grew from no longer exists. Despite appearances, it perished the moment it was about to become intelligent and self-conscious, and was instantly replaced by a person—you—who did not exist before. It is metaphysically impossible for *any* conscious being to survive the developmental stage that creates the properties sufficient for personhood. Nor can any being survive the sort of cognitive decline that destroys these properties, even if its other mental powers remain intact and there is a good degree of psychological continuity. It must cease to exist and be replaced by a new conscious being that did not exist before. The normal development of intelligence and self-consciousness is necessarily fatal, and senile dementia generates a new conscious being.

We might call this a ‘metaphysical’ solution to the conscious-foetus problem, as it has to do with existence and identity over time. It says that a foetal nonperson cannot be a person in its prime because it cannot even *exist* in its prime. It can survive birth and childhood only if, owing to some severe neural defect, it never develops intelligence and self-consciousness. We normally call such events ‘tragic’ because we take them to deprive a being of the sort of mental life the rest of us enjoy. But on this proposal they could not deprive any being of anything: not even a perfectly healthy foetus could ever go on to acquire intelligence and self-consciousness.

Views of this sort have occasionally been proposed (Burke 1994; McMahan 2002, pp. 46f.; Sharpe 2015, pp. 66f.), but they are hard to take seriously. They go against nearly everything we thought we knew about the persistence of animate beings, and their accounts of what it does consist in are obscure (Olson 1996, pp. 389-91; Sider 2001, pp. 161-5; Witt 2020, §10).

Alternatively, person essentialists can accept that the foetus you grew from exists now, as an adult, but deny that it can ever have the mental powers sufficient for personhood. It may have *some* mental powers now—those it had before you came into being, perhaps—but it can never become intelligent and self-conscious. So no nonperson can become a person, just as person essentialism demands.

We might call this a ‘psychological’ solution, as it has to do with mental powers. It has important advantages over the metaphysical solution. Suppose you one day get severe dementia. The metaphysical solution appears to imply that the foetus you grew from, having been destroyed by the appearance of intelligence and self-consciousness, is distinct from the dementia sufferer. The psychological solution allows them to be the same, as they appear to be. It would also solve the thinking-animal problem: whatever prevents an ex-foetus or potential dementia sufferer from being intelligent and self-conscious in its adult prime could also prevent a human organism from having those properties. (Presumably the foetus you grew from *is* a human organism.)

But the psychological solution is no less mysterious than its metaphysical counterpart. It’s hard to see what could prevent a healthy, mature ex-foetus or human organism from being intelligent and self-conscious, given that it can have other mental properties (as well as being physically and behaviourally identical to a person).

The best-known response to this question is Lynne Rudder Baker’s. She claims that human animals can be conscious and modestly

intelligent, but are necessarily incapable of self-consciousness—that is, of thought requiring what she calls a ‘first-person perspective’.<sup>6</sup> This is because ‘the appearance of a first-person perspective makes an ontological difference in the universe’—that is, a difference in what there is (Baker 2000, p. 163; see also pp. 17, 218–20). If, *per impossibile*, a foetus or organism could first lack a first-person perspective and then have one, the advent of this perspective would not create a new thing, but merely alter an old one, like someone’s becoming a lumberjack. It would make no difference to what there is. Likewise, if you could survive the loss of your first-person perspective, nothing would thereby cease to exist. Given that having such a perspective is what it is to be a person, you could become a nonperson. It would merely change you. Baker takes this to imply that ‘persons as such have no ontological significance’ (Baker 2000, p. 220), which she finds manifestly absurd.

Though this may sound like an argument for Lockean person essentialism in addition to those mentioned in §2, it isn’t really. The crucial claim is that if we were people contingently, people as such ‘would have no ontological significance’. No one can hear that without feeling repelled. How could you and I fail to have ontological significance? But ‘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.

The important thing, however, is not Baker’s argument for person essentialism, but her account of why a healthy, mature human organism or former foetus could never be intelligent and self-conscious, despite having other mental properties such as consciousness. Her answer appears to be that this is because human organisms are not intelligent and self-conscious essentially. They might otherwise be ideally suited to have these properties, having just the right brain, surroundings, education, and behaviour. They are not prevented from being intelligent and self-conscious by their materiality, their

<sup>6</sup> Baker concedes that an organism can be self-conscious in the derivative sense of ‘constituting’ another being that is self-conscious nonderivatively, but not in the straightforward sense that you and I are (2000, pp. 197f.).

parts, or their persistence conditions. What prevents it is the mere fact that they could exist without having those properties.

But although the claim that Fs are essentially F entails that things not essentially F cannot be F at all, it does not by itself explain *why* they can't be. If intelligence and self-consciousness are essential to the beings that have them, then human organisms and former foetuses, which are not essentially intelligent and self-conscious, cannot be so even contingently. But that is no explanation of their surprising mental incapacity. To take a similar case, suppose an infallible source told us that whatever has mass has it essentially. Given that photons are not essentially massive, it follows that it is impossible for them to have mass. But again, this would not explain why not: it would not tell us what it is about being a quantum of light that rules out having mass. Likewise, Baker's essentialist claim, even if it's true, does not tell us what it is about being a human organism that rules out intelligence and self-consciousness.

It may be that the fact simply has no explanation. Maybe there is no point in asking why a merely conscious being cannot become self-conscious, as there are no other, 'underlying' facts in terms of which it could be explained. But not even Baker has ever made this unlikely claim.

## 7. Alternatives

The problem, then, is that the foetus you grew from would appear to exist now, and to have the same mental properties now that you have, making it a person but only contingently. Person essentialists must either deny that that being now exists—the metaphysical solution—or deny that it now has your mental properties—the psychological solution.

What if we don't like either of these proposals? Suppose it's possible for a foetal nonperson to become a mature human adult, and in this condition to have the mental powers sufficient for personhood. Then something can be a person contingently and temporarily, contrary to person essentialism in both its strong and weak forms. In particular, it must be possible for *us*. And not merely possible: we actually begin as nonpeople. If *any* human foetus can become a person, those that we grew from have actually done so. And in that case these beings are none other than you and me. (No one would say that the foetus's becoming intelligent and self-conscious made it a *second* person in



addition to you.) In fact, each of us must exist as soon as the foetus we grew from has any mental properties at all, if not earlier. (Otherwise, again, it must have either ceased to exist or survived but never become intelligent and self-conscious.)

And it would not only be the property *being a person* that we must have only accidentally, but also the property *being a person at some time*. Each of us could have existed without ever being a person. If you were once a foetus that was not yet a person, you could have died before becoming one and thus remained a nonperson throughout your life.

Suppose all this is right: we start out as embryos lacking the mental powers necessary for personhood; we can survive without those powers in a state of severe dementia; and we might have existed without ever having them. They are only contingent and temporary properties of us. Would anything remain of the widespread view that we are in some way fundamentally psychological beings—that psychology is woven into our deepest metaphysical nature?

It may seem not. It may seem to follow, in fact, that we are biological organisms ('animalism'). We should persist by virtue of some sort of brute-physical continuity. We should differ from dogs and trees by having certain mental properties for part of our lives, but even if this is our normal condition, it would be no more essential to us as individuals than our having two legs. The natural alternative to person essentialism is a combination of animalism and a brute-physical account of our persistence. That would avoid all the problems we have discussed.

But there is another possibility. We might be fundamentally psychological beings in a weaker sense. Even if no particular mental powers such as intelligence or self-consciousness are essential to us, *having some mental powers or other* might be. Presumably this would hold for all beings with mental powers: whatever has mental powers has mental powers essentially, and could not exist without any. We might call this *psychological essentialism*. It follows from person essentialism but does not entail it, and comes in analogous weak and strong forms.

On this view, our persistence might still consist in some sort of psychological continuity. We could even accept a standard psychological-continuity view of personal identity: that we persist by virtue of continuity of mental properties generally, not only those required for personhood. (This would entail psychological essentialism: if you could not persist without psychological continuity, you could not exist without mental powers.) This would be a view

intermediate between the two familiar ‘extremes’: Lockean person essentialism combined with a standard psychological-continuity view, which we have shown to be inconsistent, and animalism combined with a brute-physical account of our persistence.<sup>7</sup>

Whatever its appeal, however, this intermediate view has a number of drawbacks. Most obviously, it implies that no biological organism, or at least none that we know of, could have mental powers: not even the capacity for basic consciousness. It says that whatever has mental powers has them essentially. But all organisms we know of can exist without the sophisticated nervous system required for mental powers, and actually do in their embryonic state. To explain what prevents organisms from having such powers, we should have to return to the proposals of §4, to do with their materiality, their parts, or their persistence conditions.

It also implies that acquiring the properties that make something a person creates no new entity, and losing them destroys none: people as such ‘have no ontological significance’.

Further, it rules out any account of personal identity as such—that is, any persistence conditions that necessarily apply to all and only people in the Lockean sense. Any account of personal identity compatible with the intermediate view will apply to foetuses and dementia sufferers that are not people. That’s because the intermediate view says that such an account is true of us, and that we were once such foetuses and could become severely demented. In fact, we should expect it to apply to all beings with mental powers, dogs included.<sup>8</sup> Some deny that such a view is ‘even a contender as an account of *personal* identity over time’ (Baker 2000, p. 124, emphasis in original); see also Gert 1971, pp. 475f.; and Johnston 2010, p. 256) and find this unacceptable.

And it is incompatible with Parfit’s influential account of personal identity briefly mentioned in §3: that a person persists for a day only if she exhibits a certain number of direct psychological connections.<sup>9</sup> When a human foetus acquires its first mental powers there are very

<sup>7</sup> Unger (2000) and McMahan (2002, pp. 66-9) have made proposals of this sort, though not for the reasons we discuss here.

<sup>8</sup> Or more precisely, beings that stand to canine organisms as we stand to human organisms.

<sup>9</sup> 1984, p. 206. Parfit says only that such direct psychological connections are required for someone to be the *same person* as you. But as we argued in §2, he does not take this to be required for you to persist only while you are a person, or to allow you to survive without it as a nonperson.

few such connections—certainly fewer than Parfit’s account requires. His account rules out the intermediate view’s claim that we were once such foetuses.

Though the intermediate view may be worth exploring, we doubt whether it will find many friends. What are the alternatives? One is to stick with person essentialism and try to solve the problems we have identified. That will require a novel account of our persistence compatible with person essentialism along the lines sketched in §3. And it means developing either the ‘metaphysical’ or the ‘psychological’ solution of §6. The first requires an explanation of why no human foetus can survive normal development into a healthy child even though it is physically and psychologically continuous with one. The second requires an account of why a foetus that survived this could never become intelligent and self-conscious, despite having had mental properties previously. Any of these will be a large and daunting research project.

The other alternative is to give up person essentialism and accept a view that avoids the problems, the most natural being that we are biological organisms. There is no easier way out.<sup>10</sup>

## References

- Baker, Lynne Rudder 2000, *Persons and Bodies* (Cambridge: Cambridge University Press)
- Blackburn, Simon 2005, ‘Personal Identity’, in his *The Oxford Dictionary of Philosophy*, 2nd ed. (Oxford: Oxford University Press)
- Buchanan, Allen E., Dan W. Brock 1990, *Deciding for Others: The Ethics of Surrogate Decision Making* (Cambridge: Cambridge University Press)
- Bunnin, Nicholas 2004, *The Blackwell Dictionary of Western Philosophy* (Malden, MA: Blackwell)
- Burke, Michael 1994, ‘Dion and Theon: An Essentialist Solution to an Ancient Puzzle’, in *Journal of Philosophy* 91
- DeGrazia, David 2005, *Human Identity and Bioethics* (Cambridge: Cambridge University Press)

<sup>10</sup> We are grateful to Eli Pitcovski, Radivoj Stupar, three anonymous reviewers, and the Editors of *Mind* for generous comments on earlier versions. We also thank the *Deutsche Forschungsgemeinschaft* for supporting Karsten Witt’s work on the paper.

- Flew, Antony, and Stephen Priest (eds.) 2002, *A Dictionary of Philosophy*, 3rd ed. (London: Macmillan)
- Foster, John 1991, *The Immaterial Self* (London: Routledge)
- Garrett, Brian 1998, 'Personal Identity', in Edward Craig (ed.), *Routledge Encyclopedia of Philosophy* (London: Routledge)
- Gert, Bernard 1971, 'Personal Identity and the Body', in *Dialogue* 10
- Glover, Jonathan 1988, *I: The Philosophy and Psychology of Personal Identity*, (London: Penguin)
- Hasker, William 2010, 'Persons and the Unity of Consciousness', in Robert C. Koons and George Bealer (eds.), *The Waning of Materialism* (Oxford: Oxford University Press)
- Hershenov, David 2006, 'Shoemaker's Problem of Too Many Thinkers', in *Proceedings of the American Catholic Philosophical Association* 20
- Hudson, Hud 2007, 'I Am Not an Animal!', in Peter van Inwagen and Dean Zimmerman (eds.), *Persons: Human and Divine* (Oxford: Oxford University Press)
- Johnston, Mark 2010, *Surviving Death* (Princeton: Princeton University Press)
- Kuhse, Helga 1999, 'Some Reflections of the Problem of Advance Directives, Personhood, and Personal Identity', in *Kennedy Institute of Ethics Journal* 9
- Lewis, David 1976, 'Survival and Identity', in Amélie Oksenberg Rorty (ed.), *The Identities of Persons* (Berkeley: University of California Press)
- Locke, John 1975, *An Essay Concerning Human Understanding* (Oxford: Oxford University Press) (2nd ed. first published 1694)
- Lowe, E. J. 2012, 'The Probable Simplicity of Personal Identity', in Georg Gasser and Matthias Stefan (eds.), *Personal Identity: Complex or Simple?* (Cambridge: Cambridge University Press)
- Mackie, David 1999, 'Personal Identity and Dead People', in *Philosophical Studies* 95
- Mautner, Thomas (ed.) 2005, *Dictionary of Philosophy* (London: Penguin)
- McMahan, Jeff 2002, *The Ethics of Killing* (Oxford: Oxford University Press)
- Noonan, Harold 1998, 'Animalism Versus Lockeanism: A Current Controversy', in *The Philosophical Quarterly* 48
- 2003, *Personal Identity*, 2nd ed. (London: Routledge)
- Nozick, Robert 1981, *Philosophical Explanations* (Cambridge MA: Harvard University Press)

- Oderberg, David 2005, 'Hylemorphic Dualism', in *Social Philosophy and Policy* 2
- Olson, Eric 1996, 'Composition and Coincidence', in *Pacific Philosophical Quarterly* 77,
- 1997, *The Human Animal* (New York: Oxford University Press)
- 2015, 'What Does it Mean to Say that We are Animals?', in *Journal of Consciousness Studies* 22, 11–12
- Parfit, Derek 1984, *Reasons and Persons* (Oxford: Oxford University Press)
- 2012, 'We Are Not Human Beings', in *Philosophy* 87
- Penelhum, Terence 1967, 'Personal Identity', in Paul Edwards (ed.), *The Encyclopedia of Philosophy* (New York: Macmillan)
- Reas, Emilie 2017, 'Can Alzheimer's Disease Steal One's Consciousness?', in *PLOS Neuro Community* <<https://blogs.plos.org/neuro/2017/04/03/can-alzheimers-disease-steal-ones-consciousness/>> (accessed on 10 Jan. 2020)
- Reisberg, Barry, Steven H. Ferris, Mony J. De Leon, and Thomas Crook 1982, 'The Global Deterioration Scale for Assessment of Primary Degenerative Dementia', in *American Journal of Psychiatry* 139
- Rosenberg, Jay 1998, *Thinking Clearly about Death* (Indianapolis: Hackett)
- Sharpe, Kevin W. 2015, 'Animalism and Person Essentialism', in *Metaphysica* 16
- Schechtman, Marya 2010, 'Personhood and the Practical', in *Theoretical Medicine and Bioethics* 31
- 2014, *Staying Alive: Personal Identity, Practical Concerns, and the Unity of a Life* (Oxford: Oxford University Press)
- Shoemaker, David 2019, 'Personal Identity and Ethics', in Edward Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), <<https://plato.stanford.edu/archives/win2019/entries/identity-ethics/>>
- Shoemaker, Sidney 1984, 'Personal Identity: A Materialist's Account', in Shoemaker S. and Swinburne R. (ed.), *Personal Identity* Oxford: Blackwell)
- 2011, 'On What We Are', in Shaun Gallagher (ed.), *The Oxford Handbook of the Self* (Oxford: Oxford University Press)
- 2015, 'Personal Identity', in Robert Audi (ed.), *The Cambridge Dictionary of Philosophy*, 3rd ed. (Cambridge: Cambridge University Press)

- Sider, Theodore 2001, *Four-Dimensionalism* (Oxford: Oxford University Press)
- Snowdon, Paul 1990, 'Persons, Animals, and Ourselves', in Christopher Gill (ed.), *The Person and the Human Mind* (Oxford: Oxford University Press)
- 2005, 'Personal Identity', in Ted Honderich (ed.), *The Oxford Companion to Philosophy*, 2nd ed. (Oxford: Oxford University Press)
- Strawson, Galen 2009, *Selves: An Essay in Revisionary Metaphysics* (Oxford: Oxford University Press)
- 2011, *Locke on Personal Identity* (Princeton: Princeton University Press)
- Strawson, Peter F. 1995, 'Personal Identity', in Ted Honderich (ed.), *The Oxford Companion to Philosophy* (Oxford: Oxford University Press)
- Unger, Peter 2000, 'The Survival of the Sentient', in *Philosophical Perspectives* 14
- van Inwagen, Peter 1990, *Material Beings* (Ithaca: Cornell University Press)
- Wiggins, David 1980, *Sameness and Substance* (Oxford: Blackwell)
- Witt, Karsten 2020, 'In Defence of Advance Directives in Dementia', in *Pacific Philosophical Quarterly* 101
- WHO 2019, Dementia: Key Facts <<https://www.who.int/news-room/fact-sheets/detail/dementia>> (accessed on 10 Jan. 2020)

