Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Animalism, persistence conditions, subjecthood, and aporia

(Chris Hughes)

I

Animalism has been ably defended by (at least) Peter van Inwagen, Eric Olson, and Paul Snowdon, and has many attractions. Indeed, when I read *The Human Animal*, my initial reaction was something like: why isn’t everyone an animalist? In what follows, I shall set out some of the reasons I am less surprised than I once was about the controversiality of animalism.

II

What exactly animalism is, is an interesting and delicate question. But for our purposes, it will be enough to say that (as I understand it) animalism will be a view that has the following consequences:

(A) Necessarily, human persons are essentially human animals.
(B) Necessarily, animals (and hence human animals) have biological persistence conditions, and go on only as long as their biological lives continue.
(C) We are human persons.

Why think of animalism as implying all these views? Why not say, instead, that animalism is just the view that human persons are animals? For various reasons, some of which will emerge later in this paper. But for now, suppose someone had the following view:

Human persons are currently human animals, and indeed human persons are permanently human animals. Be that as it may, a human person has both a material body and an immaterial mind. And if a human being were
‘pared down’ to its immaterial mind, it would go on existing (in virtue of retaining its immaterial mind), but would cease to be a human animal, or for that matter an animal, (in virtue of the fact that embodiment and materiality are required for something to be an animal (though not for something that is (in fact, currently) an animal to exist)).

On this view, each human person is identical to some human animal (in the same way that each human being is identical to some resident of earth (at least, if only human beings that existed in the past or present or included in the domain of quantification). Even so, the view just described is surely not animalism in the usual sense: animalists (from Thomas Aquinas right through to van Inwagen, Olson, and Snowdon) deny that human persons are, but only accidentally are, human animals, in the way that (past or present) human persons are, but only accidentally are, terrestrial residents. Moreover, as well as holding that human persons are essentially human animals, animalists hold that human persons, qua human persons, are essentially human animals. Consider the statement: every (material) human artefact is visible to the naked eye. At one time, certainly, all the (material) human artefacts there were visible to the naked eye. And if history had gone differently, and human beings had become extinct before they ever became technologically advanced enough to make (material) artefacts that are too small to be seen with the naked eye, then it would have been true that all the (material) human artefacts there ever have been or ever will be are visible to the naked eye. More than that, I imagine it would under those circumstances be true that every (material) human artefact there ever was or ever will be is essentially visible to the naked eye. (Consider, say an ordinary table, or button. Surely that very artefact could not have been too small to be seen by the naked eye). Still, even in the circumstances envisioned, it would not be true that (material) human artefacts, qua (material) human artefacts, are visible, or essentially visible, to the naked eye, because it would not be true that being a (material) human artefact implied being visible to the naked eye or being essentially visible to the naked eye. Thus, even in the envisioned circumstances, it would not be true that necessarily, all (material) human artefacts are essentially visible. I take it that animalists want to say:

It’s not as though — as it happens — all the human persons there actually are (all the human persons there ever actually have been or ever actually will be) are essentially human animals, although there might have been
human persons that were not essentially human animals (in much the way that, perhaps, all the art works there ever actually have been or ever actually will be essentially do not coincide with any thing which is both a sphere with a diameter of one mile, and made of pure gold). It is instead that human persons, as such, are essentially human animals: *being a human person implies being essentially a human animal*.

That is why I’ve said that animalism implies (A), and not just the much weaker claim that every human person is (identical to) a human animal. But, although I take it that you cannot be an animalist without accepting (A), you could accept (A) without being an animalist, if you had the following view:

Human persons are purely immaterial beings. Also, human persons are not accidentally but essentially human persons. Moreover, human persons are essentially capable of being embodied by a human body. And necessarily, whatever is possibly embodied by a human being eo ipso actually a human animal (in the same way that anything that could possibly die is eo ipso actually mortal). So even though human persons are purely immaterial beings, it is necessarily true that human persons are essentially human animals, because necessarily, whoever is a human person will be capable of embodiment by a human body in any world in which he or she exists, and so will be a human animal in any world in which he or she exists.

Animalists want to exclude that a human person could be a human animal simply in virtue of his or her capabilities (or simply in virtue of his or her history): for that would allow someone to grant (A), but insist that human persons can go on existing, and indeed go on being animals, without the right sort of biological events going on (say, without a certain living human body continuing to be a living human body). And (B) rules out this last possibility.

Suppose, now, that a philosopher held that (necessarily) a human body = a human person = a human animal, and that a human body/person/animal has biological persistence conditions. Suppose, though, that this philosopher also maintained that what we are, are the immaterial Cartesian minds that ‘control’ (and more generally are causally related in the right way to) human bodies/persons/animals. This view is compatible with the conjunction of (A) and (B), but surely is incompatible with animalism: as Snowdon emphasizes, animalism is a
thesis about what we human person are (not just about what human persons are). Hence (C).

III

A number of philosophers, from Descartes to Richard Swinburne, have argued from the premiss that our surviving disembodiment is (metaphysically) possible to the conclusion that we do not have a purely material constitution. One could also move from that premiss to the conclusion that animalism is false, via (what I shall call) Argument 1:

- A human person’s surviving disembodiment (that is, still existing, despite now being disembodied) is metaphysically possible.
- A human animal’s retaining its human animality despite disembodiment (that is, still being a human animal, despite now being disembodied), is metaphysically impossible. [Since human animality implies embodiment (that is to say, it cannot be true – at a time, in a possible world – that an individual is a human animal, unless it is true – at that time, in the world – that that individual is embodied).]
- So it is not necessary that human persons are essentially human animals: it is at least possible that there are human persons who either are only accidentally human animals, or are not human animals at all. In which case (A) is false; in which case animalism is false.

The core of Argument 1 has the following form:

- It is (metaphysically) possible for a (kind of thing) $K$ to still exist, despite having ceased to be $F$.
- It is (metaphysically) impossible for a $K'$ to still be a $K'$, despite having ceased to be $F$.
- So, it is not (metaphysically) necessarily that all $K$s are essentially $K$'s – i.e. it is at least (metaphysically) possible for a $K$ to be either only accidentally a $K'$, or not a $K'$ at all.

Since the above form is (logically) valid, so is Argument 1. The animalist will accordingly have to deny that it is (logically) sound. I take it that the
animalist will want to endorse the second premiss. After all, if a human animal could retain its human animality in a disembodied state, then—even assuming (A) — it’s hard to see how we’ll get (B): if human animals could go on existing without going on being embodied, how could it be that human animals could not go on existing, unless their biological life went on?

So it appears that the animalist will have to (and will) take issue with the argument’s first premiss — that is, the claim that it is (metaphysically) possible that a human person’s surviving disembodiment (that is, still existing, despite now being disembodied) is (metaphysically) possible. Why, an animalist might say, should anyone accept that premiss?

Well, the anti-animalist might reply that a human person’s surviving disembodiment looks possible. Consider Argument 2:

- An oak tree’s surviving (complete) “dematerialization” is (metaphysically) possible. (A thing (completely) dematerializes if it loses all its material parts.)
- A plant’s retaining its plantthood despite (complete) dematerialization is (metaphysically) impossible (since plantthood implies materiality).

So it is not necessary that oak trees are essentially plants; it is at least possible that there are oak trees that either are only accidentally plants, or are not plants at all. So “plantism” (about oak trees) is false.

No one would endorse Argument 2, because its first premiss does not appear to be true: an oak tree’s surviving (complete) “dematerialization” does not look possible. By contrast, various philosophers, from Descartes to Swinburne would endorse the modal argument against animalism under consideration, because its first premiss appears to be true. To be sure, the fact that something looks possible is not a conclusive reason for thinking that it is possible. But it is a defeasible reason for thinking that it is possible. In the absence of reasons to deny or at least doubt that it is possible, it is reasonable to accept that it not only looks but is possible. And in the case at issue, there are no reasons to deny or doubt that the premiss (a human person’s surviving disembodiment...is (metaphysically) possible) is true.

At this point, a defender of animalism might naturally object that we need to distinguish the question of whether it is possible for a human
person to survive disembodiment from the question of whether it is *metaphysically* possible for a human person to survive disembodiment. There are a number of ways the defender of animalism might develop this point.

Suppose, for example, that she accepts the sort of ‘two-dimensional’ account of modality articulated and defended by David Chalmers. On Chalmers’ account of modality, we depart from the Carnapian idea that both terms and sentences have a unique intension. Instead, we suppose that terms and sentences have both a primary and a secondary intension. Oversimplifying in ways I hope will be immaterial to our concerns (e.g. neglecting the distinction between uncentered and centered worlds), the primary intension of the term *water* picks out, at a given possible world \( w \), the clear drinkable liquid in the rivers, lakes, and oceans of \( w \) (whether that liquid is \( \text{H}_2\text{O} \), or, say, Putnam’s XYZ); by contrast the secondary intension of ‘water’ picks out at a given possible world \( w \), what the primary intension of the term ‘water’ picks out in the actual world—which is to say, \( \text{H}_2\text{O} \). Analogously, the primary intension of the sentence *water often freezes* is the set of possible worlds \( w \) such that the substance picked out by the primary intension of *water* in \( w \) is an element of the set picked out by the primary intension of *often freezes* in \( w \). Hence the primary intension of *water often freezes* will include possible worlds in which there is no \( \text{H}_2\text{O} \), and XYZ is the clear, drinkable liquid in the rivers lakes and oceans, and XYZ often freezes. The secondary intension of the sentence *water often freezes* is the set of possible worlds \( w \) such that the substance picked out by the secondary intension of *water* in \( w \) is an element of the set picked out by the secondary intension of *often freezes* in \( w \). Hence the secondary intension of *water often freezes* will not include any \( \text{H}_2\text{O} \)-less possible world \( w \), even if the primary intension of *water* picks out something in \( w \) that is an element of the set picked out by the primary intension of *often freezes* in \( w \). A sentence of the form *water is F* will be *primarily necessary* just in case every possible world is an element of its primary intension, and *secondarily necessary* just in case every possible world is an element of its secondary intension. Chalmers calls primary necessity *epistemic necessity* (though to my mind, *conceptual necessity* might be a more apt term) and identifies epistemic necessity with a priori knowability; he calls secondary necessity *metaphysical necessity*. Thus *water is extended* will be epistemically as well as metaphysically necessary, but *water is non-homeomerous* will be epistemically contingent, though metaphysically necessary.
An animalist who thinks of modality in the way Chalmers does might challenge argument 1 as follows:

For sure, a human person’s surviving disembodiment looks possible. Indeed, a human person’s surviving disembodiment is possible. That is to say, it is primarily or epistemically possible (unlike a tree’s surviving dematerialization). After all, it cannot be excluded a priori that human persons are purely immaterial Cartesian beings that are embodied as long as, and only as long as, they stand in the right causal relations to some body. But why suppose that a human person’s surviving disembodiment is secondarily or metaphysically possible? As Locke emphasized long ago in his discussion of the possibility of “thinking matter”, we cannot exclude a priori that the beings human person refers to in the actual world are purely material beings. And if the beings that human person refers to are purely material beings, then there is no obvious reason to think that the secondary intension of human person picks out any individual in any possible world that survives disembodiment. Hence there is no obvious reason to think it is secondarily or metaphysically possible for a human person to survive disembodiment.

Naturally, an animalist does not need to find Chalmers’ (controversial) “two-dimensionalist” account of modality congenial. But even an animalist who does not sign up to two-dimensionalism about modality can still distinguish what cannot be excluded a priori (and (a priori) looks possible) from what is in fact metaphysically possible, and argue that we cannot move from the non-a-priori-excludability of human persons survive disembodiment to its metaphysical possibility, any more than we can move from the non-a-priori-excludability of water is homeomerous to its metaphysical possibility. Just as water’s being homeomerous is metaphysically impossible, despite its non-a-priori-excludability, given what water actually refers to, a human person’s surviving disembodiment may well be metaphysically impossible, despite its non-a-priori-excludability, given what human person actually refers to.

To see how the proponent of argument 1 might respond to this objection, it will be helpful to consider Argument 3:

- A human person’s surviving (simultaneous) disembodiment and “dismindment” (that is, still existing, despite no longer having a body or a mind) is metaphysically possible.
A human animal’s retaining its human animality despite (simultaneous) disembodiment and “dismindment” (that is, still being a human animal, despite no longer having a body or a mind), is metaphysically impossible - since human animality implies having a body or a mind (that is to say, it can not be true (at a time, in a possible world) that an individual is a human animal, unless it is true (at that time, in that world) that that individual has a body or a mind).

So it is not necessary that human persons are essentially human animals: it is at least possible that there are human persons that either are only accidentally human animals, or are not human animals at all. In which case (A) is false; in which case animalism is false.

Argument 3, unlike argument 2, would not be endorsed by dualists such as Descartes or Swinburne, and is manifestly unconvincing, inasmuch as a human person’s existing in a disembodied and disminded state does not look possible. We can make sense of the idea of a human person’s still existing, but no longer having a body. And arguably (as animalists suppose) we can make sense of the idea of a human person still existing, but no longer (or not yet) having a mind. What I – and I take it we – have difficulty making sense of, is the idea of a human person’s still existing, despite being both bodiless and mindless (or the idea of a human person now in existence having already existed, despite then being both bodiless and mindless).

There is a more general point here. If we think that there are individuals that have neither “bodily” characteristics (in a broad enough sense of “bodily” to allow that non-biological as well as biological beings have bodily characteristics), nor mental characteristics, then we will obviously have no trouble with the idea that individuals of that kind can exist in a mindless and bodiless state: a mathematical Platonist will think it’s true, and a fortiori metaphysically possible, that the number π exists without having either a body or a mind. But, so far as I can see, if we think of individuals as having either bodily or mental characteristics, we have great difficulty thinking of individuals of that kind existing in a bodiless and mindless state. Incidentally, I think this comes into the explanation of why we find the argument against plantism from possible dematerialization hopeless. We think of an oak tree as unlike π, inasmuch as it has bodily characteristics. But we think of an oak tree as like π, inasmuch as it lacks—and, I am inclined to say, essentially lacks—
mental characteristics. According to ancient mythologies, oak trees were “inhabited by” things with minds (things that had or perhaps were minds)—dryads. And I think I can conceive of trees being inhabited by dryads. But it’s the dryad inhabiting the tree, and not the tree itself, that would have the mind (and the mental characteristics): I can’t make anything of the idea of a tree’s having a mind.

If, however,

(a) whatever has either bodily or mental characteristics, could not exist without having either bodily or mental characteristics
(b) a tree essentially lacks mental characteristics,
then
(c) a tree could not survive (complete) dematerialization, given that
(d) (complete) dematerialization implies the loss of all bodily characteristics.

To return to our main point: if we can make sense of the idea of a human person’s existing in a disembodied state, that is only because we can make sense of the idea of a human person’s existing in a minded-though-disembodied state.

And if we can make sense of the idea of a human person’s existing in a minded-though-disembodied state, then we can make sense of the idea of a human person’s existing, not just without bodily parts, but also without any “extra-mental” parts at all—that is, without any parts that are disjoint from (do not overlap with) her mind. After all, it doesn’t look as though human persons have some essential part which is neither mental nor bodily: as we think of it, there is “nothing more to” a human person than her body and her mind. So someone who thinks that a human person’s existing without bodily parts is metaphysically possible, should and will likewise think that a human person’s existing without extra-mental parts is metaphysically possible.

Now anyone who thinks it is metaphysically possible for a human person to exist without extra-mental parts is in a position to offer Argument 4 against animalism:
A human person’s still existing, despite not having extra-mental parts, and not inhabiting anything with extra-mental parts, is metaphysically possible.

A human animal’s retaining its human animality despite not having extra-mental parts, and not inhabiting anything with extra-mental parts, is metaphysically impossible.

So it is not metaphysically necessary that human persons are essentially human animals: it is at least metaphysically possible that there are human persons that either are only accidentally animals, or are not animals at all. In which case (A) is false; in which case animalism is false.

What is the relation between argument 4 and argument 1? This depends on how we understand what it is for a human person to be embodied. For the sake of brevity, let us say that a human person has something if that human person either has that thing (as a part), or inhabits something that has that thing (as a part). On a stronger understanding of what it is for a human person to be embodied, a human person is embodied at a time in a possible world only if she has a (whole) (biological) body at that time in that world. On a weaker understanding what it is for a human to be embodied, a human person is embodied at a time in a possible world only if she has some “bodily” (i.e. material) parts at that time in that world. (In modal arguments against materialism from possible disembodiment, embodiment needs to be understood in the weaker way, since a purely material being needn’t be embodied, in the strong sense of “embodied”).

If we understand embodiment in the first and stronger way, it seems that not having extra-mental parts implies disembodiment: at least, I don’t see how an individual who neither has extra-mental parts, nor inhabits anything with extra-mental parts, could either have a (whole) (biological) body as a part, or inhabit a (whole) (biological) body. Assuming that not having extra-mental parts implies being disembodied, we may conclude that if argument 4 is a sound argument against animalism, so too is argument 1. If on the other hand, we understand embodiment in the second and weaker way, there is no obvious reason to suppose that the soundness of argument 4 would imply the soundness of argument 1. Perhaps there is a metaphysically possible scenario in which (a) a human person loses all her extra-mental parts, (b) the
person’s mind continues to exist, and continues to be “involved in” thoughts and experiences, and (c) the human person herself continues to exist. (a) – (c) could all be true, even if, given what human person actually refers to, it is not metaphysically possible for a human person to exist, without having “bodily” or material parts. If that is how things are, argument 4 against animalism will be sound (assuming (as it seems safe to assume) there could not be a human animal that is as it were “all mind”), but argument 1 will be unsound (since human persons will be essentially embodied (in the weak sense of embodiment).

We noted earlier that an animalist will challenge the idea that a human person’s surviving disembodiment is secondarily or metaphysically possible, as well as primarily possible (non-excludable a priori). In particular, an animalist might say that, absent a commitment to something like Cartesian dualism, there is no reason to suppose think that, given the individuals that human person actually refers to, it is primarily or metaphysically possible for a human person to survive disembodiment.

We can now see that the proponent of argument 1 can respond to this challenge in either of two ways, depending on whether she has a stronger or weaker conception of embodiment. If she has the stronger conception of embodiment, she can say that, even assuming that the individuals human person actually refers to are purely material beings, there is still a reason to suppose that a human person’s surviving disembodiment is metaphysically possible: the reason is that it is metaphysically possible for a human person’s to go on existing, despite ceasing to have* any extra-mental parts, and for a human person to go on existing, despite ceasing to have* any extra-mental parts, is for that human person to survive disembodiment. If on the other hand, the proponent of argument 1 has the weaker conception of embodiment, she can concede to the animalist that argument 1 is unconvincing, inasmuch as a human person’s surviving disembodiment might or might not be metaphysically possible (depending on what sort of individuals human person actually refers to), but fall back on argument 4. For as we have seen—irrespective of whether the proponent of argument 1 has a stronger or a weaker conception of embodiment—she believes it is metaphysically possible for a human person to survive disembodiment, only because she believes it is possible for a human person to exist in a minded-though-disembodied state, and if she believes it is possible for a human person to exist in a minded-though-disembodied state, she
should and will believe that it is metaphysically possible for a human person to exist without having* any extra-mental parts. If, however, it is metaphysically possible for a human person to exist without having* any extra-mental parts, then argument 4 is sound.

Moral: proponents of argument 1, independently of how they construe embodiment, and independently of their views on whether human beings are wholly immaterial, only partly immaterial, or wholly material, can endorse argument 4.

I used to think that some and perhaps many anti-animalists rejected animalism, because—under the influence of certain conceptions of the afterlife or Meditation 6— they held that human persons will or at any rate could continue to exist without continuing to have* material parts (and they saw the incompatibility of this view with animalism). But I now think this gets things backwards. Those who think that a human person’s continuing to exist without continuing to have* material parts is (metaphysically) possible, think that, only because they think that a human person’s continuing to exist without continuing to have* extra-mental parts is (metaphysically) possible—and they hold this last (animalism-excluding) view, for reasons independent of whether a human person could go on existing, even though she no longer had* material parts. If, like Aquinas, they rejected the idea that a human person could continue to exist without continuing to have* extra-mental parts, they would, like Aquinas, reject the idea that it is metaphysically possible for a human person to go on existing, without continuing to have* material parts—even if, like Aquinas, they thought that a human person’s mind was an immaterial entity that could and indeed would survive the destruction of that human person’s body.

IV

Returning to our main concern, how might the animalist oppose argument 4? Again, I take it that the animalist does not have the option of taking issue with its second premiss: a view on which human persons are human animals, but a human person can survive as (something coincident with) an immaterial mind, or as (something coincident with) the ‘thought-involved’ part of a human brain , and a human animal can hold on to its animality, despite now coinciding with an immaterial mind, or the ‘thought-involved’ part of a human brain, is not worthy of
the name “animalism”. Thus the animalist will need to challenge the idea that a human person’s existing without having* extra-mental parts is metaphysically possible. How might this go? Argument 4’s first premiss is true as long as

(i) It is metaphysically possible for a human person’s mind to continue to exist, despite that human person’s no longer having* any extra-mental parts (either because that human person no longer exists (then), or because that human person still exists, but no longer has* any extra-mental parts (then)).

And

(ii) a human person’s mind’s continuing to exist, despite that human person’s no longer having* any extra-mental parts is metaphysically compossible with that human person’s continuing to exist.

So the animalist can concede (i), but refuse to concede (ii), or refuse to concede either (i) or (ii). (Conceding (ii) and refusing to concede (i) is not an option, since (ii) implies (i).)

The animalist could dig in her heels at (i), and say that, even if we cannot exclude a priori that a human person’s mind will continue to exist, despite that human person’s no longer having* any extra-mental parts, we have no reason to think that, given what human person’s mind actually refers to, it is metaphysically possible for a human person’s mind to continue to exist, despite that human person’s no longer having* any extra-mental parts. Perhaps, if we had a better understanding of (the nature of) what human person’s mind actually refers to, we would see that a human person’s mind could only exist as a (proper) part of the human person to whom it belongs.

I doubt, though, that animalists are well advised to oppose Argument 4 simply by casting doubt on its first premiss. Given that there are no uncontroversially actual cases of a human person’s mind continuing to exist, even though the human person who had it no longer has* any extra-mental parts, one might wonder how we can know that such a thing is metaphysically possible, as well as non-excludable a priori. Be that as it may, it’s hard to see what reason we have to think it is metaphysically impossible for a human person’s mind to continue to exist, even though the person who had it no longer has* any extra-mental properties. If the animalist has good reasons to doubt the truth of the first premiss of Argument 4, but lacks good reasons to deny it,
then it would seem that the animalist lacks good reason to deny the soundness of argument 4—unless she has good reason to deny that Argument 4’s second premiss is true. And if the animalist lacks good reason to deny the soundness of argument 4, the animalist lacks good reasons to be an animalist.
So it seems as though the animalist will need reasons (not simply to doubt but) to deny (ii). What might such reasons be?
Consider the following (Cartesian) argument for dualism—henceforth, Argument 5:

- Human persons could make it into the (immediate) future, despite having no material parts in the (immediate) future.
- Nothing wholly material could make it into the (immediate) future, despite having no material parts in the (immediate) future.

--------------------------
- Human persons are not wholly material beings (they are either partly or wholly immaterial).

A materialist—or at least, a materialist who is an animalist—might naturally respond to this argument as follows:

A human person’s making into the (immediate future), despite not having material parts in the (immediate) future is not excludable a priori, and hence primarily possible. But we have reason to think it is not secondarily or metaphysically possible. What reasons? Well, there’s nothing more to a human person than a body and a mind. A human person’s body is manifestly a wholly material thing. But the systematic correlations between mental states and cerebro-cortical states show that a human person’s mind—the part of that person she thinks with—is, or at any rate, shares a decomposition with, a part of her (wholly material) brain. So a human person’s mind, like her body is wholly material—in which case a human person is wholly material. Moreover (just as the proponent of Argument 5 maintains), nothing wholly material could make it into the (immediate) future, despite having no material parts (then). Hence, the sort of beings that human person actually refers to (wholly material beings) could not make it into the (immediate) future, despite having no material parts in the (immediate) future.

This response to Argument 5 is an instance of what William Rowe called a Moore shift. In a Moore shift, the defender of not-r concedes the validity
of the argument, \( p, q, \text{ therefore } r \), but denies its soundness, on the grounds that \( \text{not}-r, q, \text{ therefore not}-p \) is sound. (In the case at hand, \( p = \text{human persons could make it into the (immediate) future, despite having no material parts (then)}, \ q = \text{nothing wholly material could make it into the (immediate) future, despite having no material parts (then)}, \) and \( r = \text{human persons are not wholly material beings}. \)

The animalist can analogously concede the validity of Argument 4, but deny its soundness on the grounds that the argument that is sound is (not Argument 4 but rather) Argument 6:

- It is (metaphysically) necessary that human persons are essentially human animals.
- A human animal’s retaining its human animality despite not having* extra-mental parts, is metaphysically\(\square\)possible.
- A human person’s still existing, despite not having* extra-mental parts, is metaphysically\(\square\)impossible.

If the animalist makes this move, the anti-animalist will naturally protest that the issue under debate (or rather, given that animalism implies (B) and (C) as well as (A), one of the issues under debate) is whether it is in fact (metaphysically) necessary that human persons are essentially human animals: why should we think it is? I take it the animalist will say that if all human persons are essentially human animals is true, it is also metaphysically necessary (so that all human persons are essentially human animals is like all water is composed of molecules, and unlike all works of art are essentially either not made of pure gold or not spheres with a diameter of a mile). On this assumption, the animalist will be able to motivate acceptance of the first premiss of argument 6, as long as she can motivate acceptance of human persons are essentially human animals. But how is she going to motivate acceptance of this last claim?

Suppose it is evident, even before we start doing philosophy, that human persons are human animals, and it becomes evident, upon philosophical reflection, that whoever is a human animal, is essentially a human animal. Then, (on the assumption that human animals are essentially human animals is (metaphysically) necessary if true), we may conclude that Argument 6, rather than Argument 4, is sound.
Some animalists think that it is in fact pre-philosophically uncontentious that human persons are human animals—or at least, that it is pre-philosophically uncontentious that, if human persons are wholly material beings, then they are human animals:

We can understand how someone might deny we are...animals by rejecting materialism. Many philosophers have argued that we are not material objects of any kind, but immaterial objects, or abstract objects akin to computer programs, or the like; or that we have some immaterial part. But once it is conceded that we are material beings of some sort, it seems quite obvious what sort of material beings we are: we are living animals. . . . [T]hat we are members of the species *Homo Sapiens* is certainly something we ordinarily assume when we are not doing philosophy. (E. Olson, *The Human Animal*, p. 95).

But I am not so sure that when we are not doing philosophy, we take it for granted that we (human persons) are (human) animals. When we are “outside the study” (whether temporarily, in the case of philosophers, or permanently, in the case of non-philosophers), we certainly take it for granted that we are human *persons*, and human *beings*, but it seems much more doubtful that we take it for granted that we are human animals. To start with, I think that many non-philosophers regard it as obvious that animals are wholly material beings (share a decomposition with their (wholly material) body), but think it is an open question whether human persons are wholly material. And I find it hard to believe that all the non-philosophers who take it for granted that animals are wholly material beings, but do not take it for granted that human persons are wholly material beings, think it for granted that human persons are human animals. Also, so far as I can see, even a non-philosopher who takes it for granted that human persons are wholly material beings needn’t take it for granted that human persons are human animals. For a non-philosopher (as much as a philosopher) might (a) be confident that human persons are wholly material beings, (b) be confident that that being a person is being one kind of thing, and being an animal is being a different kind of thing, but (c) not be confident that one and the same individual could simultaneously be each of those two different kinds of things. So far as I can see, not all materialist non-philosophers are confident that human beings are in the intersection of the extension of *person* with the extension of *animal*, because not all
materialist non-philosophers are confident that the intersection of those extensions in non-empty. An animalist might say here that the problem here is that non-philosophers typically use the word “animal” to mean “non-human-animal”. If non-philosophers fail to distinguish “animal” from “non-human animal”, then they obviously won’t be happy with the idea that human beings are in the intersection of the extension of person with the extension of animal. Similarly, an animalist might say, non-philosophers who confidently assent to *animals are wholly material beings* but do not assent to *human persons are wholly material beings* are failing to distinguish, or failing to clearly distinguish *animal* and *non-human animal*.

Well, I don’t know. I doubt that, if you asked a non-philosopher to define *animal*, she would necessarily bring non-humanity into the definition (my (admittedly uninformed) guess is that most non-philosophers would say that animals are living beings that can do certain things (that, say, plants cannot).

What is true, I think, is that, with the right sort of priming, at least a great many non-philosophers can be led to affirm that human persons are human animals. The priming might go: what kind of thing are you? Are you, say, a human being? And what are human beings? Are they, say, a kind of animal? If a philosophy professor primes her (first-year) undergraduates this way, I think she will find that at least a good number of her students will affirm that they are human animals. Once an undergraduate has affirmed that he is a human animal, the professor can ask him whether he is a human person, and he’ll no doubt answer affirmatively. On this basis, the professor will be able to get the student to say that the human person that he is, is a human animal, and then generalize, and affirm that (we) human persons are human animals. By contrast, if the professor starts off by asking: what kind of thing are you? Are, you say, a person, and in particular a human person? And what kind of thing is a human person? Is a human person a kind of animal—a human animal?

Then, I think, the professor will have a harder time getting her students to affirm that they are human animals as well as human persons, or that (we) human persons are human animals. The question, *are human persons a kind of animal?* will trigger (in non-philosophers) more reluctance to assent than the question *are human beings a kind of animal?* (An anti-
animalist might suspect that this is because human being is ambiguous between human person and human animal.)
In short, I think it is an oversimplification to say that non-philosophers, or even materialist non-philosophers, regard it as “quite obvious” that we (human persons) are animals, and in particular human animals. (In fairness to Olson, if I have understood the rest of the chapter of The Human Animal from which I took the quote above, he would not disagree).

Of course, the question of whether human persons are human animals is obvious, even before we start doing philosophy, is of no great moment, if it can be made evidently true, once we do philosophy. And animalists have philosophical arguments that purport to make it at least reasonably obvious that human persons are (identical to) human animals. In what follows, I shall concentrate on what seems to me one of the most promising candidates for the “kernel” of this sort of argument.—Olson’s (very neat) “thinking animal” argument. The gist of this argument is:

(1) Some animals think. In particular, the human animal with my body thinks.
(2) The human person with my body thinks.
(3) At most one individual with my body thinks.
(4) So the human person with my body = the human animal with my body.

As Olson would recognize, (4) falls short of each human person is (identical to) a human animal. Still, we might think that there is a reasonably uncontroversial line of reasoning that will take us from (4) to the identity of each human person with some human animal. To start with, it would on the face of it be surprising if some human persons (e.g. the author of this piece and the reader of this piece) were (identical to) human animals, but other human persons were not: one would have expected the class of human persons (as opposed, say, to the class of persons) not to be heterogeneous in that way. Moreover, it is at least arguable that if (4) is true, so is:

(5) Any thinking human person with a body is identical to the human animal with that body.
After all, it seems plausible that any thinking human person with a body will be a thinking human person with a human body (what would make a person with a non-human body a human person?) And it seems plausible that whenever there is a human person HP with a human body, there is also a human animal HA with that human body. Assuming the thinking animal argument is sound, it seems that HA will think, and be identical to HP (lest HA and HP be two different thinkers with the same body).

Also, it is at least arguable that any thinking human person is a thinking human person with a body (again, what would make a thinking person without a body human?) If this is right, we could move from (5) to

(6) Any thinking human person is identical to the human animal with that person’s body.

Suppose we say that, in order to be a “thinking human person”, a human person doesn’t have to be (“occasionally”) thinking, but only needs to be “a thinker” -- in Cartesian parlance “a thinking thing”. Then it is at least arguable that every human person is (in that sense of “thinking human person”) a thinking human person (that is, a thinker or “thinking thing”). If, however, every human person is a thinking human person, and every thinking human person is identical to some human animal (to wit, the animal with the body of that human person), it seems we may conclude that

(7) Each human person is identical to some human animal.

So, the animalist can say, perhaps human persons are (identical to) human animals is not pre-philosophically obvious. But (1) – (3) are pre-philosophically obvious. (The canonical non-philosopher on the Clapham omnibus has no doubt that (“prototypical”) members of the species canis familiaris and (“prototypical”) members of the species homo sapiens think, and no doubt that a human person with her body thinks, and no doubt that there is only one thinker with her body). With just a little philosophical reflection upon (1) – (3), we can make it obvious that (4) is true. As we have seen, (4) is not the claim that each human person is identical to some human animal, but it is not a million miles from it, either; perhaps with more philosophical reflection (1) – (4), we can make it obvious that each human person is identical to some human animal
(as was suggested above). And perhaps, with a bit more philosophical reflection we can make the truth of (the rest of) animalism obvious. In fact, though, it’s not so clear to me that since (1) – (3) are all obviously true and obviously jointly imply (4), (4) is obviously true as well. True enough, even those (or perhaps especially those) who have not done much philosophical reflecting are strongly inclined to assent to each of (1) – (3), and the path from (1) – (3) to (4) is short and straight. But I think this provides rather less in the way of support for (4) than we might have thought. To see this, consider Argument 7:

(1’) A sweater is a material object, and a thread is a material object.
(2’) It never happens that two different material objects are in the very same place at the very same time.
(3’) So a sweater knitted from one (very long) thread and the (very long) thread that the sweater was knitted from are the same material object (since they are manifestly in the same place at the same time).

(Judging from my experience with sixth-form students and first-year undergraduates) Clapham bus-riders are (by and large) strongly inclined to assent to both (1’) and (2’), and (1’) and (2’) jointly imply (3’). But this provides significantly less support for (3’) than one might have thought, given that Clapham bus-riders are also (by and large) strongly inclined to assent to:

(4’) Only different material objects can have different makers, histories, or potentialities.

And

(5’) The sweater knit from the (very long) thread have different makers, different histories, and different potentialities.

--and (4’) and (5’) jointly imply that the sweater and the (very long) thread it was not knit from are not the same material object. Non-philosophers—or as we might call them, ‘the folk’--are strongly inclined to assent to things that (jointly) imply that the sweater and the (very long) thread it was knit from are the very same material object, and also strongly inclined to assent to (other) things that (jointly) imply that the sweater and the (very long) thread it was knit from are not the very same material object. In light of this, the fact that the folk are strongly inclined to assent to things that (jointly) imply the identity of the thread
with the sweater is far from conclusive evidence in favour of the identity of the thread with the sweater. (For this reason, (to vary the example), a professor who is an able “primer” can get a good many undergraduates to either affirm or deny that a paper airplane = the sheet of paper it “coincides with”, depending on which “primer questions” she chooses). Similarly, it might be that the folk are strongly inclined to accept things that imply that human persons are (identical to) human animals, and also strongly inclined to accept (other) things that imply that human persons are not (identical to) human animals. Why suppose that? Remembering that strong embodiment implies having* a (whole) (biological) body, and weak embodiment only implies having* some material parts, I shall say that things that continue to exist, despite no longer having a (whole) (biological) body, survive weak disembodiment, and that things that continue to exist, despite no longer having any material parts at all, survive strong disembodiment. It is at least arguable that the folk—or at least a great many of the folk—think that human persons could survive weak disembodiment (assuming their mind survived, and continued to “work”) but human animals could not (and think that whenever Ks could survive something that K*’s could not, Ks ≠ K*’s). (Compare: it is at least arguable that the folk—or at least a great many of the folk—think that a very long thread that currently coincides with a knitted sweater could survive being completely “unknit”, but the sweater currently knit from that thread could not (and think that whenever Ks could survive something that K*’s could not, Ks ≠ K*’s).)

The idea here is that, as the folk see it, just as there is nothing impossible in the idea of a non-human person’s coinciding with and sharing a decomposition with a non-human mind, there is nothing impossible in the idea of a human person’s coinciding with and sharing a decomposition with her human mind. Just as the archangel Gabriel might (permanently) coincide with and share a decomposition with his angelic mind, a human person might (temporarily) coincide with and share a decomposition with her human mind. For the folk, persons as such are the kind of things that could coincide and share a decomposition with their mind, in the same way that kitchen knives, as such, are the kind of things that could coincide with their (largest) central part. (Imagine a kitchen knife consisting of a central metal or ceramic part, and two wooden or plastic handles. The knife will “outvolume” its central part; but if the knife loses it handles, it will
shrink into coincidence and “co-decomposability” with its (largest) central part.)

Here a champion of animalism might object that we must be careful not to over-attribute beliefs to the folk. The folk no doubt believe that a human person’s coinciding with and being co-decomposable with her mind is not something that can be excluded a priori. So they believe it is primarily possible, or “epistemically possible”, in Chalmer’s sense of “epistemically possible”. They may also believe that it is epistemically possible in the more straightforward sense of the term. That is, they may believe that we cannot exclude a human person’s mind coinciding with and being co-decomposable with her mind (at some (possibly future) time) on either a priori or a posteriori grounds. But why should we suppose that they think it is secondarily or metaphysically possible? In other words, why should we suppose that the folk judge that some human person has coincided with and been de-composable with or will coincide with and be co-decomposable her mind is not only not excludable a priori, and for all we know true, but is also genuinely possible, given what human person and human mind actually refer to? The question is crucial to the issue under discussion, because if we cannot attribute to the folk the belief that a human person’s coinciding with and being co-decomposable with her mind is secondarily possible, there is no obvious reason to think that, although some things the folk take for granted (to wit, (1) – (3)) imply that the person with their body = the animal with their body, other things the folk take for granted (concerning the persistence conditions of human persons and the persistence conditions of human animals) imply that the person with their body ≠ the animal with their body.

The question of which modal beliefs we can attribute to the folk is a large and difficult one. I think (as I take it Kripke thinks, judging from his remarks about the “intuitive content” of essentialist claims, and about the epistemic and metaphysical senses of it could have turned out that) that even if the folk do not talk explicitly about primary and secondary possibility, they do understand what it is for \( p \) to be non-excludable independently of experience, what it is for \( p \) to be non-excludable given everything we know, and what it is for \( p \) to be the sort of thing that could have been be true (given what the words in \( p \) actually refer to). Moreover, I think that the folk have views about which statements \( p \) are in fact non-excludable independently of experience, and about which statements \( p \) are in fact non-excludable, given everything we know, and
about which statements \( p \) could in fact have been be true (given what the words in \( p \) actually refer to).

We can elicit folk views about what is possible (in the various sense of \textit{possible} under discussion) by asking the folk the right questions. If we want to know whether the folk think that water’s existing though hydrogen doesn’t is primarily possible (epistemically possible in Chalmers’s sense) we can ask them whether water’s existing though hydrogen does not could be ruled out independently of experience (if they answer ‘no’, then they think water’s existing though hydrogen does not is primarily possible; if they answer ‘yes’, then they think water’s existing though hydrogen does not is not primarily possible).

If we want to know whether the folk think water’s existing though hydrogen does not is epistemically possible in the (stronger and more natural) sense of “epistemically possible”, we ask them whether, for all they know, water exists, though hydrogen doesn’t (if they answer ‘yes’, then they think water’s existing though hydrogen does not is epistemically possible (in the stronger and more natural sense); if they answer ‘no’ (as presumably some better chemically informed members of the folk will), then they think water’s existing though hydrogen does not is not epistemically possible (in the stronger and more natural sense). Finally, if we want to know whether the folk think that water’s existing though hydrogen did not is secondarily or metaphysically possible, we ask them whether there could have been water without hydrogen, given what \textit{water} and \textit{hydrogen} actually refer to (given the kinds of things water and hydrogen actually are).

Similarly, we can ask the folk whether a human person’s coinciding with and sharing a decomposition with her mind can be ruled out independently of experience, whether it can be ruled out, given everything they know, and whether it could have happened that some human person coincided with her mind, given what \textit{human person} and \textit{mind} actually refer to. My sense is that many if not most of the folk would be happy to say that a human person’s coinciding with and sharing a decomposition with her mind is not just something that cannot be excluded a priori, but also something that might have been so, given what \textit{human person} and \textit{mind} refer to. For my money, the reason that so many philosophers think that a human person’s coinciding with and sharing a decomposition with her mind is secondarily or metaphysically possible, is that they believed that it was secondarily or metaphysically possible when they belonged to the folk, and they didn’t give up that
belief when they became philosophers. In any case, I know of no good reason to suppose that the folk would not want to say that a human person’s coinciding with and sharing a decomposition with her mind is something that might have been so, given what *human person* and *mind* actually refer to. If there is no good reason to suppose that the folk would not regard a human person’s coinciding with and sharing a decomposition with her mind as secondarily or metaphysically possible, then we cannot rule out that some of the things non-philosophers take for granted imply that human persons are (identical to) human animals, and some of the things non-philosophers take for granted that human persons are not (identical to) human animals, and that the “thinking animal” is accordingly no more conclusive than the parallel argument for the identity of the sweater knitted from the one long thread with the one long thread the sweater was knitted from.

At this point, an animalist might say:

Perhaps many if not most of the folk accept things that imply the metaphysical possibility of a human person’s coinciding with and sharing a decomposition with her mind, as well as the metaphysical impossibility of a human animal’s coinciding with and sharing a decomposition with its mind. After all, not a few of the folk accept that they not only could but in fact will someday coincide with and share a decomposition with their mind (in the interval between their death and last-day resurrection, perhaps). These members of the folk accept the truth of some human person coincides with and shares a decomposition with her mind at some time, and truth implies weak epistemic possibility, strong epistemic possibility, and metaphysical possibility. Be that as it may, it is evident that the human person with my body = the human animal with my body is incompatible with human persons have different persistence conditions from human animals (only different things have different persistence conditions). And once the ‘thinking animal argument’ is on the table, we can see that that if we have to choose between the human person with my body = the human animal with my body and human persons have different persistence conditions from human animals, the one to plump for is the thinking human person with my body = the thinking human animal with my body. Surely the conjunction of the human animal with my body is a thinker with there is at most one thinker with my body is more “un-give-up-able” than it is (metaphysically) possible for a human person to coincide with and share a decomposition with her mind.
Why surely? I am happy to admit that I find it hard to doubt the conjunction of human animals as well as human persons may be said to think with there is only one thinker with my body. So I find it hard to doubt that the person with my body = the animal with my body. That said, I don’t find it easy to doubt that it is metaphysically possible for a human mind existing at the present time $t_{pr}$ to exist and “work” (i.e., be involved in thoughts and experiences) at a future time $t_f$ even though the human person who has* extra-mental parts at $t_{pr}$ no longer has* any extra-mental parts at $t_f$, and even though no other human person, and indeed no other subject, has that mind (as a proper part) at $t_f$. Now it might seem that the question of whether a mind currently had* by a human person could exist (and work) “in isolation” at a time when the human person no longer has* extra-mental parts, is independent from the question of whether a human person who now has* extra-mental parts could at a future time coincide with and share a decomposition with her mind. Compare: it seems true that a car engine that currently belongs to a car could exist and work at a future time (exist, and be involved in the right sort of events (controlled explosions taking place in cylinders, etc.)), even though that car engine no longer belonged to that car, or anything else (so that the engine existed “in isolation”), and even though that car no longer had “extra-engine” parts. But it does not seem true that a car that now has all the usual parts could at some future time coincide with and share a decomposition with its engine: a car engine can at most partly constitute a car. (Hence it would be true at the future time that the car no longer had extra-engine parts, not because the car had shrunk into “coincidence and comateriation” with its engine, but because the car no longer existed, and thus no longer had any parts).

In fact, though, as Olson is aware (cf. “A Counterattack” in chapter 5 of The Human Animal) the question of whether a human mind now had* by a human person could exist (and work) in isolation at a (future) time when the human person no longer had* extra-mental parts is linked to the question of whether a human person could at some (future) time coincide with and share a decomposition with her mind in the following way: giving an affirmative answer to the first question and a negative answer to the second has seriously counterintuitive consequences.

I take it that if a mind is “working”—that is, involved in thoughts and experiences—the thoughts and experiences that the mind is involved in must be thoughts and experiences of some subject (no thought or experience is so to speak res nullius). Assume that at some future time $t_f$,
the mind you currently have* is isolated, but still involved in thoughts and experiences, and the human person you are no longer has* any extra-mental parts. If thoughts and experiences necessarily belong to some subject, then the thoughts and experiences the (then) isolated mind is involved in at $t_f$ belong to some subject. Who or what is the subject of those thoughts and experiences? Assuming that a human person cannot coincide with and share a decomposition with her mind, that subject cannot be you: if the subject of the thoughts and experiences an isolated mind is involved were you, then you would coincide and share a decomposition with that isolated mind. So the subject must be someone ≠ you. Is it a “brand new” subject of experience—that is, a subject who didn’t exist when you had* extra-mental parts and your mind was un-isolated? As Olson grants, it seems deeply weird to suppose that making a human person cease to have* extra-mental parts, while keeping the mind of that person existing and working, necessarily involves bringing a new subject—or indeed, anything at all—into existence: the process under discussion seems subtractive, rather than additive.

What are the other options? Well, we could say that the subject of the thoughts and experiences the isolated mind is involved in is something that already existed when you had* extra-mental parts, and your mind became isolated. That thing would presumably be your mind. If we go this route, we can say that making a human person cease to have* extra-mental parts and isolating her mind and keeping it “functioning” need not involve bringing anything new into existence. But then we face the question: when your mind was had* by you, was it (already, even then) a subject of experience? If it wasn’t, then making you cease to have* extra-mental parts, and isolating your mind, and keeping it functional, “subjectified” something that had not previously been a subject: it turned something that didn’t think or experience into something that did. (At least as I read Aquinas, this is his view: he supposes that in statu viatoris the human person with a material body and an immaterial intellective soul is the thinking subject, and the intellective soul is the proper part she thinks with (the proper part “involved” in her thoughts), but when the human person dies, the human person ceases to exist (and thus ceases to have any extra-mental parts) and the immaterial intellective soul continues to exist in isolation (say, in purgatory), and undergoes “subjectification”: a disembodied soul in purgatory, rather than being involved in the thoughts of a new (“death-
born”) subject of experience, is a “freshly subjectified” subject of experience.) This again seems weird. Suppose that your mind has up to now been involved in predominantly enjoyable experiences, and that after you have ceased to have* extra-mental parts, and your mind has been isolated, it will continue to be involved in predominantly enjoyable experiences. Then it seems that your mind “has something to gain” from isolation: it thereby gets to do something good it could never do before—(to wit, have (to be the subject of) enjoyable experiences! Supposing that, in the envisioned circumstances, a (mind that was) a non-thinker and non-experiencer becomes a (mind that is) a thinker and experiencer seems wrong in exactly the same way it seems wrong to suppose that when we cut off Tibbles’ tail, (a bit of matter that was) a non-cat becomes (a bit of matter that is) a cat.

This suggests that perhaps we should say that the now isolated but still working mind not only is a subject of experience now, but was one all along (even before its isolation). That way, we avoid both the weirdness of “subject creation” by a process that don’t look “creative”, and subjectification by isolation. But it seems non-negotiable that, when the now isolated was had* by a human person, that human person was a subject of experience. Assuming that the human person ≠ the mind that is first un-isolated, and subsequently isolated, we end up with the “too many thinkers” problem again: on the face of it, it is no less counterintuitive to suppose that I and my mind are two different thinkers, than it is to suppose that the human person with my body and the human animal with my body are two different thinkers.

We could avoid the too many thinkers problem by insisting that I and my are not different thinkers, but the very same thinker, inasmuch as I just am my mind. If, however, I = my mind, then, irrespective of whether my mind is a Cartesian substance, or a Thomistic immaterial substantial form, or something identical to or at least coincident and comateriate with a part of my brain, it will turn out that I (and other human persons) don’t walk, or weigh more than one hundred pounds, or the like. And it seems no less counterintuitive to deny that some human persons walk, weigh more than one hundred pounds, and so on, than it does to deny that some animals think.

(Suppose on the other hand that you think both animals and minds are the wrong kind of thing to think or experience. Then you can accept that a mind once had* by a human person can continue to be involved in thoughts and experiences, even though it has been isolated, and the
human person that had* it no longer has any extra-mental parts, and reject all the unpalatable options discussed above. You can say: when a human person ceases to have* any extra-mental parts, the human animal that shares a body with that human person goes out of existence. The human person does not go out of existence; instead she shrinks into “coincidence and co-materiation” with a part of her brain, or perhaps shrinks into “coincidence and co-immateriation” with her immaterial intellective soul. Either way, the mind’s undergoing isolation, and the person’s ceasing to have* extra-mental parts, does not bring a new subject of experience into existence, or subjectify something that was previously only potentially a subject. And either way, we have just one thinker (the same thinker), both before and after the brain is isolated, and the human person ceases to have* extra-mental parts).

Let’s take stock. The animalist and the anti-animalist agree that, even before we start doing philosophy, or even if we never do philosophy, we are strongly inclined to accept things that imply the human person with my body = the human animal with my body. The animalist holds that, when we do philosophy, we should accordingly regard it as obvious that the human person with my body = the human animal with my body. A non-animalist will naturally respond that this is too quick, because it arguably is the case that as well as being strongly inclined to accept things that imply that the human person with my body = the human animal with my body, non-philosophers and “pre-philosophers” are also strongly inclined to accept things that imply that it’s not the case that the human person with my body = the human animal with my body (viz. claims about the conditions under which human persons and human animals continue to exist).

In the case of persons and animals, as in the case of threads and sweaters, non-philosophers and “pre-philosophers” are plausibly regarded as inclined to make judgments about the link between coincidence or something akin to coincidence and identity (there are no two material objects in exactly the same place, there are no two thinkers with my body) that are incompatible with the judgments they are inclined to make about the persistence conditions of the things that coincide (or something like coincide). The animalist can then counter that even if the non-animalist is right about pre-philosophers and non-philosophers having conflicting intuitions, it should be clear upon reflection that we have to go with the intuitions that favor the truth of
the human person with my body = the human animal with my body (the intuitions the thinking animal argument relies on). This I do not see. The animalist argues along these lines:

- Once we have reflected on the thinking animal argument, it becomes obvious that \( \text{the human person with my body} = \text{the human animal with my body} \) is true.
- It is obvious that if the human person with my body = the human animal with my body, then human persons and human animals are not beings with different persistence conditions.

So (once we have reflected on the thinking animal argument) it becomes obvious that human persons and human animals do not have different persistence conditions.

Suppose someone thinks it is clearly metaphysically possible for a human mind to exist and ‘work’ in isolation, after the human person to whom it belonged has ceased to have* extra-mental parts. Then he can argue as follows:

- Once we reflect on what would and would not be true, if a human mind existed and worked in isolation, after the human person to whom it belonged had ceased to have* extra-mental parts, we can see that human persons and human animals have different persistence conditions.
- It is obvious that if human persons and human animals have different persistence conditions, then the human person with my body ≠ the human animal with my body.

So (once we have reflected on the (metaphysical) compossibility of my mind’s continuing to exist and work with my ceasing to have* extra-mental parts) it is obvious that human animals and human persons have different persistence conditions, and that the human animal with my body ≠ the human animal with my body.

Suppose, though, that someone thinks it is not clear whether a human mind’s continuing to exist and work is (metaphysically) compossible with the human person to whom that mind belonged ceasing to have* extra-mental parts. She can still resist the animalist’s argument for the
(upon reflection) obviousness of the human person with my body = the human animal with my body. For, without committing herself to the metaphysical possibility of a mind’s continuing to exist and ‘work’ after it has been isolated, and the human person to whom it belonged has ceased to have* extra-mental parts, she can still Moore-shift the animalist’s argument for the (upon reflection) obviousness of the human person with my body = the human animal with my body as follows:

- It is not obvious that it is metaphysically impossible for a human mind to go on existing, and go on being involved in thoughts and experiences, even though it is isolated, and the human person to which it belonged no longer has* extra-mental parts.
- Once we have reflected on what would and would not be true if the scenario just described obtained, we can see that if it is not obvious that a human mind could not exist in isolation, even after the human person to which it belonged has ceased to have* extra-mental parts, then it is also not obvious that human persons and human animals have the same persistence conditions.
- Moreover, it is obvious that if human persons and human persons do not have the same persistence conditions, then the human person with my body ≠ the human animal with my body.

--------------------------

- So--despite the animalist’s argument to the contrary--it isn’t after all obvious (after philosophical reflection) that the human person with my body = the human animal with my body.

To my mind, this Moore-shifted argument for the non-obviousness (even after philosophical reflection) of the human person with my body = the human animal with my body is at least as plausible as the animalist argument for its (post philosophical reflection) obviousness. But isn’t it crazy to think there are two thinkers with my body, and crazy to think that (human) animals don’t think? I don’t think either option is crazy, exactly. A four-dimensionalist could tell a story about how two thinkers could share my body, and—pace Judith Jarvis Thomson—I wouldn’t go so far as to say four-dimensionalism is a ‘crazy metaphysic’. But suppose that the view that two thinkers share my body is utterly indefensible (say, because, as Olson thinks, it has disastrous epistemological implications). I don’t see that the same goes for the view that (ovine, canine, cetacean...and human) animals don’t think.
Here is an analogy. Physicists often talk about things they call “bodies”—things that move, accelerate, orbit, collide, and so on. Now what sort of things are, say, the bodies involved in a collision? It is very natural to suppose that two colliding bodies might both be naturalia—an asteroid might collide with the moon—but equally one or both of two colliding bodies might be artificialia—Russell’s “celestial teapot” might collide with a comet, or two celestial teapots might collide with each other. So, we might say, both naturalia and artificialia are bodies.

But if—as I believe—artefacts with different persistence conditions can be coincident and comateriate, and if the way physicists count bodies involved in collisions is right, then artefacts cannot after all be “bodies” in the physicist’s sense. For consider a collision between a dog and a sweater knitted from one very long thread. A physicist would describe this as a two-body collision. If, however, the sweater is a body, and the thread is a body, then the sweater and the thread are different bodies (since $x$ and $y$ are different bodies if and only if $x$ is a body, and $y$ is a body, and $x \neq y$)—in which case the collision that involves the dog and sweater involves (at least) three bodies. So either the sweater or the thread must not be a body. But neither of the pair {the sweater knit from the one long thread, the one long thread the sweater was knitted from} has a better claim than the other to be a body. So neither the sweater nor the long thread is a body. Both the sweater and the thread are things that coincide with and are comateriate with the one body that collides with the dog. (What sort of thing would that one body be? Something like a mass of ultimate matter, in the Lockean sense of “mass”, perhaps).

Now if we can say that artefacts aren’t proprie loquendo bodies, but only material objects that are coincident with and comateriate with bodies, why can’t we say that (ovine, canine, cetacean....human) animals are not proprie loquendo thinkers, but only organisms that are coincident and comateriate with thinkers? Is the view that—at least for all we know—animals at most coincide with thinkers, just as artefacts at most coincide with bodies, unmotivated? Again, not if—at least for all we know—a human can mind can exist and work in isolation, even though the human person to whom it belong no longer has* any extra-mental parts, so that—at least for all we know—human persons and human animals, like artefacts and bodies in the physicist’s sense, are different kinds of things.

I have tried to explain—I hope in less than tedious detail—why, although I used to consider the thinking animal argument convincing, I now only
consider it ‘respectable’ or ‘defensible’ (in the sense that it is a logically valid argument that has considerable intuitive appeal, and is not evidently unsound).

But suppose for the sake of argument that the thinking animal argument is sound. It seems to me there is more daylight between the soundness of the thinking animal argument and the truth of animalism than one might suppose (and than I used to suppose).

Imagine that it is the thirty first century, and you have a canine pet called Argo. Sadly, Argo contracts a very bad disease. The disease will (soon) make Argo’s heart, lungs, and liver unviable. A veterinary surgeon can transplant new vital organs “into” Argo, but the disease will (swiftly) make those transplanted organs unviable as well. However much transplanting surgeons do, in the attempt to ensure that Argo still has viable vital organs, in the end, Argo will no longer have a functioning heart, or functioning lungs, or a functioning liver. But the disease does not affect the tissues that constitute Argo’s mind. And in the thirty first century, veterinary neurosurgeons can ‘isolate’ that mind, and keep it alive and involved in thoughts and experiences. (With the help of computer scientists), they can also ensure that the thoughts and experiences that Argo’s isolated mind is involved in are very like the thoughts and experiences Argo had up to the time at which she contracted the disease: they can put Argo’s mind in a simulated reality of a more benign sort than the one in “The Matrix”.

If veterinary neurosurgeons do that, there will be a subject of the thoughts and experiences that Argo’s mind is involved in. For reasons already set out, that subject will either be a new subject brought into existence by the isolation of Argo’s mind, or an already existing individual “subjectified” by the isolation of Argo’s mind, or Argo herself. But again, it seems strange to suppose that the neurosurgeons bring a new subject of experience into existence, or “subjectify” a canine mind that had previously been a subject in potentia tantum.

The alternative is that the neurosurgeons relocate Argo in a vat, and generate a simulated reality for her. And it seems that you might naturally enough think of things this way. For suppose that the doctors at the veterinary hospital have told you that there is no way to make it go on being true (for very long) that Argo has a heart, or lungs, or a liver, but there is a way to make it go on being be true (for a long time) that Argo’s mind exists in isolation, and is involved in (non-veridical) experiences that mimic her happy (pre-illness) experiences. You want
to do what’s best for Argo. And you might well think that what’s best for Argo is that you ask the neurosurgeons (and the computer technicians) to get to work. That way—you might well think—Argo will go on having happy experiences for a long time.

It is true that many of the experiences had by the subject of the experiences Argo’s isolated mind is involved in will be “sham” experiences (that subject won’t really chase balls, or chew on bones, or cool off on summer days by jumping into the Thames, but only “have the experience of” doing those things). So, assuming that Argo = the subject of the experiences Argo’s isolated mind is involved in, Argo has “sham” experiences. Yes: but even if we assume that it’s a pro tanto bad thing for Argo to have sham experiences, isn’t it better for Argo to go on having happy experiences, at the cost of having many sham experiences, than for Argo to avoid sham experiences, at the cost of ceasing to exist (or being put in a non-simulated reality that is toto coelo different from (and much grimmer than) the one she knew before she became ill)?

Also, suppose that you (outside the vat) can interact with Argo (in the vat). Suppose that Argo, in her vat, has a kind of “avatar” that you can see and interact with on your computer or TV screen, in the way that someone can interact with a video-game character on a screen when he is playing Nintendo Wi. Using your “super-Wi” equipment, you can “virtually” toss a stick. When you do that, Argo (in her vat) will seem to see you tossing a stick. She will accordingly perform the basic actions of trying to run after the stick, trying to pick it up, and trying to bring it back to you. After she has performed those actions, you can “virtually” take the stick from the Argo-avatar, and “virtually” scratch the Argo-avatar behind her ears. That will make Argo (the real Argo in her vat, not Argo’s avatar on the screen) happy. So, it is not unnatural to think that if you let the neurosurgeons and computer scientists get to work, the result will be not just that Argo has lots more happy experiences, but also that you and Argo go on sharing a life. Argo will have genuine rather than sham experiences of playing with you (that is, she really will be playing with you, even if she’s not really holding sticks with her teeth), and these genuine experiences will contribute meaning to Argo’s life (and to yours).

I have already argued that, when we reflect on human mind-isolating scenarios, we can see that it is not obvious that you could not come to coincide with and be comateriate with your mind. If I am right, by the
same reasoning, it is not obvious that Argo could not come to coincide with and be comateriate with her mind. But if Argo is an individual that could shrink into coincidence and comateriation with her mind, then what is Argo? Well, she is a “canine subject” (a “canine quasi-person”). But is she a dog? Is she an animal?

We have a number of options here. We might say that dogs are canine animals, and canine animals cannot shrink into coincidence and comateriation with their minds; hence in the same way that a sweater is not a physical body, and a human person is not a human animal, Argo is not a dog, or a canine animal. But we might instead say that until or unless Argo’s mind is isolated, Argo is a dog, and a canine animal, though Argo is only accidentally a dog and a canine animal, inasmuch as there are (Argo-mind-isolating) metaphysically possible scenarios in which Argo continues to exist, but ceases to be a dog or a canine animal, as a result of having lost (all) her extra-mental parts. Alternatively, we might say that Argo is essentially a dog and a canine animal, but goes on being a dog and a canine animal, even after she has shrunk into coincidence and comateriation with her mind.

Although may be clearly less attractive than—it arguably stretches our concepts of caninity and animality past the breaking point, I don’t think the same can be said about). If we say that Argo is not a dog, or a canine animal, any more than a human person is a human animal, then we’ll presumably also have to say that Aurora—my older daughter’s “canine pet”—is not a dog or a canine animal (but only a canine subject that coincides with and is comateriate with a dog and canine animal). But it’s at best very surprising that Aurora is not a labrador retriever, a dog, and an animal.

The moral would seem to be that it is not a foregone conclusion that the things the folk call both “dogs” and “animals” (e.g. Aurora) have the persistence conditions that animalists attribute to organisms. More generally, it is not a foregone conclusion that all the things the folk call “animals” have the persistence conditions that animalists attribute to organisms. Perhaps the things that the folk are happy to call “unthinking animals” have the persistence conditions animalists attribute to organisms, and the things that the folk are happy to call “thinking animals” (ovine animals, canine animals, feline animals….and human animals) have different persistence conditions.

If this is so, however, then there is no short and straight path from the human person with my body = the human animal with my body—or even from
the more general human persons = human animals—to full-fledged animalism. An anti-animalist will be able to say that animalists are right to maintain that human animals and human persons have the same persistence conditions, but wrong about what those persistence conditions are (and, more generally, are wrong about what the persistence conditions of thinking animals are). Moreover, she can say, either (i) animalists are wrong about the persistence condition of various sorts of organisms (ovine, canine, cetacean...and human organisms), or (ii) (if animalists have got the persistence conditions of (all) organisms right), human animals ≠ human organisms, and canine animals ≠ canine organisms (in which case, it seems plausible to suppose, human or canine organisms do not think, lest we end up with too many thinkers with my (or Aurora’s) body). An animalist might object that option (ii) above is hopeless, because on it, even if there aren’t two thinkers with my body, there will be two living beings with my body (a human animal, and a human organism) and this is crazy. But why is there are two different living beings with my body any less entertainable than there are two different material objects made from this yarn (where the objects in question might be the sweater knit from the long thread, and the long thread the sweater had been knit from)? An animalist might say that if the human organism with my body were a different living being from the human animal with my body, the human organism would have to be a different thinker from the human animal with my body. But this last conditional is not obvious to me. It seems that we can make sense of the idea that although an isolated mind coincides with and is co-materiate with (or co-immateriate with) someone who was once a human person, the isolated mind is not a thinker, and the (ex-human) person the mind coincides with is. If so, then presumably we can make sense of the idea that although a human organism coincides with and is co-materiate with a human animal, the organism is not a thinker, although the animal is.

V

For the avoidance of doubt: nothing I have said implies that animalism is indefensible, or even clearly less defensible than at least one of the alternatives to it. And I don’t believe that animalism is clearly less defensible than the alternatives to it that are usually discussed. (On the
contrary, I think animalism is more defensible than any “neo-Lockean” view on which (a) the persistence of a human person implies the persistence of a mind, and (b) no “compositional continuity” is required for the persistence of a human person, inasmuch as the persistence of a human person could be secured simply by the (non-branching) transfer of “information” from something existing at an earlier time to something compositionally disjoint from it existing in the (immediate) future.

That said, for reasons I have tried to make clear, I am much more receptive than I once was to the idea that animalism has strongly counterintuitive consequences. It may be that the same is true about all the competitors to animalism. Perhaps no theory of the nature and persistence conditions of human persons is one we’ll ever be entirely happy with. Like Olson (and unlike Shoemaker, Lewis, et multi alii aliaeque) I want to say that unless or until we start doing philosophy, we have no inclination to think that the existence or persistence of a human person implies the existence or persistence of her mind. (Outside the study) we think that necessarily, we have been in existence for as long as we have been alive, but we do not think that necessarily, we have had a mind for as long as we have been alive (pre-philosophically, you were alive then, but you didn’t yet have a mind sounds just as possible to us as you were alive then, but you didn’t yet have fingernails). But unlike Olson (and like Shoemaker, Lewis, et multi alii aliaeque), I want to say that unless or until we start doing philosophy, we have no inclination to think that the existence or persistence of a human person implies her being or continuing to be a human animal. (Outside the study) we think that a human person’s existing/continuing to exist without being/continuing to be a human animal is possible (whether or not we think it’s something that ever happens). (This is because we are inclined to think that a human person could coincide with mind, and disinclined to think a being without extra-mental parts could be a human animal). If, however, our existing (persisting) requires neither that we are (continue to be) human animals, nor that we have (continue to have) minds, it is hard to see what our nature and persistence conditions could be. The difficulty is that we intuit that the nature and persistence conditions of human persons are as it were “non-disjunctive”, and it’s hard to see how a theory of our nature and persistence which allows us to antedate our minds, and postdate our human animality (or at any rate postdate our
coinciding with or inhabiting a human animal), can avoid the consequence that human persons have a “disjunctive” nature, and “disjunctive” persistence conditions. So, when we try to set out a philosophical account of the nature and persistence conditions of human persons, we end up in a state of aporia. And there’s no guarantee that, with a little (or even a lot) more reflection, we will be able to get beyond our aporia. Compare: thinking about how—if at all—free choice is possible or thinking about how—if at all—there can be both heaps and non-heaps generates aporiai. Philosophers have been trying—to my mind unsuccessfully—to resolve these aporiai for a long time, and they may or may not someday succeed.

On the other hand, my pessimism may be misplaced, and there may be an account of the persistence conditions of human persons (and canine quasi-persons) that (1) either accommodates or satisfyingly explains away our intuition that a human person could antedate her mentality, (2) either accommodates or satisfyingly explains away our intuition that a human person could postdate her animality (or at any rate postdate coinciding with or inhabiting a human animal), and (3) is not unappealingly “disjunctive”. If there is such an account, I suspect it will be neither animalist nor “mentalist”, and will be more accommodating to our pre-philosophical intuitions about the persistence of you and me and Argo and Aurora than both animalist accounts (whether of the Aristotelian or the contemporary variety) and mentalist accounts (whether of the Cartesian, or the Thomistic, or the Lockean or neo-Lockean variety). But I shall leave these questions for another time and place.