dialectica

Vol 78 : 1
March 2024

The Plurality of the Self

Harry Cleeveley

doi:10.48106/dial.v78.i1.03

Harry Cleeveley. 2024. "The Plurality of the Self."
Dialectica 78(1): 1–1. doi:10.48106/dial.v78.i1.03.

EUROPEAN SOCIETY FOR ANALYTIC PHILOSOPHY

The Plurality of the Self

Harry Cleeveley

There are good reasons to think that I am a Lockean person, and there are good reasons to think that I am an animal. But there is a problem: an animal cannot be identical to a Lockean person, as the two types of entity have different persistence conditions. And so the standard response of Lockeans has been to reject the claim that I am, strictly speaking, an animal. Similarly, animalists, while they can accept that we are persons in a broader sense, deny that we are Lockean persons. I find this unsatisfactory. Just as I think it is true of me that I would survive the teleporter, so I think it is true of me that I would continue to exist in a persistent vegetative state. But how can this be so? The answer, I argue, is that there are two thinkers at my location, a Lockean person and a human animal, and that my use of the first-person pronoun is ambiguous between these two entities. Thus there is one sense in which I am a Lockean person, and would survive teleportation, and another sense in which I am an animal. I also show how my proposal avoids the problem of there being too many thinkers at my location.

There are many problems of personal identity, but two are perhaps the most important. The first is the question of our fundamental nature: what sort of things are we? The second is the question of our persistence through time: what makes a person at one time the *same* person as a person at another time? Although these are separate questions, the answer we give to one will tend to influence the answer we give to the other.

Probably the two most influential theories of personal identity are Lockeanism and animalism. Lockeanism, named for the theory of personal identity put forward by John Locke in *An Essay Concerning Human Understanding* (1689), is the claim that we are particular kinds of objects with psychological persistence conditions, namely persons.[1] To put it in terms of many neo-Lockean thought experiments of recent times, my personal identity would go

1 Views differ on exactly what a person is and precisely what the relevant psychological conditions are for its persistence. For Baker (2000), a person is an entity that essentially has a first-person perspective on the world. This incorporates the idea that, to be a person, an object must have

with my transplanted brain. Versions of this view have been put forward by Shoemaker (2004, 2008) and Parfit (1971, 2012), amongst others. Animalism, on the other hand, is just the claim that we are animals. This view has been advocated by Olson (2003, 2007) and Snowdon (2003), amongst others.

I start from the perspective of a neo-Lockean. I agree with the thought experiments that say that I would go with my transplanted brain, and I think that it is at least metaphysically possible that I could persist in a non-physical afterlife. But I also think that we are animals. I feel the force of Olson's (2007) point: when I look in the mirror, an animal looks back at me. Surely, then, I am the animal – specifically, a primate of the species *homo sapiens* – looking back?

But there is a problem, which is that an animal cannot be identical to a Lockean person. The two types of entity have different persistence conditions: a Lockean person can (in principle, as a matter of metaphysical if not natural possibility) persist beyond the destruction of its associated animal; but an animal cannot survive its own destruction. Therefore a Lockean person and its associated animal will have different – in fact, incompatible – modal properties; and therefore, whatever their metaphysical relationship may be, it cannot be one of numerical identity.

Thus a common response that Lockeans make to animalism is to deny that I am, strictly speaking, an animal. If I am an animal at all, then it is only in the derivative sense of being, for example, materially constituted by an animal, or inhabiting an animal, or some other such formulation; and if I have animal properties, then I do so only in a derivative sense, in virtue of being closely metaphysically related to – but not identical to – the animal that directly bears those properties.

I find this response unsatisfactory. Whilst I do not claim to *prove* that I am both an animal and a Lockean person, I do think that there are good reasons to accept both claims, and that they can be made compatible. So I want to carve

phenomenal consciousness – that, following Nagel (1974), there is something it is like to be a person (at least some of the time) – but it is a stronger requirement. On this view, for an entity to be a person, it is not enough that there should be something it is like to be that thing: its consciousness also needs to be organised into a coherent, unified viewpoint on the world. The conditions for the survival of the person are thus the conditions for the persistence of this unified, conscious first-person perspective. Broadly, this notion of a person and its survival conditions is the notion that I have in mind in this paper. (Baker also suggests that, to count as a person, an entity with a first-person perspective must also be self-aware and, moreover, aware of its own mental states *as* its mental states. However, this further requirement seems to me to be too strong and too exclusive. But I digress.)

out a philosophical space in which it is literally true both that I would go with my transplanted brain, and that I existed as a pre-conscious foetus. If I am successful, this will allow Lockeans to accept that we are literally and strictly speaking animals, and it will remove a potential objection to Lockeanism, namely that it does not allow us to make such a claim.

But how is this possible? The answer cannot be that I am one thing that is both Lockean person and animal – since, as we have seen, a Lockean person cannot be identical to an animal. So my proposed solution is that there are two thinking entities at my location, a Lockean person and a human animal, and that my use of the pronoun 'I' is ambiguous in its reference between these two entities. Informally, we might say that *I am not one thing, but two*. We must be careful, though, not to take this slogan literally, since there is no disambiguation of 'I' on which it is true of me that I am two things: rather, there is a disambiguation on which I am an animal, and a disambiguation on which I am a Lockean person.

Perhaps the main objection to this position is that it (supposedly) means that there are at least two thinkers, with two minds, at my location, and that this would give rise to spurious epistemological worries about which one of them is really me. This problem – the too-many-thinkers problem – is at the heart of Olson's (2003) *thinking animal* argument for animalism. I will show how my theory can meet this objection. My view does indeed entail a sense in which there are (at least) two *thinkers* at my location. But, I argue, this is not in itself a problem. The important thing is that on my view there are not two *minds* at my location. There is only one mind, or one stream of consciousness – but there are two overlapping entities whose mind it is, and which, in their own different ways, are thinking my thoughts. I also develop an account of first-person reference to support this view.

A note on the scope of my claims about personal identity: they are not intended to apply to all possible persons (or even to all actual persons), but to *human* persons – and, by extension, to the persons of non-human animals, if there are any. Primarily, my thesis is a claim about myself: that my use of the first-person pronoun can refer to two things, namely a Lockean person and an animal. By extension, it is also a theory about all those persons who are of the same nature as me.

doi: 10.48106/dial.v78.i1.03

1 Can a Lockean Person Be Identical to an Animal?

I start from the perspective that some form of Lockeanism is true. I will not argue for this claim, since others, such as Parfit (1971, 2012), and Shoemaker (2008), have already done so, and I am broadly in agreement with their views. I agree that I have psychological persistence conditions – that I would go with my transplanted brain, and that I could survive teleportation. Crucially, I also think that it is metaphysically possible for me to survive the destruction of my physical animal – that I could in principle survive, for example, as a computer simulation or in a non-physical afterlife – as long as the relevant psychological continuity persists.[2]

Of course, it is an empirical question whether my transplanted brain, or any part thereof, would in fact take with it my psychological continuity. The point is just that, if it would, then I would go with it. Similarly for teleportation, and for the more exotic visions of my post-animal survival.

A point to note: my persistence through time requires that some future entity should in the relevant sense be me – that is, my present self. However, in this paper, I will take no view on whether sameness of person across time is a matter of strict numerical identity or of some other metaphysical relationship, such as being temporal parts of the same four-dimensional object. For the record, I am inclined against numerical identity and towards a temporal parts view – but nothing that follows depends on the answer to this question.

My aim in this paper is to show how, starting from a Lockean perspective, we can – indeed should – accept that we are animals. But what does it mean to say that we are animals? As Olson (2015b) argues, animalism is not merely the claim that I am an animal in some derivative or indirect sense, such as by being embodied in an animal or constituted by an animal. For in neither of

2 To be clear, neo-Lockeans disagree amongst themselves regarding which scenarios I would survive in, and which I would not. For example, Shoemaker (2004) thinks that I could survive brain transplantation but not the teleporter, because in the latter case there would not be the right sort of causal continuity for the Lockean person to persist. Nor, of course, do Lockeans have to accept that it is metaphysically possible for me to survive in a non-physical form.

Conversely, one could think that I could survive brain transplantation without being a Lockean – for example, if one holds that I am just identical to my living brain. Similarly, one could think that I could survive in a non-physical afterlife without being a Lockean, for example if one holds that I am an immaterial soul. So there is no agreed upon set of circumstances where all the Lockeans – and only the Lockeans – think I would survive. My point here is just that I think I could survive the brain transplant and the teleporter, and could in principle continue to exist in a non-physical form, precisely because I think that in all these cases the psychological continuity that is necessary and sufficient for personal persistence could be realised.

these scenarios would it be strictly true to say of me that I am an animal. If I am just a thing that is embodied in an animal (such as a Cartesian immaterial substance), then I am not really an animal at all; and if I am just a thing that is constituted by an animal, then I must be something other than the animal that constitutes me, for a thing is not identical to the thing that constitutes it.[3] So animalism is not the claim that we are animals in some derivative or indirect sense. Rather, as Bailey (2015) argues, it is the claim that each of us is *numerically identical* to an animal. In my case, there exists a human animal, and my use of the first-person pronoun refers to that animal: I am identical to that thing, in the classical sense that all its properties are my properties, and all my properties are properties of that human animal. So, the question is this: given the starting point that I am a Lockean person, is it possible that I am identical to an animal?

Now, there is no reason why an animal cannot be a person in the broader sense of having a unified first-person consciousness. That is because the concept of a person in this broader sense can plausibly be understood as a *phase sortal*. A phase sortal concept refers to a phase through which an object may pass, and which can cease to apply to the object, without the object in question ceasing to exist. Snowdon (2003) draws a distinction between phase sortal concepts and what he terms an *abiding sort*, where an abiding sort is a category such that something of that sort cannot cease to be a member of the sort, without ceasing to exist. For example, the concept of a philosopher is a phase sortal, rather than an abiding sort: a thing can pass through a phase of being a philosopher, then cease to be one, without ceasing to exist. To be a philosopher, one does not have to have philosophy-related persistence conditions. But the concept of a lump of rock, on the other hand, is plausibly an abiding sort: anything that is a lump of rock cannot cease to be a lump of rock and yet continue to exist. Now, there is no reason to think that to be a person in this broader sense – as opposed to being a *Lockean person* – an

3 There is room for debate here, but I take Olson to have in mind something like the classical view of material constitution, as in Wasserman (2021), which I will also adopt in this context. On this view, a statue is materially constituted by the lump of clay from which it is formed, but the relationship is asymmetrical – the statue does not constitute the lump of clay – and therefore constitution is not the same as identity. Moreover, the existence of that particular lump, shaped in that particular way, is metaphysically sufficient for the existence of that particular statue, but it is not metaphysically necessary, since the same statue could have been formed from a different lump.

doi: 10.48106/dial.v78.i1.03

entity must be a person essentially.[4] Thus it is plausible that being a person in the broader sense is a phase sortal: a thing, such as a human animal, can instantiate the properties of personhood for a period of time, and then cease to do so – thus ceasing to be a person – without ceasing to exist. Indeed, this is the account that most animalists give of our personhood.

However, being a *Lockean* person – as opposed to merely a person in the broader sense – cannot be a mere phase of an animal. It is tempting to think that this is because Lockeanism is a theory of our persistence conditions, and that there is a general principle that a thing cannot change its persistence conditions and yet continue to exist. However, I do not think we are entitled to assume any such general principle. A thing cannot cease to *satisfy* its persistence conditions whilst continuing to exist; but this does not mean a thing's persistence conditions themselves cannot change over time, while the thing persists.

For example, consider the following case. A person P_1 at time t_1 has persistence conditions C_1, where C_1 is the appropriate type and degree of psychological continuity with P_1. Thus some later person, P_2, at time t_2, will be the same person as P_1 just if they satisfy C_1. And now the persistence conditions of P_2 will be C_2, where C_2 is the appropriate type and degree of psychological continuity with P_2; and some later person P_3 at t_3 will be the same person as P_2 just if they satisfy C_2; and so on. So the question is whether the sequence of successive persistence conditions, C_1, C_2...C_n, must be unchanging?

If sameness of person across time is transitive, then any C_n must include C_1, in the sense that satisfaction of C_n must entail satisfaction of C_1. This is because if sameness of person is transitive, then, if P_1 is the same person as P_2 and P_2 is the same person as P_3, then P_3 is the same person as P_1 – and thus the fact that P_3 satisfies C_2 means that they also satisfy C_1. However, if sameness

4 Against this, Baker (2000) and Sutton (2014) both argue that an entity must have its person-making characteristics essentially if it is to count as a person. On this view, animals cannot be persons in any sense. I think this is mistaken. An animal, on my view, can have a phase of being a person in the broader sense, in virtue of constituting (or being otherwise metaphysically related to) a Lockean person – but an animal cannot be a *Lockean* person. Of course, this means that there can be two persons at my location – one Lockean, with psychological persistence conditions, the other merely a phase of the animal. But does this not lead to a problem of having *too many persons* at my location? Not necessarily. Certainly, we do not want there to be more than one first-person perspective at the same location. But I think it is possible to have an animal that has a phase of having a first-person perspective (and which therefore has a phase of being a person in the broader sense), and for this very same first-person perspective to be an essential feature of a Lockean person.

of person across time is not transitive, but is rather a matter of degree that can (so to speak) fade out over time, then the persistence conditions of a person may change over time – even as the person remains the same. To take the current example, it could be that P_3 is the same person as P_2, and P_2 is the same person as P_1, but P_3 is not the same person as P_1. But this would mean that P_2 and P_1, despite being the same person, have different persistence conditions – since P_3 satisfies C_2 but not C_1.

Now, for present purposes I will not take any view on whether or not sameness of person across time is a transitive relation; and nor am I claiming that it is possible for something to go from having psychological persistence conditions to having non-psychological persistence conditions. The point is just that we cannot help ourselves to a general principle that a thing cannot change its persistence conditions without ceasing to exist. So this is not the reason why being a Lockean person cannot be a mere phase of an animal.

Fortunately, there is a simpler reason, which is that a phase cannot outlast the object of which it is a phase. And, if we are Lockean persons, then we can in principle outlast our animals – and so our being Lockean persons cannot be a mere phase of our animals. By the same logic, being an animal cannot be a mere phase of a Lockean person – because, if something is an animal, then it can survive the permanent loss or destruction of its associated Lockean person, for example if it enters a persistent vegetative state. This does not mean, of course, that a Lockean person cannot have a phase of being *associated* with an animal, such as by being materially constituted by or inhabiting one – just that being numerically identical to an animal cannot be a phase of a Lockean person.

So an animal cannot be a phase of a Lockean person, and a Lockean person cannot be a phase of an animal. But is it nonetheless possible for a Lockean person and an animal to be numerically identical? That is, is it possible for something to simultaneously have both the property of being an animal and the property of being a Lockean person? I will argue that it is not.

At first sight, it may not seem obvious that a Lockean person cannot be identical to an animal. After all, Lockeanism is not first and foremost a theory of our fundamental nature, but rather an account of our persistence conditions. Indeed, Lockeanism is consistent with both physicalist and dualist accounts of consciousness, and of our nature.[5] In itself, it is silent on whether we have

5 Indeed, on some definitions, physicalism is compatible with the claim – made by some Lockeans – that I could persist in an immaterial form. For example, Jackson (1998) defines physicalism as the claim that my physical properties are metaphysically *sufficient* for my consciousness –

doi: 10.48106/dial.v78.i1.03

parts, and if so which ones, or on how we relate to human animals. However, this does not mean that Lockeanism is completely silent on our nature: if Lockeanism is true, then our fundamental nature, whatever it may be, must be such that we have psychological persistence conditions.

But the nature of animals is such that their persistence conditions, whatever they may be, are not psychological.[6] For example, some humans survive in a persistent vegetative state, with no natural possibility that their previous psychological continuity will ever be restored. In such situations, the human animal persists, even though the psychological continuity of the person has been extinguished. Similarly, every animal was once an embryo without any mental life. Therefore psychological continuity (or even the mere presence of psychological states) is not a *necessary* condition for an animal to persist in existence; but nor, as shown by the neo-Lockean thought-experiments involving brain transplants, teleporters and so forth, is psychological continuity *sufficient* for the survival of the animal (let us suppose that the brain-donor animal and the human animal that steps into the teleporter are both destroyed in the course of the experiment). So animal persistence conditions, whatever else they may be, are not psychological persistence conditions.

So the persistence conditions of animals and those of Lockean persons are completely different: for Lockean persons, psychological persistence is necessary and sufficient for continued existence; for animals, animal conditions (be they somatic or organic) are necessary and sufficient. And these conditions are not merely different – they are incompatible. If psychological

and therefore, from a Lockean perspective, that they support my persistence conditions. But this does not entail that the continuation of my physical properties is metaphysically *necessary* for the continuation of my consciousness: it leaves open the possibility that my consciousness might continue – and thus, by Lockean standards, that I might persist – in an immaterial form. (Interestingly, this would seem to entail that being a material object is not an abiding sort: if both Lockeanism and physicalism are true, I could in principle cease to be a material object and yet continue to exist.)

6 What then are the persistence conditions of animals? The main options are: first, that the persistence conditions of animals relate to the persistence of organic life – that is, an animal is a mass of particles organised in a common life, and the survival of the animal is a matter of the continuity of these organic processes of life; second, that the persistence conditions of animals are somatic – that is, that the survival of the animal is a matter of the physical persistence of its body, whether or not organic life persists. If the persistence conditions of animals are organic in the aforementioned sense, then the lifeless body of an animal is not strictly speaking an animal, since its organic life has ceased; but if they are somatic, then the lifeless body of an animal is a dead animal. Either way, the important point is that the persistence conditions of animals have nothing to do with the continuity or otherwise of their psychological properties.

persistence is sufficient for survival, then animal persistence is not necessary; and if animal persistence is sufficient for continued existence, then psychological persistence is not necessary. So nothing can have necessary and sufficient persistence conditions that are psychological and necessary and sufficient persistence conditions that are animal; and therefore nothing can be both a Lockean person and an animal.

Now, it may be objected that an animal and a Lockean person might be numerically identical despite their irreconcilable persistence conditions, if they are both phase sortals. We have already seen that one cannot be a mere phase of the other – but what about the possibility that both may be phases of some underlying entity? For example, the necessary and sufficient conditions of being a philosopher are that one does philosophy – it is neither necessary nor sufficient that one plays football. And the necessary and sufficient conditions of being a footballer are to play football – it is neither necessary nor sufficient that one does philosophy. But the fact that the property of being a footballer and the property of being a philosopher have completely different essential conditions does not mean that something cannot instantiate both simultaneously. Thus a single object, namely a human being, can have overlapping phases of being a philosopher and a football player (ignore, for this example, the small matter of whether it is the Lockean person or the animal who philosophises and plays football).

But the situation with animals and Lockean persons is not like this. Something that is a footballer, or a philosopher, can cease to be a footballer, or a philosopher, and yet continue to exist. When we specify the necessary and sufficient conditions for something to persist as a footballer or a philosopher, we are not thereby specifying the conditions for it to persist in existence. But the persistence conditions of a Lockean person are the conditions for something that is a Lockean person to continue to exist; and the persistence conditions of an animal are the persistence conditions for something that is an animal to continue to exist. In other words, *Lockean person* and *animal* are abiding sorts, not phase sortals. And it cannot be the case that a single object is such that psychological continuity is sufficient for its continued existence, yet animal persistence is necessary, or that animal persistence is sufficient, yet psychological continuity is necessary.

To be clear, this does not rule out the possibility of something having persistence conditions that involve the *conjunction* of psychological and animal conditions. The difference here is between psychological and animal conditions being individually necessary and sufficient, in the case above, and them

doi: 10.48106/dial.v78.i1.03

being individually necessary but only *jointly* sufficient, in the conjunction case. However, I will rule out this conjunction view on the grounds that it is incompatible with Lockeanism: if animal and psychological persistence conditions are only jointly sufficient for my survival, then I could not, by definition, survive the destruction of my animal.

And what about *disjunctive* persistence conditions? The idea here is that my persistence consists of the persistence of *either* psychological *or* animal conditions. Thus Langford (2014) argues for the view that we are what he terms *bio-psycho continuers* – that is, entities that will survive if either the appropriate psychological continuity or the appropriate biological continuity occurs. The problem with this view, though, is that while it allows that psychological and biological continuity are both individually sufficient for my survival, it denies that either is individually necessary. But if this is true, then we are *neither* Lockean persons *nor* animals. The Lockean intuition is not just that I could survive the teleporter; it is also that there is a sense in which I would *not* survive in a persistent vegetative state. Because my psychological continuity would be extinguished, there is an important sense in which I would cease to exist. Similarly, the animalist intuition is not just that animal continuity is sufficient for my persistence – it is also the intuition that, in some sense, I would *not* survive a brain transplant that destroyed my donor body, and could *not* continue to exist in an immaterial afterlife. Now, in the case of animalism, it may be objected that this intuition goes beyond the core claim that we are animals. Perhaps, one might think, I am identical to an animal, but the animal to which I am identical could persist in an immaterial form, or would go with my transplanted brain. And in that case, my being identical to an animal would not be inconsistent with the disjunctive account of my persistence conditions. But to close this loophole, we need only add the premise (highly plausible, in my view) that animal persistence requires a degree of somatic or organic continuity that is not satisfied in these Lockean scenarios. And so if we think, as I think we should, that there is a sense in which we would not survive in a persistent vegetative state, and a sense in which we, being animals, could not exist in an immaterial form, then we should reject disjunctivism about our persistence conditions.

So, to summarise: anything that is a Lockean person has psychological necessary and sufficient persistence conditions; anything that is an animal has non-psychological necessary and sufficient persistence conditions (be they organic or somatic); and, because being an animal and being a Lockean person are not phase sortals, nothing can have both these sets of persistence

conditions; and therefore nothing can be both a Lockean person and an animal.

2 Why Think That We Are Animals?

Perhaps unsurprisingly, a common Lockean response (see e.g. Baker, 2002) is to deny that we are identical to animals. If I am an animal at all, then it is only in a derivative sense, such as in virtue of by being materially constituted by one. And if I am an animal in at most a derivative sense, then I will have certain animal properties in at most a derivative sense, in virtue of the close metaphysical relationship that I have to the animal that, strictly speaking, has them. Suppose, for example, that metabolism and immunity to certain diseases are, strictly speaking, properties of my animal, and not of the Lockean person that I am identical to. On this view, although I do not strictly speaking have these properties, I may be said to have them in a derivative sense because I am materially constituted by, or otherwise metaphysically related to, the entity that properly has them.[7]

I find this position unsatisfactory. My view is that we are animals in a non-derivative sense, and that we have our animal properties directly and non-derivatively. But why should we think this? There are of course numerous arguments for animalism. In Section 3, I will consider Olson's (2003) *thinking animal* argument in some detail. Here, I will briefly consider some other arguments for animalism.

First, the *animal ancestors* argument (Blatti, 2020) is supposed to be a *reductio ad absurdum* of any theory that denies that we are animals, since this would lead – so the argument goes – to the denial of evolution. In essence, the argument runs as follows: if I am not an animal, then nor were my ancestors; and if my ancestors were not animals, then nor were their ancestors; and so on; but, in that case, I am not the product of evolution through natural

7 Can we be more definitive about which of my animal's properties would, on this view, be had by me in at most a derivative sense? We cannot simply say that *any* property had by my animal will be had by me in at most a derivative sense, since my animal will – derivatively, and in virtue of constituting a Lockean person – have some properties that I have non-derivatively. But nor can we say that any property had non-derivatively by my animal will be had by me (at most) derivatively, since there will be some properties that my animal and I will both have non-derivatively (such as our temporal properties, perhaps). But do I think it is reasonable to say that, on this view, any property that my animal has in virtue of being an animal, but which I do not have just in virtue of being a Lockean person, will be had by me, the Lockean person, in at most a derivative sense.

doi: 10.48106/dial.v78.i1.03

selection; but this is absurd, since we know evolution to be true; therefore I am an animal.

However, I do not think this argument is particularly compelling. Even if I am not identical to an animal, it remains the case that my animal is an animal, and that my animal's ancestors are themselves animals, and so on. It is my animal that is the product of evolution through natural selection, regardless of whether or not I am numerically identical to it. Moreover, it is animals that have ancestors, since reproduction is a function of animals; and, if I am not my animal, then it is not clear that, strictly speaking, I have ancestors. Now, if I am not my animal, then there will no doubt be questions about how and at what point in evolutionary history animals came to be metaphysically connected to Lockean persons. But these are part and parcel of the wider question of how (on the Lockean view) Lockean persons and animals are metaphysically related.

The *association* argument for animalism (Bailey, 2015) proceeds from the observation that my animal and I seem always to be closely associated with each other: wherever I go, my animal goes; whenever my animal is tired, or sick, I am tired, or sick; and so on. The simplest and best explanation of this phenomenon, so the argument goes, is that my animal and I are constantly associated because we are identical. But, once again, I do not think this argument is compelling. If I am not an animal, but a Lockean person that is (let us say) materially constituted by an animal, then I would expect to be associated with my animal in the way that we actually observe. Now, it is true that material constitution (for example) is a somewhat more complex explanation of the association than strict identity would be. But it is not much more so, and the fact remains that both explanations are able to account for the observed datum of the association. And if we have independent reason to think that we are Lockean persons – if, for example, we are persuaded by the thought-experiments of Locke, Parfit and others that we have psychological persistence conditions – then the mere fact that I am associated with an animal does not provide a compelling reason to think that I am identical to that animal as opposed to being a Lockean person.

A stronger reason, in my view, for thinking that we are animals is that the standard Lockean account – which says that I am a thing that is constituted by an animal, but not an animal (cf Baker 2002) – cannot explain how I existed as a foetus, or how I could continue to exist in a persistent vegetative state. In other words, whilst it may be plausible that a Lockean person could have animal properties such as metabolism and disease immunity in a derivative

sense, in virtue of its close connection to the animal that properly has them, it is much less plausible that a Lockean person can have animal persistence conditions in a derivative or secondary sense. How can we meaningfully say that I, being a Lockean person, would not *strictly speaking* exist if my animal were in a persistent vegetative state, but that I would exist *in a derivative sense*? If, as per the standard Lockean view, the first-person pronoun refers to a Lockean person and not to an animal, then it has no reference at all when my animal is a foetus or in a persistent vegetative state. In those circumstances, I will not be around to have any properties at all, whether derivative or otherwise. So the problem for the standard Lockean view is that our ordinary commitments about ourselves, if taken at face value, entail not just that we have animal properties, but that we have animal persistence conditions.

Of course, the standard Lockean is free to reject the claim that I existed as a foetus and would continue to exist in a persistent vegetative state, just as the pure animalist is free to reject the claim that I would survive in brain transplant and teleportation scenarios. But if we accept both sets of claims, then we must accept that the first-person pronoun refers to a thing with psychological persistence conditions, namely a Lockean person, and to a thing with animal persistence conditions, namely an animal.

3 I Am Not One Thing, but Two

It seems that we are at something of an impasse: there is good reason to think that we are Lockean persons, with psychological persistence conditions; and there is good reason to think that we are – literally and strictly, and not merely derivatively – animals; and yet an animal cannot be numerically identical to a Lockean person.

But this impasse arises only because we have made the implicit assumption that my use of the first-person pronoun refers to just one thing – and that the questions of our fundamental nature and persistence conditions concern the fundamental nature and persistence conditions of that one thing. But what if my use of the first-person pronoun does not just refer to one entity, but is ambiguous between two? And so I propose the following solution to the impasse: that there are two metaphysically distinct entities at my location – a Lockean person and a human animal – and that my use of the first-person pronoun is ambiguous between these two entities. Thus the first-person pronoun has a sense in which it will be true that *I am a Lockean person*, and a sense in which it will be true that *I am identical to an animal*. (By the same token, it will also

doi: 10.48106/dial.v78.i1.03

have a sense in which it is false that I am a Lockean person, and a sense in which it is false that I am a human animal – I will return to this point in due course.)

It is important to be distinguish the claim that I am making here from certain things that I am *not* saying. I am not saying that we are complex entities containing both Lockean persons and human animals as parts: for if I am identical to a complex containing parts, then I am not identical to any of the parts – and so I would not, in that case, be either a Lockean person or a human animal.[8]

Another important point is that, while the title of this section proclaims that *I am not one thing but two*, we have to be careful with this slogan. Strictly speaking, as I clarified in the Introduction, there is no disambiguation of 'I' on which it is true that I am two things. Rather, there is a disambiguation on which it is true that I am one thing, and a disambiguation on which it is true that I am another.

Moreover, it is important to be clear that the ambiguity attaches to the first-person pronoun itself, and not to the verb 'to be'. So it is not that I *am* a Lockean person in one sense and I *am* an animal in another sense, such as by being numerically identical to a Lockean person, and materially constituted by an animal. Rather, my claim is that the first-person pronoun has one disambiguation on which we are strictly identical to Lockean persons, and another on which we are strictly identical to human animals. In this respect, I agree with Sutton (2014), who argues that, in ordinary usage, the first-person pronoun is ambiguous in this way, referring sometimes to ourselves as persons, sometimes to ourselves as animals, and often to both without distinction. The core idea about self-reference that I am advocating is the same as Sutton's, although we arrive at the same place via different (but complementary) routes: Sutton's aim is to show how it is not necessarily problematic for there to be more than one thinker (and, by extension, more than one candidate for reference of 'I') at my location, as long as the thinkers in question share a

8 This position is similar to the conjunctive account of our persistence conditions, which I rejected above. There is a subtle difference, though: on the conjunctive account, there is one thing, which will persist if and only if there is both psychological and animal continuity, and I am identical to that thing; on the present, 'complex' account, there are two things – one with psychological persistence conditions, one with animal persistence conditions – and I am identical to the complex entity consisting of both things. But the difference is a fine one, and I will reject both positions on the grounds that they are incompatible with my being a Lockean person.

supervenience base; my aim is show how we can respect the intuitions that motivate both animalism and Lockeanism.

If self-reference is indeed ambiguous in the way that Sutton and I propose, then there is likely to be a division of labour, so to speak, between those properties that we have as animals (or those predicates that are true of us as animals), and those properties that we have as Lockean persons. Some properties we will have as animals, but not as Lockean persons; some we will have as Lockean persons, but not as animals; and some we will have as both animals and Lockean persons. So, plausibly: my animal has organic properties such as metabolism and immunity to certain diseases; my Lockean person has the property of being able to survive teleportation; and (as I shall argue below) both my entities have the property of thinking my thoughts. The precise division of labour, of course, will depend on one's metaphysical theory of Lockean persons and their relation to animals. It is an open question, for example, whether my Lockean person has the same mass as my animal, or indeed any mass at all; similarly whether, if my animal is walking, my Lockean person is also walking in a non-derivative sense. The exact details of this division of labour need not concern us for present purposes. Under normal circumstances, we do not even notice that there is a division of labour with respect to our properties, because the two entities overlap in time and space and in their causal relations to the outside world. But there are certain circumstances under which the two entities have the potential to come apart, and the division of labour is important. For example: I, the animal, but not I, the Lockean person, existed as a foetus; and I, the Lockean person, but not I, the animal, could in principle survive the destruction of my animal.[9]

9 We can imagine hypothetical situations in which the reference of 'I'-thoughts is complicated by the spatial separation of an animal and its Lockean person. For example, Dennett (1978) imagines a scenario in which his living brain is separated from the rest of his body and kept alive in a vat. The neural connections between his brain and his still-living body are maintained by means of implanted radio transmitters and receivers, connecting the brain to the nerve stumps in his now-brainless skull, so that the neural pathways between brain and body are perfectly preserved. Sensory inputs travel from the body to the brain, and motor signals flow in the other direction. Dennett then imagines himself in his now-brainless body, looking at his envatted brain. Or should he say that he is in the vat, being observed by his body? Where is he?

What should we say about the reference of Dennett's 'I'-thoughts in this hypothetical scenario – and, in particular, about the spatial location of their referent? With respect to Dennett's animal, the situation seems relatively straightforward: if the separated brain is not deemed to be part of the animal, then we should say that the animal is located wherever the brainless body is located; but if, as may seem more plausible, the living brain is deemed a part of the animal, then we ought to say that the animal has a part in the vat and a part outside of it. Hence if Dennett's

doi: 10.48106/dial.v78.i1.03

But what about thought? As I mentioned above, my view is that both the animal and the Lockean person are thinking my thoughts. However – to anticipate an objection that I will address in more detail in the next section – this does not mean that there are two parallel streams of thought at my location, one belonging to the animal and the other to the Lockean person. Rather, there is only one stream of thought at my location; but there are two entities at my location, each of which is metaphysically related to my thoughts in such a way that each is thinking them (albeit in slightly different senses of 'thinking'). So there is one stream of thought, but two thinking entities.

This brings me to the question of how we should understand thoughts about ourselves. How does self-reference work? Which thinking entity is being referred to, and which is doing the referring? First, we must distinguish between, on the one hand, a thought in the sense of an occurrent mental state that is the bearer of semantic content (a thought in this sense may be, depending on one's view, a brain state, or a non-physical token of phenomenal consciousness) – and, on the other, the semantic content that the thought bears. As already noted, there is only one stream of thoughts at my location, which is common to both the animal and the Lockean person. Thus, any time I think an 'I'-thought, there is only one occurrent mental state, which is common to both thinking entities. However, the semantic content of this single mental token is ambiguous. The idea, in outline, is that the first-person

'I'-thoughts are taken to refer to Dennett's animal, the location of their referent seems relatively unproblematic. But where, in this scenario, is Dennett's (Lockean) person located? It seems to Dennett that he, the person, is located with his body – that is where is first-person perspective resides, after all. He has no sense whatsoever of being in the vat. And yet, as a physicalist, he takes the view that his conscious states are identical to his brain states – which are physically located in the vat. So which is the true location of Dennett, the person?

I am not sure we can give a single, unambiguous answer to this question. It seems to me that there are two ways to understand the location of the Lockean person. First, there is the location in physical space of the causal (or metaphysical, depending on whether dualism or physicalism is true) substratum of consciousness – that is, the location of the relevant brain-states. Second, there is the apparent location of the Lockean person's first-person perspective. This is determined not by the actual location of the causal substratum of consciousness, but by the content of one's sensory experience. In Dennett's case, the apparent location of the first-person perspective is at his brainless body. I say *apparent* location to emphasise that it is determined by the content of consciousness, not by the spatial properties of the entities that are the bearers or the underlying causes of consciousness. Moreover, the apparent location of the first-person perspective may not even correspond to a location in physical space – for example in scenarios where we are immersed in a convincing but virtual world. Normally of course, these two locations of the Lockean person – the apparent location of the first-person perspective, and the physical location so the causal substratum – coincide. But there is no reason to suppose that they have to always coincide.

pronoun has a narrow content that, for any token occurrence of the pronoun in thought, will pick out whatever entity is thinking the thought in which the token is embedded. In other words: *I* am the thinker of *this thought*. But of course, on the model I am proposing, for any thought, there are two entities that are thinking it, namely a Lockean person and a human animal. This means that any occurrence of the first-person pronoun will be ambiguous in its reference between these two entities, at least in the absence of any contextual information to disambiguate the reference.

And if the first-person pronoun is ambiguous, then statements containing it will also be ambiguous. When I talk about myself, I could be talking about a Lockean person, or about a human animal (or both, without distinction). As noted above, the result is a sort of division of labour: some statements that I make about myself will be true regardless of how the first-person pronoun is disambiguated; but some will be true only if the first-person pronoun is interpreted as referring to the Lockean person, and some will be true only if it refers to the human animal. Or, to put it another way, some statements will be true in one sense, and false in another. Hence it is true that I, the Lockean person, could survive teleportation, but false that I, the animal would survive; and it is true that I, the animal, would persist in a persistent vegetative state, but false that I, the Lockean person, would persist.

This, though, raises another worry. The statements 'I am a human animal' and 'I am a Lockean person' will each have a disambiguation of 'I' on which they are true, and a disambiguation on which they are false: it is true of a human animal that it is a human animal, but false that it is a Lockean person; and it is true of a Lockean person that it is a Lockean person, but false that it is a human animal. And, moreover, this means that there is a sense in which it is true that I am *not* a human animal, and another sense in which it is true that I am *not* a Lockean person.

While this may seem alarming, we need to keep in mind what is going on here: the first-person pronoun is ambiguous, and refers both to a thing that is a Lockean person (and not a human animal), and to a thing that is a human animal (and not a Lockean person). The truth-conditions of the various disambiguated 'I'-statements are the just consequence of this ambiguity. So, while there is a sense in which it is true that I am not a Lockean person, this just means that 'I' can refer to a thing that is not a Lockean person. But this is very different from saying that it does not refer to a thing that is a Lockean person. And similarly, *mutatis mutandis*, for the human animal.

doi: 10.48106/dial.v78.i1.03

Now, if the first-person pronoun is ambiguous in this way, then it is tempting to think of each disambiguation as belonging to – that is, as being thought by – the corresponding thinker. That is, it is tempting to suppose that the human animal thinks the 'I' that refers to the human animal, and the Lockean person thinks the 'I' that refers to the Lockean person. But this is a mistake. As Sutton (2014) argues, we should not think that each disambiguation of an 'I'-thought must be assigned to either the Lockean person or the animal. Rather, there is only one token occurrence of the 'I'-thought, in the sense of one content-bearing mental state, and this thought is common to both the animal and the Lockean person. Thus each disambiguation of the thought will belong to – will be thought by – both thinkers.

Still, the idea may persist that each thinking thing is referring only to itself, and not to the other entity. Consider the following scenario: two protestors, A and B, are together holding up a sign that says, "I am angry". By analogy with my view, we ought to say that each protester is expressing both that A is angry, and that B is angry. But would it not be more natural to regard each protestor as expressing only of himself that he is angry?[10] The fact that they are jointly holding one sign, rather than each holding their own individual signs, hardly seems to be the decisive factor. So why should we think that the situation is different when it comes to having two thinking entities thinking one thought?

But now suppose that the sign, instead of saying, "I am angry", had said' "the holder of this sign is angry". In that case, it might seem more natural to regard each protestor as intending to communicate, not just that he himself is angry, but that both he and the other sign-holder are both angry – to have intended the ambiguity, so to speak. (If the sign had read, "the holders of this sign are angry", then there would of course be no ambiguity, and we would interpret each sign-holder as communicating something about both). However we interpret the intentions of the sign-holders, the point is that it makes sense for us to assess – using whatever pragmatic cues are available in the situation – what each individual sign-holder intends to communicate, and what the intended reference of his communication is.

However, it is not clear that we can perform the same sort of assessment when we are dealing with the case of a Lockean person and a human animal, both of whom are thinking the same token thought. What we are able to do – at least in some situations – is to make a pragmatic assessment of the

10 I am grateful to an anonymous referee for posing the challenge in this way.

intended *reference* of a human person's use of the first-person pronoun. For example, if someone says, "I could survive the teleporter", and, given all the circumstances, we take them to intend to express a truth (and we know that they are schooled in the philosophy of personal identity), then we might take them as intending to refer just to the Lockean person, and not to the human animal. If, on the other hand, the human person had said, "I would continue to exist in a persistent vegetative state", then we might take them as intending to refer just to the human animal, and not the Lockean person. So it is not the case, in my view, that *every* occurrence of the first-person pronoun as used by a human person has to be ambiguous in its reference: some occurrences will be intended to refer one way, and some the other, and we can often use pragmatic cues to work out the best interpretation of a speaker's intentions.

But it is one thing to disambiguate the intended reference of a particular use of the first-person pronoun; it is another to draw a distinction between the person's act of reference and the animal's act of reference. In the case of the sign-holding protesters, it makes sense for us to assess the communicative intentions of each individual separately. Protestor A has a communicative intention, and protestor B has a communicative intention, and both of their intentions are numerically distinct things (and they are also distinct from the sign itself, which is analogous to the mental state shared by the animal and Lockean person that is the bearer of their content). But this is not the case when a human animal and a Lockean person are thinking the same token thought. Their shared thought is not like a sign that they are jointly holding up. It is not as if there are three token thought-entities in the picture: the animal's communicative intention, the Lockean person's communicative intention, and the mental state that they have in common; rather, there is just the one occurrent thought, which they are both thinking. So whilst it may be possible, depending on the context, to disambiguate the intended reference of that thought, the act of reference itself will belong equally to both the animal and the person.

Moving on, we see that this theory of first-person reference can be broadened in scope in two ways. First, we can expand the scope to cover more than just first-person reference. Our self-reference does not in itself cause there to be both a thinking animal and a Lockean person at my location; nor does it cause both of these entities to be thinking the same numerical stream of thoughts. The metaphysical facts are independent of our self-reference: the ambiguity of the pronoun 'I' merely tracks (or rather, fails to distinguish between) these metaphysical facts. So there is no reason why this account

doi: 10.48106/dial.v78.i1.03

cannot be extended to second and third-person reference as well. If I am both a Lockean person and an animal (to speak informally), then so are *you*, and so are *they*.

Second, we might ask: what about non-human animal persons? If non-human animals do not have Lockean persons, then the issue that this theory is intended to solve will not arise, for they will simply be animals (and perhaps also persons in some broader, phase sortal sense). But if non-human animals do have Lockean persons, then the same account is available for them as well.[11]

4 Too Many Minds? Too Many Thinkers?

Perhaps the main objection to my thesis – that my use of the first-person pronoun is ambiguous in its reference between two thinking things at my location – is that it will supposedly lead to what Shoemaker (2008) calls the *too many thinkers* problem. In simple terms, the worry is that if my use of 'I' can refer to either a Lockean person or a human animal, then there will be two distinct entities that are both thinking my thoughts. There may even be two minds at my location, one belonging to the animal and the other to the Lockean person. Thus the claim that 'I' is ambiguous seemingly collapses into the view that there are two distinct persons or minds – one animal, one Lockean – both inhabiting the same portion of space and time, and each thinking my thoughts. This seems to raise epistemological worries about which of these entities I really am. If I think the thought 'I am a Lockean person', then how do I know that I am the Lockean person thinking truly, and not the human animal, thinking falsely? And similarly for the thought 'I am a human animal'?

This worry is at the heart of Olson's (2003) *thinking animal* argument for animalism. This argument can be summarised as follows: (i) there is an animal at my location; (ii) the animal at my location is thinking; (iii) I am

11 For non-human animals that are capable of self-reference (if there are any), then the same principles that I outlined above in respect of our self-reference here will apply to them. For non-human animals that are not capable of self-reference (which one imagines is most, if not all of them), the issue of how to understand their self-reference will of course not arise. But there is no reason to think (without further argument) that an absence of self-reference entails the absence of a Lockean person. And if non-human animals do have Lockean persons associated with them, then it is also possible that our third-person reference to them has the potential to be ambiguous between the non-human animal and the associated Lockean person (at least in the absence of contextual information that removes any ambiguity).

at my location; (iv) I am thinking; but (v) there is only one thinker at my location; and therefore (vi) I am the thinking animal at my location.

I will take it that the claims (i), (iii) and (iv) are true – that there is an animal at my location, that I am at my location, and that I am thinking. I will also take it that the argument is essentially valid. So the important questions are whether the animal at my location is thinking, and whether there can only be one thinker at my location. If both of these claims are true, then the thinking animal argument will indeed show that I am identical to an animal. Moreover, this would not just show that I am an animal: it would also show that I am *not* a Lockean person. This is because Lockean persons, whatever else they may be, are certainly thinkers; but, as I have argued above, a Lockean person cannot be identical to an animal; and so, if there is a thinking animal at my location, and if there can only be one thinker at my location, then there cannot also be a Lockean person at my location – and therefore I cannot be a Lockean person. So the *thinking animal* argument presents a problem not just for traditional Lockeans, who deny that I am identical to an animal, but also for my claim that my use of 'I' can refer to *both* a Lockean person *and* an animal.

How then should we respond to the *thinking animal* argument? Lockeans who wish to resist the conclusion that I am an animal have several options. One is to deny premise (ii) – that is, to deny that the animal at my location is really thinking at all (see, for example, Shoemaker 2004). Another option is to reject premise (v) – that is, to argue that there can be more than one thinker at my location. Thus Baker (2000) argues that the Lockean person to which I am identical and the animal at my location are both thinking, but in different senses: the Lockean person is thinking in a non-derivative sense, and the animal is thinking in a derivative sense, in virtue of constituting the Lockean person.

But unlike both Shoemaker and Baker, I do not wish to deny that we are animals. Unlike Shoemaker, I do not deny premise (ii). In fact, I think it implausible to deny that the human animal at my location is thinking – after all, this animal has all of the brain states that are (at the very least) causally correlated with my thinking. Yet what is the animal doing with all of these brain states, if not thinking?

But what about the constitutionalist version of Lockeanism, as advocated by Baker (2000, 2002)? I agree with this view in its rejection of premise (v) of the *thinking animal* argument – like Baker, I think there are (at least) two thinkers at my location. Moreover, it can be plausibly argued that this claim –

doi: 10.48106/dial.v78.i1.03

that there are two thinking things at my location – does suggest some version of constitutionalism, even if it does not strictly entail it. How so? If there are two thinking things at my location, this raise the question of whether one of them could be thinking without the accompaniment of the other. Now, it seems a fairly straightforward consequence of Lockeanism that a thinking Lockean person can exist without an associated thinking animal. The more interesting question is whether, on my view, we ought to say that there could in principle be a thinking animal without an accompanying Lockean person. Here we face something of a dilemma. We can ask: would a zombie duplicate of a thinking animal – that is, a being that is a physical duplicate of a thinking animal, but which lacks phenomenal consciousness, per Chalmers (2010) – count as thinking? If the answer is 'Yes', then it seems to follow that a thinking animal can exist without an accompanying thinking Lockean person – since, plausibly, a Lockean person requires at least some degree of phenomenal consciousness. But the cost of this is that animal thinking would be a very different sort of thing than the thinking done by Lockean persons. The latter would necessarily involve phenomenal consciousness, but the former would not. And we might then wonder whether it is even appropriate to refer to both sorts of activity as 'thinking'. But if, on the other hand, the answer is 'No', then we are arguably closing in on some sort of constitution view: if we have a thinking animal, then we have an animal that is phenomenally conscious; but if we have phenomenal consciousness, then it is hard to see how we do not have a Lockean person; and therefore the existence of a thinking animal would metaphysically entail the existence of an accompanying Lockean person.[12]

So I agree with Baker in one important respect, namely that there can be more than one thinker at my location; and it is at least plausible this in itself leads to some form of constitutionalism. However, there is one crucial respect in which my view differs from the constitutionalist account offered by Baker: unlike her, I think that my use of the first-person pronoun does refer to the thinking animal at my location (because, as noted earlier there is a sense in which I would still exist if this animal ceased thinking and entered a persistent vegetative state).[13]

12 Note also that one could accept the constitution view without having to embrace some form of physicalism about phenomenal consciousness: we could hold that the Lockean person is constituted by the physical properties of the animal plus appropriate contingent psycho-physical connecting laws.

13 Another point of disagreement with Baker is that, in her view, our Lockean persons are necessarily embodied. So, while I, the Lockean person, could in principle survive the destruction of my

But how can there be more than one thinker at my location? The first thing is to remember, as discussed in the previous section, that there is a distinction between having multiple thinkers and having multiple minds at the same location. The paragraph at the start of this section, in which I outlined the supposed problem of having too many thinkers, elided that distinction – but it is a crucial one, because although having more than one mind at my location would be a problem, having more than one thinker need not be.

Consider the following two scenarios. The first is that there are two entities at my location, namely a Lockean person and a human animal, each of which is thinking numerically distinct occurrences of the same qualitative type of thought. Thus there are two parallel streams of consciousness, one of which is being thought by each thinker. So in this scenario, there are not only two thinkers at my location, but two distinct minds. And so there seems to be a genuine question about which thinker is me, and which mind is mine. And not only is there a genuine question but – even worse – there seems to be no way of answering it. *Ex hypothesi*, the animal's mind and the Lockean person's mind are subjectively the same, even though they are numerically distinct. To return to the question posed at the beginning of the section: if I think 'I am an animal', then how do I know that I am the animal thinking truly, and not the Lockean person thinking falsely?

But another scenario is that while there are two thinkers, there is only one mind. In this scenario, while there is only one stream of consciousness at my location, there are two entities – a Lockean person and a human animal – that are metaphysically related to the single stream of consciousness in such a way that each is thinking it. Now, while there being two minds at my location would give rise to the epistemological worries just noted, it is much less clear that having two thinking entities for one stream of consciousness would – in itself – be a problem. For example, at the same time as my animal is thinking my thoughts, it is also the case that my living brain is thinking my thoughts, as is my un-detached head, as is my left-hand complement – and so on. For any given stream of thoughts, there are multiple different and overlapping ways to delineate the immediate causal ground of those thoughts, and this does not (or does not obviously) give rise to a problem of having too many

animal, I could not continue to exist without being embodied in some form. My view is that, while it is plausible that my living animal is metaphysically *sufficient* to constitute my Lockean person, there is no reason to think that continued embodiment in some form is metaphysically *necessary* for my Lockean person to survive. But we can ignore this point of disagreement for present purposes.

doi: 10.48106/dial.v78.i1.03

minds. Whichever way we cut things, it seems that any reasonable ontology must allow for more than one thinking entity at my location – and we do not normally find such part-sharing scenarios to be problematic in terms of containing too many thinkers. This point is also made by Sutton (2014), who argues that having both a thinking animal and a Lockean person at my location is no more problematic than the parts-sharing cases. The essential point, for Sutton, is that where a Lockean person and her animal (or body) share a supervenience base for thinking, the thinking that results is not *summative* – that is, we do not have to count the Lockean person's thinking separately from the animal's.

This seems quite right to me.[14] The essential point, as I see it, is that having two thinking things in our ontology need not commit us to more than one stream of thought, and it is only this that would be problematic. Moreover, in the two-thinkers-one-mind scenario, there is no epistemological worry about which mind is my mind, because there is only one mind – there is one stream of consciousness that is common to both thinkers, not two parallel streams of consciousness. But is there an epistemological worry about which *thinker* I am though? No. If I think 'I am a human animal', then, as I argued above, both the Lockean person and the human animal are thinking this thought. But then is the thought true or false? Because of the ambiguous reference of the first-person pronoun, the thought will have a disambiguation on which it is true (because the human animal is a human animal), and a disambiguation on which it is false (because the Lockean person is not a human animal).

14 There is a tangential – but important – matter on which I disagree with Sutton's account, but it does not affect the present argument. In short, his supervenience solution to the too-many-thinkers problem is supposed to apply only if we assume a physicalist account of persons – that is, if persons supervene *metaphysically* on animals (or bodies). But if persons are not physical, then there is an easy solution, according to Sutton – the Lockean can reject the claim that animals (or bodies) think, on the grounds that their brain states will not be sufficient for thought.

Now, I think this is mistaken on a number of points. First, I do not think that thought needs to supervene *metaphysically* on an animal's physical properties in order for us to say that animals think. It might plausibly be enough for thought to supervene *causally* on the material base. That is, we can be permitted to say that animals (or bodies) think, even if we believe that naturalistic dualism is true. But this means that a non-physicalist who is also a Lockean and who believes that animals can think (I confess that I am in this category) must be able to avail him-or-herself of the two-thinkers-one-mind solution. Fortunately, there is no more reason to think that this solution is only available to physicalist accounts of persons than there is to think that, if dualism is true, then animals will not think. The essential point is that an animal's thoughts and those of a co-located Lockean person need not be counted separately, but can be numerically identical. But this issue is independent of whether physicalism is true.

But, as we have seen, this does not mean that the human animal thinks the true disambiguation of the thought and the Lockean person thinks the false disambiguation. Rather, both entities are thinking the one thought, in the sense that they have in common a single bearer of content – but it just so happens that the content that the thought bears is ambiguous in its reference, and it further so happens that the ambiguity in question relates to the very entities that are thinking the thought.

Now, with all this talk of ambiguity, we might wonder how many semantic contents an 'I'-thought will have. Here we have a choice. We can either say that there is one narrow content, which is ambiguous between two references; or we say that there are two broad contents. I happen to favour some form of phenomenal intentionality theory of mental content (cf Bourget & Mendelovici, 2019), and am a believer in narrow content; moreover, I think that two-dimensional semantics, along the lines developed by Chalmers (2010), provides a useful way to understand narrow content. If one is sympathetic to such theories, then one could take the view that the first-person pronoun, 'I', has a primary intension roughly equivalent to 'the thinker of this thought', where this thought refers to the token content-bearing mental state in which the occurrence of 'I' is embedded. This primary intension would constitute a narrow content that is ambiguous in its reference between the Lockean person and the animal.

But this internalist and two-dimensional account is optional: one need not accept the existence of narrow content to accept the general metaphysical account of persons and of self-reference that I am proposing here. An alternative possibility, as already noted, is that any occurrence of the first-person pronoun will have two broad contents, one relating to the Lockean person and the other to the animal, with no narrow content involved. However, this debate need not concern us for present purposes. For simplicity's sake, I have spoken in terms of one (narrow) content that can be ambiguous in its reference, rather than in terms of two broad contents – but it should be understood that one could re-phrase my overall account according to one's preferred theory of mental content.

Another potential objection to my central claim is that, even if I have avoided the problem of there being too many *minds* at my location, having more than one *thinker* – even for the same mind – raises problems of its own. Arguably, we now seem to have two entities in competition to be the true cause of my thoughts. Is it the person or the animal that causes the thoughts to occur? If only one entity is causally effective, then it is hard to see how we

doi: 10.48106/dial.v78.i1.03

can legitimately say that the other entity is thinking my thoughts. But in that case, then surely my use of 'I' refers to the entity that is actually doing the thinking of my thoughts, in the sense of causing them to occur, and not to anything else? But if, on the other hand, both entities are causally effective with respect to my thoughts, then we seem to face worries, familiar from the literature on mental causation, about causal over-determination. What, in this scenario, ensures that the person and the animal have the same causal output? Must there be causal determinism of both entities' outputs to ensure they remain in synch?

But I think these concerns would be misplaced. We should not see the person and the animal at my location as being in some sort of competition to be the true cause of my thoughts. Recall the earlier point that there are multiple overlapping entities that, at any given moment, are thinking my thoughts: just as my brain is thinking, so is my living body, and my left-hand counterpart, and so on. We do not thereby regard these entities as being in causal competition, but as being metaphysically related in such a way that all are participants in one causal process.

Consider an analogy. There is a computer on my desk whose physical architecture is performing various calculations. But there is also a token of a computer program that is performing these very same calculations. Is there a too-many-calculators problem? Do we have to choose between saying that it is really the physical computer that is doing the calculating, or the computer program? No, because we can readily grasp that the computer and the program are each calculating in a different sense: roughly, the program is a functional system whose states represent information states, and the transitions between these states represent operations performed on that information; the computer itself is the physical system that realises or implements the functional system. The same token program – bearing the same information – could survive the destruction of the computer, if it were transferred to some other physical architecture; and the same computer would persist if the program were wiped, and another installed in its place. We can understand the difference between hardware and software, and that, when there is calculating going on, both hardware and software are involved, and both can be said – albeit in different senses – to be calculating.

Now suppose that one of the states of the program, S, is such that S represents whatever entity produces state S. But what does S represent – the physical computer or the computer program? Plausibly, it can represent either, depending on how we understand the term 'produces' with respect to S. We

can legitimately say that the program produces S because the program is a system of functional operations of which state S is one result. But we can also say that the physical computer produces state S because it implements the program. By analogy, if state S were a phenomenally conscious state, we could equally well say that the computer or the program is the 'thinker' of S.

If we wish to press the analogy, we could take the view that there are different senses of *thinking* in which the person and the animal at my location are thinking my thoughts. But what are these two senses of thinking? Roughly, we might say that the animal is thinking the stream of consciousness in the sense of being the immediate causal basis (or metaphysical ground, depending on one's view) for the occurrence of those thoughts. However, this is not an essential fact about the animal: it would still be an animal were this not the case (if, say, it were in a persistent vegetative state). Thus the animal not only has the property of thinking in an animal-specific sense: it also has this property only non-essentially. But the Lockean person, on the other hand, is thinking the (numerically) same stream of thoughts as the animal in the sense that it is the unified, first-person conscious perspective in which the stream of thought occurs. And so not only does the Lockean person think in a different sense from the animal: unlike the animal, it has the property of thinking essentially.[15]

Now, it may be objected that this proposal – that there are two sorts of thinking, one applicable to the Lockean person and the other to the animal – weakens the idea that animals are really thinking. However, I think this worry would be misplaced. There is no reason, absent further argument, to hold that the sense in which the animal thinks is any less real or meaningful than the sense in which the Lockean person thinks. Consider the constitutionalist model advocated by Baker (2000, 2002), on which my animal is thinking in a derivative sense, in virtue of constituting me, the Lockean person, who thinks non-derivatively. I am sympathetic to Baker's insistence that the animal's thinking is no less real than that of the Lockean person. My objection to Baker is not that, on his view, the animal is not *really* thinking: my objection is that,

15 Moreover, if we combine this view, according to which there are two senses of 'thinking', with the two-dimensional account of first-person reference that I outlined above, according to which the first-person pronoun has an intension roughly equivalent to 'the thinker of this thought', then the ambiguity of self-reference will be grounded in an underlying ambiguity in the notion of *thinking*. Thus, if 'I' refers to whatever entity is thinking the particular thought in which the occurrence of the pronoun is embedded, there will be one sense of 'thinking' in which the animal at my location is thinking my thoughts, and I am that animal, and another sense of 'thinking' in which the Lockean person at my location is thinking my thoughts, and I am the Lockean person.

doi: 10.48106/dial.v78.i1.03

on his view, the animal's thinking is not really *my* thinking. I say that my use of 'I' refers ambiguously to both the Lockean person and the animal – and that, although there may be different senses in which each entity is thinking my singular stream of thoughts, any sense in which they are thinking is a sense in which I am thinking.

5 Conclusion

As I have argued, a Lockean person cannot be identical to an animal. For this reason, it is common for those who accept some form of Lockeanism to deny that we are animals, at least strictly speaking. In this paper, I have hopefully shown that this is not necessary – that is quite possible for Lockeans to accept the claim that we are animals, and not just in a derivative sense. I have argued that there are two thinking entities at my location – a human animal and a Lockean person – and that my use of the first-person pronoun is ambiguous between these two entities. Thus there is one sense in which I am a Lockean person, and another sense in which I am a human animal.

For those of a Lockean persuasion, this approach has certain advantages over the outright rejection of animalism. It permits a plausible division of labour with respect to the properties we attribute to ourselves, such that some are instantiated by the animal, some by the Lockean person, and some by both entities. And I think it also captures the ambiguity that we often feel about our own natures: that in one sense we are animals who are destined to return to the dust from which we are made; and that in another sense we could, in principle at least, persist beyond the destruction of our animal selves.

6 References

*

Harry Cleeveley
0000-0002-3008-8328
London
UK
harrycleeveley@yahoo.co.uk

* I am grateful to Charlie Pelling, Stacie Friend, David Papineau, and several anonymous referees for their helpful comments on earlier versions of this paper.

References

AGAR, Nicholas. 2003. "Functionalism and Personal Identity." *Noûs* 37(1): 52–70, doi:10.1111/1468-0068.00428.

BAILEY, Andrew M. 2015. "Animalism." *Philosophy Compass* 10(12): 867–883, doi:10.1111/phc3.12263.

BAKER, Lynne Rudder. 2000. *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press, doi:10.1017/cbo9781139173124.

—. 2002. "The Ontological Status of Persons." *Philosophy and Phenomenological Research* 65(2): 370–388, doi:10.1111/j.1933-1592.2002.tb00207.x.

BLATTI, Stephan. 2019. "Animalism." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/fall2019/entries/animalism/.

BLOCK, Ned, ed. 1980. *Readings in the Philosophy of Psychology. Volume I*. Language and Thought. Cambridge, Massachusetts: Harvard University Press, doi:10.4159/harvard.9780674594623.

BOURGET, David and MENDELOVICI, Angela. 2019. "Phenomenal Intentionality." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/spr2019/entries/phenomenal-intentionality/.

BRADDON-MITCHELL, David and MILLER, Kristie. 2020. "Conativism About Personal identity." in *Derek Parfit's* Reasons and Persons*: An Introduction and Critical Inquiry*, edited by Andrea SAUCHELLI, pp. 129–159. London: Routledge, doi:10.4324/9780429488450.

CHALMERS, David J. 2010. *The Character of Consciousness*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780195311105.001.0001.

DENNETT, Daniel C. 1978a. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, Massachusetts: The MIT Press. Second edition: Dennett (2017).

—. 1978b. "Where I Am?" in *Brainstorms: Philosophical Essays on Mind and Psychology*, pp. 310–323. Cambridge, Massachusetts: The MIT Press. Second edition: Dennett (2017).

—. 2017. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, Massachusetts: The MIT Press. Anniversary edition, wiht a new introduction by the author; first edition: Dennett (1978a), doi:10.7551/mitpress/11146.001.0001.

GODMAN, Marion, MALLOZZI, Antonella and PAPINEAU, David. 2020. "Essential Properties are Super-Explanatory: Taming Metaphysical Modality." *Journal of the American Philosophical Association* 6(3): 316–334, doi:10.1017/apa.2019.48.

JACKSON, Frank. 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press, doi:10.1093/0198250614.001.0001.

KAPLAN, David. 1989. "Demonstratives." in *Themes from Kaplan*, edited by Joseph ALMOG, John R. PERRY, and Howard K. WETTSTEIN, pp. 481–563. Oxford: Oxford University Press. Widely circulated from 1977 on.

KOVACS, David Mark. 2016. "Self-Made People." *Mind* 125(500): 1071–1099, doi:10.1093/mind/fzv176.

LANGFORD, Simon. 2014. "On What We Are and How We Persist." *Pacific Philosophical Quarterly* 95(3): 356–371, doi:10.1111/papq.12035.

LOCKE, John. 1689. *An Essay concerning Human Understanding*. London: Eliz. Holt, for Thomas Basset, at the George in Fleet Street.

—. 1975. *An Essay concerning Human Understanding*. Oxford: Oxford University Press. Edited with a foreword by Peter H. Nidditch, doi:10.1093/actrade/9780198243861.book.1.

MILLER, Kristie. 2024. "Assessor Relative Conativism." *Journal of the American Philosophical Association* 10(1): 96–115, doi:10.1017/apa.2022.35.

NAGEL, Thomas. 1974. "What is it Like to be a Bat?" *The Philosophical Review* 83(4): 435–450. Reprinted in Nagel (1979, 165–180) and in Block (1980, 159–169), doi:10.2307/2183914.

—. 1979. *Mortal Questions*. Cambridge: Cambridge University Press, doi:10.1017/cbo9781107341050.

OLSON, Eric T. 2002. "What does Functionalism Tell us About Personal Identity?" *Noûs* 36(4): 682–698, doi:10.1111/1468-0068.00407.

—. 2003. "An Argument for Animalism." in *Personal Identity*, edited by Raymond MARTIN and John BARRESI, pp. 318–334. Blackwell Readings in Philosophy. Oxford: Blackwell Publishers.

—. 2007. *What Are We? A Study in Personal Ontology*. Oxford: Oxford University Press, doi:10.1093/acprof:oso/9780195176421.001.0001.

—. 2010. "Personal Identity." in *The Stanford Encyclopedia of Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the Study of Language and Information, https://plato.stanford.edu/archives/win2010/entries/identity-personal/.

—. 2015a. "On Parfit's View That We Are Not Human Beings." in *Mind, Self and Person*, edited by Anthony O'HEAR, pp. 39–56. Royal Institute of Philosophy Supplements n. 76. Cambridge: Cambridge University Press.

—. 2015b. "What Does it Mean to Say that We Are Animals?" *Journal of Consciousness Studies* 22(11/12): 84–107.

PARFIT, Derek. 1971. "Personal Identity." *The Philosophical Review* 80: 3–27, doi:10.2307/2184309.

—. 2012. "We Are Not Human Beings." *Philosophy* 87(339): 5–28, doi:10.1017/s0031819111000520.

SAUCHELLI, Andrea, ed. 2020. *Derek Parfit's* Reasons and Persons*: An Introduction and Critical Inquiry*. London: Routledge, doi:10.4324/9780429488450.

968 SHOEMAKER, Sydney S. 2004. "Functionalism and Personal Identity – A Reply [to
969 Agar (2003) and Olson (2002)]." *Noûs* 38(3): 525–533, doi:10.1111/j.0029-4624.200
970 4.00481.x.
971 —. 2008. "Persons, Animals, and Identity." *Synthese* 162(3): 313–324, doi:10.1007/s112
972 29-007-9253-y.
973 SNOWDON, Paul F. 2003. "Objections to Animalism." in *On Human Persons*, edited by
974 Klaus PETRUS, pp. 47–66. Heusenstamm b. Frankfurt: Ontos Verlag, doi:10.1515/
975 9783110324648.47.
976 SUTTON, Catherine S. 2014. "The Supervenience Solution to the Too-Many-Thinkers
977 Problem." *The Philosophical Quarterly* 64(257): 619–639, doi:10.1093/pq/pqu036.
978 WASSERMAN, Ryan. 2021. "Material Constitution." in *The Stanford Encyclopedia of*
979 *Philosophy*. Stanford, California: The Metaphysics Research Lab, Center for the
980 Study of Language and Information, https://plato.stanford.edu/archives/fall2021
981 /entries/material-constitution/.

doi: 10.48106/dial.v78.i1.03