

The Supervenience Solution to the Too-Many-Thinkers Problem

By C. S. Sutton

Keywords: animalism, thinkers, psychological continuity, personal identity, thinking animal, first-person

Abstract

Persons think. Bodies, time-slices of persons, and brains might also think. They have the necessary neural equipment. Thus, there seems to be more than one thinker in your chair. Critics assert that this is too many thinkers and that we should reject ontologies that allow more than one thinker in your chair. I argue that cases of multiple thinkers are innocuous and that there is not too much thinking. Rather, the thinking shared between, for example, persons and their bodies is exactly what we should expect at the intersection of part sharing and the supervenience of the mental on the physical. I end by responding to the overcrowding objection, the personhood objection, the personal-pronoun reference problem and the epistemic objection.

1. Introduction

Persons think. Neo-Lockeans argue that persons are not identical to their bodies. Bodies also think. Why wouldn't they? They have brains and all of the necessary equipment. Neo-Lockeans thus seem to be committed to the claim that in your chair, there are two thinkers: a person and a body. Critics assert that this is too many thinkers and conclude that neo-Lockeanism is wrong. I argue that having two thinkers in your chair is innocuous and not a problem for neo-Lockeanism. When we understand the relationship between the person and her body properly as a case of part sharing, we see that there is no problem of too many thinkers.

I have framed the problem as one for neo-Lockeanism, but in fact, several positions face this problem. *Perdurantists* hold that you are the sum of your temporal parts (time slices). So right now, you are thinking, and your temporal part is thinking (cf. Noonan 1999: 284 ff.; Merricks 2001: 48, n. 10; Olson 2002: 191–2). *Animalists* hold that a person is identical to a human animal.¹ A prominent defence of animalism uses the too-many-thinkers problem against neo-Lockeanism, so it is perhaps surprising that animalism itself faces two versions of the problem. First, the corpse problem is that the when the animal

¹ In the too-many-thinkers literature, the non-person thinker is sometimes called the 'body' and sometimes called the 'human animal'. Neo-Lockeans tend to use 'body', and animalists tend to use 'human animal'. Whether there is a difference between the body and the human animal is debated by philosophers, but that question can be ignored for most of this paper. The exception is that the corpse problem, discussed in section 5 below, brings to the fore the relation of the body and the human animal. In other parts of the paper, I use 'body' and 'human animal' interchangeably.

dies, a body remains. Because death signals the end of the animal's existence, but the body still exists, the body and the animal are not identical. Before the animal dies, surely the body can think as well as the animal, so it seems that there are two thinkers, the animal and the body (Shoemaker 1999; Baker 2000: 207–208; Lim 2012; Árnadóttir 2013). Animalists face a second version of the too-many-thinkers problem. Your brain is not identical to the human animal. If the brain thinks, and you as a human animal think, then there are two thinking things in your chair (cf. Olson 2007: 215–9; Kovacs 2010). In this paper, I offer a solution to the too-many-thinkers problem that can be applied to any of these varieties of the problem. Along the way, I address the vanishing-weight problem.

I round out the paper by responding to objections that have been bundled with the too-many-thinkers problem. (1) The overcrowding problem charges that the neo-Lockean ontology is overcrowded with thinkers. (2) The personhood objection charges that if the body thinks, then it must count as a person, so the neo-Lockeans' distinction between persons and bodies collapses. (3) The epistemic problem says that you cannot know whether you are the person or the body. (4) And finally, the first-person reference problem says that allowing too many thinkers results in an unhealthy competition for the referent of 'I'.

2. The Too-Many-Thinkers Problem Stated

The too-many-thinkers problem has been raised by a number of philosophers (Olson 2001, 2007; Merricks 2001: 47 ff.; Árnadóttir 2010, 2013). The most vocal advocate of this problem has been Eric Olson, so let us consider one of his statements of the problem as representative:

There is a human animal intimately related to you, which some call your body. Now consider that animal's mental properties. It would seem to *have* mental properties. You have mental properties, and the animal has the same brain and nervous system as you have (and the same surroundings too, if that is relevant). [...] Consider what it would mean if you were not the animal. The animal thinks. And of course you think. [...] So if you were not that thinking animal, there would be *two* beings thinking your thoughts: there would be the thinking animal, and there would be you, a thinking nonanimal. We should each share our thoughts with an animal numerically different from us. For every thought there would be two thinkers. (2007: 29–30)

3. Some Preliminaries

For this argument, I assume a physicalist account of persons. If it turns out that the mind is non-physical or that mental states do not supervene on the physical, then there is a quick solution to the too-many-thinkers problem: reject the claim that the body must think in virtue of having the same brain and nervous system as the person.

I focus first on the too-many-thinkers problem for neo-Lockeans, who think that persons are not identical to their bodies (or not identical to human animals). If persons have psychological persistence conditions, then persons are not identical to their bodies because bodies do not have psychological

persistence conditions. For example, if a woman enters into a persistent vegetative state and never recovers, then (so says the neo-Lockean) the person no longer exists, but (so say almost all of us) the human animal still does. Because the person and the human animal have different persistence properties, they are not identical, even though they are intimately connected.

In the numbered argument below, I argue that having two thinkers would not be a problem; it would not entail too much thinking, and it would be continuous with other part-sharing cases. The problem of too-many-thinkers is blunted, so I shall argue, by the fact that the two thinkers share a supervenience base for thought. The thinking of collocated thinkers is a natural extension of *non-summative* features that we find in proper parts and wholes, such as the weight of your arm and the weight of your body.² By *non-summative*, I mean this:

Thing *t* is *non-summative* in cases in which the amount or number of *t* had by *x* and the amount or number of *t* had by *y*, taken together, is the same as the sum of the two *minus* the amount or number of *t* that is shared by *x* and *y*.

For example, consider the weight had by your body and the weight had by your arm. If your body weighs 150 lbs. and your arm weighs 9 lbs., then when you step on the scale, it registers 150 lbs. The scale did not forget your arm! The 9 lbs. of the arm is the same as 9 lbs. of the body. Weight is thus non-summative across non-identical objects where those objects share a supervenience base for weight, even though weight is summative where objects do not share a supervenience base. Likewise, thinking is non-summative where objects share a supervenience base for thinking, even though thinking is summative where objects do not share a supervenience base, such as my thinking and your thinking. We should not count the thinking of your body and the thinking of you as a person twice or else we will overcount the amount of thinking.

Overcounting the amount of thinking is an uncommon but not unheard of mistake. For example, David Mackie (1999) wrote that Harold Noonan (who defends the existence of two thinkers) does not ‘deal with the simplest objection of all—that it just seems crazy to suppose that there are two psychological lives going on where I am now’ (375). As I show in the numbered argument below, there is only one psychological life. The duplication comes in because two non-identical things share that psychological life. Dean Zimmerman (2003) worried about an overabundance of thoughts: ‘if there are two or three thinkers here, then there are two or three pains that have to be taken into account in utilitarian calculations, for instance; and two thinkers thinking “I am in pain”’ (497). Zimmerman adds that allowing for coincident entities ‘seems to introduce two or more pains, feelings of melancholy, etc., one for each coincident entity, where intuitively there should be only one’ (497). I argue below that we should recognize that there are two non-identical beings that think the same thoughts. There would be only one pain and only one feeling of melancholy. How to figure a shared pain into a utilitarian calculus is an interesting but separate question that I will leave to the ethicists, but my first-pass response is to count by quantity of pains rather than number of beings sharing that pain. The worry that there is too much

² To be accurate, I mean *mass* rather than *weight* since weight does not supervene on parts and their interrelation. But in this paper, I will use *weight* to avoid ambiguity with a *mass of matter*, which often crops up in discussions of collocation, and to provide continuity with previous discussions in the literature.

thinking is not the only, nor the most common, worry raised by the too-many-thinkers problem. But I will argue against it first because it is a primary worry and because my replies to the other facets of the too-many-thinkers depend on the fact that there is not too much thinking.

Let me note some premises that I use but do not argue for. I leave these premises unargued in order to maintain the focus of this article on the too-many-thinkers problem and because, though these premises are not undisputed, they enjoy many supporters. I take as a premise that some objects are proper parts of other objects. The list of things that exist includes hands and brains, pen caps, shirt buttons and planets, even though they are parts of larger wholes—bodies, pens, shirts and the galaxy. (For challenges to this premise, see Olson 1995; van Inwagen 1997; and Merricks 2001: 47–53, 135–6.)

I take as a premise that the person and the body share parts. There are multiple ways that this part sharing might occur. One way is by collocation or spatial coincidence. According to this view, the person and the body would have all parts in common at some breakdown of the parts (all of the same cells, for example). A second way that the person and body could share parts has been proposed by Derek Parfit (2012) and Jeff McMahan (2002). They hold that the person (or mind, in McMahan's case) is a proper part of the human being; that is, the person is a part of the human being, and the human being has additional parts that the person does not. For example, the person might extend as far as the brain and the nervous system.³ A third possibility is that a person is an 'extended mind' of the kind suggested by Carol Rovane (1993: 95–7, esp. n. 23) and by Andy Clark and David Chalmers (1998). On this view, a person's mental states are in part constituted by things in the environment, and consequently, the boundaries of the mind extend into the environment. When combined with a Lockean view, this would mean that the human being is a proper part of the person. Nothing in my argument hangs on *how* the person and the body share parts, only on the assumption that they do and that the parts that underlie the person's thinking and the body's thinking are among the shared parts. In what follows, I call the non-person thinker the 'body' or the 'human animal' for convenience, but you could substitute 'the brain and the nervous system' or some other subset of the body and the philosophical moves of my argument would be the same.

The last unargued premise to note is that the body thinks along with the person (under the proper circumstances—I do not claim that the body always can think, even in a vegetative state). Here I depart from prominent defender of the constitution view Sydney Shoemaker, who holds that the person thinks and that the body cannot (2008). My aim here is to offer an alternative solution on behalf of the neo-Lockean because some philosophers are on the verge of taking the claim that human animals cannot think as a *reductio ad absurdum* of Shoemaker's view.⁴ My aim is not to challenge Shoemaker's

³ I agree with Parfit and McMahan's reasoning that their view solves the too-many-thinkers problem, and I draw on similar reasoning to illustrate premise four of my argument below. But this reasoning need not be tethered to the view that the person is a proper part of the human being. It can be used to support a big-tent Lockean view that welcomes collocationist Lockeans and extended-mind Lockeans as well. What does the work in solving the too-many-thinkers problem is the sharing of a supervenience base, and it need not be by way of the person being a proper part of the body (though my view allows for that possibility among others). My reasoning for premise one of my argument establishes the role of a shared supervenience base in solving the too-many-thinkers problem.

⁴ Árnadóttir (2010) writes that denying that human animals can think 'is not *prima facie* appealing. Our animals would appear to share with us all the features relevant to having thoughts. They have the same brains and nervous

account—that has already been done (e.g. Árnadóttir 2010)—but rather to lay out a new response to the too-many-thinkers problem on behalf of neo-Lockeans.

Finally, let me make a note about terminology. In the argument below, I discuss *supervenience*. If x *supervenes* on y , then there cannot be a change in x unless there is a change in y . For example, the calories of a drink supervene on the ingredients and their quantities. There cannot be a change in the calories of the drink without a change in the ingredients or their quantities. Note that supervenience is not symmetric: the ingredients and the quantities do not supervene on the calories. There could be a change in the ingredients without a change in the calories; for example, one 100-calorie ingredient might be substituted for a different 100-calorie ingredient, leaving the calorie count the same.

4. Proposed Solution: The Supervenience Argument against Too Much Thinking

Here is an argument that if both the person and the human animal think, then there is not too much thinking. After the argument, I provide reasoning for each premise.

1. Where object a and object b share a supervenience base for x , x is not summative.⁵
2. A person and her body share a supervenience base for thinking.
3. Thinking is not summative where a person and her body share a supervenience base for thinking.
[from 1, 2]
4. When something is not summative, there is not too much of that thing.

Conclusion: Therefore, there is not too much thinking if both the body and the person think. If there is not too much thinking when the body and the person think, then having two thinkers is not a problem.

Premise 1: Where object a and object b share a supervenience base for x , x is not summative.

systems, the same perceptual systems, the same behavioural outputs, to name but a few. So how could they fail to think?' Árnadóttir does not rest on the prima facie implausibility of the suggestion that human animals cannot think. She gives a detailed argument against Shoemaker's view. For Shoemaker's argument that human animals cannot think, see Shoemaker's 2008 article 'Persons, Animals, and Identity'.

⁵ Supervenience is usually discussed in terms of properties supervening on properties or facts, but not always. In the *Stanford Encyclopedia of Philosophy* entry on 'supervenience', for example, the authors discuss the possibility of fusions supervening on their parts (see McLaughlin and Bennett, section 3.4). I follow that use here because our concern with the too-many-thinkers problem is about how many objects (such as thoughts) and quantities (of weight, of thinking) there are, not about how many properties there are.

Furthermore, note that non-identical objects do not share a supervenience base for instantiated properties; if object a instantiates property p , then the supervenience base for that instantiation includes the existence of object a . *Mutatis mutandis* for object b instantiating property p . Without the bearer of the property included in the supervenience base, it is unclear in what sense it would be that object's property. So premise one does not apply to instantiated properties. My claim is that the person and the body share a supervenience base for thinking and token thoughts, but they do not share a supervenience base for properties.

Premise 1 follows from understanding supervenience and part sharing. x supervenes on parts p if and only if there cannot be a change in x without a change in p or their interrelation. Imagine that parts p compose object a . Imagine then that object a later becomes a proper part of object b , though object a remains intact and its parts are the same and are interrelated the same way. Parts p are now parts of both object a and object b . Imagine for *reductio* that where object a and object b share a supervenience base for x , x is summative; that is, imagine for *reductio* that premise 1 is false. Thus, when object a becomes a proper part of object b , x increases twofold. Thus, there would be a change in x without a change in parts p or their interrelation, so x would not supervene on parts p and their interrelation. But by hypothesis, x supervenes on parts p and their interrelation. x cannot both supervene and not supervene on parts p . Therefore, we reject what was supposed for *reductio* and conclude that premise 1 is true. x remains the same whether the parts it supervenes are parts of one object or two.

Premise 1 has been resisted more often in cases of collocation (complete part sharing at some break down of the parts) than in cases of partial overlap. Critics who issue the vanishing-weight objection to collocation assume that if two objects were composed of the same parts, then the two objects together would weigh twice as much, and this implicitly denies premise 1. Because the weight of the second object vanishes, these critics conclude that there is only one object after all. David Lewis (1986: 252) first made the vanishing-weight objection as an aside in his book, *On the Plurality of Worlds*: 'It reeks of double counting to say that here we have a dishpan, and we also have a dishpan-shaped bit of plastic that is just where the dishpan is, weighs just what the dishpan weighs (why don't the two together weigh twice as much?), and so on.' Dean Zimmerman (1995) fleshed out the objection:

Some properties of physical wholes—for example, mass, weight, and shape—seem to be determined by or 'supervenient upon' the physical properties of and relations holding among an object's proper parts. Weight would seem to be thus supervenient; my living body and its constituting mass of cellular tissue are both made of physical parts so propertied and related that any whole made of such parts must weigh 140 pounds. But then how can it be so easy to lift both of those 140-pound physical objects at once? (87)

The answer to the vanishing-weight problem is part sharing, mereological overlap.⁶ In ordinary cases of part sharing (such as your arm and your body), weight is not summative. Cases of collocated objects are cases of part sharing, so it is appropriate that the weight of the body is not added to the weight of the mass of cellular tissue. This solution to the vanishing-weight problem has the virtue of providing

⁶ I place the too-many-thinkers problem in the same family as the vanishing-weight problem and the overcrowding problem. Each of these problems has a variant for anyone who accepts the existence of proper parts. I propose that this family of problems has a unified solution, namely, that cases of non-identical objects that share parts also share thoughts, weight, and occupied space.

Michael Rea (1997: 368–9) gives a similar response to the vanishing-weight objection and cites Zimmerman (who cites Peter van Inwagen) and Trenton Merricks for parts of the reply. Merricks's point was that vanishing weight, if it is a problem, is a problem for anyone who believes in composite objects that are parts of other composite objects (say, a bell tower that is part of a church) and believes that parthood is a transitive relation.

continuity between cases in which objects partially overlap and cases in which they wholly overlap. In both types of case, a shared supervenience base (partially shared or wholly shared) accounts for the fact that what supervenes is not summative. In cases of partial overlap, we recognize that where objects overlap, their weights are not summative. It would be odd if the weight of two objects that shared all but one part were *non-summative* where they overlapped, but in a case in which the one object's extra part was removed so that the two objects shared all parts (at some breakdown of the parts), the weights were suddenly *summative*.

We should accept premise 1 because it follows from an understanding of supervenience. Furthermore, those who reject premise 1 by endorsing the vanishing-weight problem against collocationism commit themselves to an unnatural discontinuity between cases of partial overlap and cases of complete overlap.

Premise 2: A person and her body share a supervenience base for thinking.

This is a premise that gets the too-many-thinkers problem off the ground in the first place. Those who pose the too-many-thinkers problem argue that whatever the person uses to think, such as the brain and nervous system, the human animal has as well, which they say, makes for too many thinkers. So although this premise is controversial (substance dualists reject it, for example), it is accepted by those whose arguments I am responding to here.

The view that the mental supervenes on the physical can be cashed out in a variety of ways. The supervenience claim could hold if mental states are *identical to* physical ones, or if the physical *constitutes* the mental, or if the mental is *multiply realized by* the physical, or if mental states are *functions of* physical states. The supervenience claim is neutral on whether the physical causes the mental, or vice-versa, or if causation is two-way. I draw on supervenience to stay neutral among various positions in the philosophy of mind. As an aside to supervenience connoisseurs, I am assuming *individual* supervenience of the *weak* variety.⁷

Subconclusion 3: Thinking is not summative where a person and her body share a supervenience base for thinking.

Subconclusion 3 follows from premises one and two.

⁷ A case might be made for *strong* individual supervenience of the mental on the physical, but for the purposes of the too-many-thinkers problem, the weaker version suffices. We are considering overlapping persons and bodies, and physical objects that overlap are in the same possible world, which is the scope of weak supervenience. I will not stake a claim to strong supervenience because, for all I know, there is a possible world that has consciousness-scrambling or consciousness-blocking effects, such that a thinking body in this world could have a non-thinking counterpart in that world.

Premise 4: When something is not summative, there is not too much of that thing.

Let us first consider an illustrative case of this premise before moving on to the reasoning behind it. Recall the case of your arm and your body. The cells that make up your arm are the same cells that make up the arm-region of your body, and those cells are the supervenience base both for the arm's weight and for part of the body's weight. You might designate the supervenience base to be at a more microphysical level (molecules, say, or smaller), but those are still shared by both the arm and the body's arm region. Based on premises one and two, we conclude that weight is not summative where your arm and your body share a supervenience base for weight. There is not too much weight when your arm and your body are weighed together. If your body weighs 150 pounds, and your arm weighs 9 pounds by itself, when you step on the scale with your arms attached, the scale registers 150 (not 159). The arm does not register twice, even though it is not identical to your body. This can be empirically verified, though I recommend against weighing your arm by itself.

If neither weight nor thinking is summative in cases of part-sharing, then we should expect parity on the question of whether the quantity is too much. The claim of 'not too much' is imprecise, but it can be established nonetheless. Those who raise the charge of too many thoughts, or too much weight, or too many thought lives propose as an alternative that there is only one feeling of melancholy, only 150 pounds when I step on the scale and only one thought life. Premises one through three establish that the things in question are non-summative, and therefore the number or quantity is the sum of the two minus the amount that is shared. The total weight of the body and arm is 150 pounds. We add the body's thought life to the person's thought life, and subtract the shared thought life, to find that there is only one thought life after all. Because the quantity or number in cases of non-summativity is the same as the number that critics consider the target number, the amount is not too much on the critics' own terms.

Conclusion: Therefore, there is not too much thinking if both the body and the person think.

The conclusion follows from premises three and four.

To conclude this section, if there are two (or more) thinkers, it is not a problem but rather exactly what we should expect at the intersection of supervenience and part sharing. Let us turn to the too-many-thinkers problem as it is faced by perdurantism.

5. Perdurantism and the Problem of Thinking Temporal Parts

Perdurantists believe that a person is spread out over time. Just as you are composed of spatial parts—including your head, torso and limbs—you are also composed of temporal parts—including your childhood parts, middle-aged parts and elderly parts. You are the sum of your *spatiotemporal* parts.

Both spatial and temporal parts can be much more fine-grained, such as cells and atoms and instantaneous 'time slices'. (For an overview of temporal parts, see Hawley 2010.)

On this view, you think right now in virtue of having a temporal part that is thinking. So you think, and your temporal part thinks. In fact, you have multiple temporal parts that think: you right now, you today and you this week, for starters. So the perdurantist faces the problem of too-many-thinkers (Noonan 1999: 284 ff.; Merricks 2001: 48, 97–9; Olson 2002: 191–2).⁸

The supervenience argument solves the perdurantist's problem. We can modify the numbered argument above to see how it applies to temporal parts. The main change is to replace 'a person and her body' with 'a person and her temporal part'. I put in bold-face the premises that change in this version of the argument, with notes on the justification for those premises.

1. Where object *a* and object *b* share a supervenience base for *x*, *x* is not summative.

2. A person and her temporal part share a supervenience base for thinking.

Recall that premise 2 is needed to motivate the too-many-thinkers problem in the first place. The perdurantist says that you think in virtue of your temporal part thinking. We could cash this out in one of two ways. First, we might say that both your thinking and your temporal part's thinking supervene on your brain activity [or enter your preferred physical base here]. Secondly, we might say that the supervenience relationship is nested: your thinking supervenes on your temporal part's thinking, which in turn supervenes on the activity in your brain. This second way of formulating the supervenience relation does not threaten premise two, for supervenience is transitive (McLaughlin and Bennett 2011, section 3.2). It follows by transitivity that your thinking supervenes on the activity in your brain, even if your thinking also supervenes on your temporal part's thinking. Either way we cash out the supervenience relationship, premise two stands.

3. Thinking is not summative where a person and her temporal part share a supervenience base for thinking.

Subconclusion 3 follows from premises one and two.

4. When something is not summative, there is not too much of that thing.

Conclusion: Therefore, there is not too much thinking if both the temporal part and the person think. If there is not too much thinking when the temporal part and the person think, then having multiple thinkers is not a problem.

⁸ Another form of four-dimensionalism, stage theory, avoids the problem. According to stage theory, you are identical to a momentary stage or temporal part. You are not the sum of your temporal parts, and there is no person extended in time that competes with the temporal part for the title of 'thinker'. For a defence of stage theory, see Sider (1996). The perdurantist version of four-dimensionalism, which is discussed above, is the more common version.

This conclusion follows from premises three and four. The supervenience solution to the too-many-thinkers argument can thus be employed by perdurantists.

6. Animalism and the Corpse Problem

Animalism's corpse problem centres on the relationship between the human organism and her body. At death, the person ceases to exist, and according to animalism, the person is identical to the human animal, so the human animal ceases to exist. But the body remains, as a corpse. So the person is not identical to her body. Though the lifeless body does not think, it once had all the same equipment to think as the person did, so before the person died, the body was a thinker. So we once again have two thinkers: the person and her body.

There are other ways to solve the corpse problem, so in this case the supervenience solution might be unnecessary. One could hold, like Olson, that death is such a drastic change that it brings something new into existence, a corpse, at the same time the person goes out of existence (2004). On this view, we understand the term 'body' to be ambiguous between an animal and a corpse. But at no time is there both an animal and a body. Or one could hold with Judith Thomson (2008) that the person is *identical* to her body, and a corpse is a person, namely, a dead person, so the body and the person are the same thinker. But for those dissatisfied with the existing solutions to the corpse problem, the supervenience argument would read just as it did for the neo-Lockeans. The key difference is that animalists understand 'person' to mean 'human animal'. So a second, equivalent way to phrase the solution would have premise two read,

2. A *human animal* and her body share a supervenience base for thinking.

Subsequent premises would also substitute 'human animal' for 'person'. The conclusion would be that the body's thinking and the human animal's thinking does not add up to too much thinking.

7. Animalism and Thinking-Brains Problem

The supervenience solution straightforwardly solves the thinking-brain problem for animalists. Recall the problem: Animalists hold that the animal is identical to the person, so there is only one thinker. But the brain is clearly not identical to the animal because, for starters, the brain is much smaller than the animal. Eric Olson, the most famous animalist, is friendly to solving this problem via denying the existence of brains (1995; 2007: section 9.3), though he seems open to other solutions in his 2007.

Denying the existence of brains is a bold move. The supervenience solution I offer is a less drastic alternative. To address the thinking-brain problem, we only need to replace 'body' with 'brain' in the supervenience argument. Once again, I put in bold face the modified premises to facilitate skimming the now-familiar argument:

1. Where object *a* and object *b* share a supervenience base for *x*, *x* is not summative.

2. A person and her brain share a supervenience base for thinking.

3. Thinking is not summative where a person and her brain share a supervenience base for thinking.

[from 1, 2]

4. When something is not summative, there is not too much of that thing.

Conclusion: Therefore, there is not too much thinking if both the brain and the person think. If there is not too much thinking when the brain and the person think, then having two thinkers is not a problem.

Let us turn now to other objections related to the too-many-thinkers problem.

8. The Overcrowding Objection

Eric Olson takes what he calls the ‘overcrowding problem’ to be the heart of the too-many-thinkers problem: ‘it means that there are far more thinking beings than we thought’ (2007: 35). We typically think that when you and your friend meet for coffee, two thinkers are sitting at the table, but if persons are not their bodies, then there are four thinkers enjoying the effects of caffeine.

In response, I should make two points. First, note that if we allow that the person and animal both think, we can still say truly and informatively that there is one thinker in your chair. We commonly restrict quantification when numerically distinct objects share parts and function. For example, we say that there is one thing holding coffee on the car, even though the travel mug with its spill-proof lid is a numerically distinct object from the tumbler (what is left when the lid is removed). The restricted quantification helps us mark an important difference between two functionally distinct travel mugs, on the one hand, and a travel mug and its parts on the other. When we count things that hold coffee, we are almost always interested in counting the former and not the latter. Likewise, when we count thinkers (e.g. How many thinkers were in the Vienna Circle?), we are almost always interested in counting thinkers with unique thought lives and not numerically distinct objects that share a thought life. Restricting quantification when we count thinkers is not an *ad hoc* metaphysical trick; instead, it gives us continuity with the way that we count in run-of-the-mill part-whole cases.

Eric Olson protests, ‘the mere existence of the hypothesis that we “count” philosophers by a relation other than numerical identity does little to make it easier to believe that there are two numerically different philosophers sitting there and reading this now’ (2007: 36). Notice that the overcrowding objection does not engage the *reasoning* that leads to the conclusion that the person is numerically distinct from his or her body. Instead, the worry seems to be that the conclusion offends our pre-theoretical or common-sense notions about how many thinkers there are.

This brings me to my second point about the overcrowding objection. The overcrowding objection is hasty; it is not at all clear that counting more than one thinking thing in your chair is out of step with common sense. Consider the common-sense answers to the following questions:

(1) Do brains think? —Yes.

(2) Do you think? —Yes.

(3) Are you more than your brain? —Yes. (This question is a pre-theoretical way to ask about numerical identity.)

These pre-theoretical answers, taken together, imply that there is more than one thing that thinks in your chair. But, one might protest, common sense answers ‘No’ to the question,

(4) Is there more than one thinker in your chair? —No!

I happily concede that this is the pre-theoretical answer to question (4), but note that this shows that (not surprisingly) common sense is internally inconsistent, as it is in other metaphysical matters (free will, anyone?). The overcrowding objection worries about offending our common-sense scruples, but in the case of whether we should count more than one thinking object in your chair, common sense damns us if we do and damns us if we don’t. (Incidentally, in order to avoid positing too-many-thinkers, Olson [1995] rejects the existence of brains. Common sense is not amused.) So let us set aside the goal of trying to please our pre-theoretical notions of how many things think and move on to other objections.

9. The Personhood Objection

The personhood objection to the view that both the person and the animal think is that this view seems to imply that the animal is a person (Snowdon 1990: 94; Olson 2007: 36–7; Árnadóttir 2013: 576). And if the animal is a person, then animalism is true. And if animalism is true, and animals can be identical to persons, then neo-Lockeanism is unmotivated. Olson writes, ‘[Your animal] satisfies every ordinary definition of “person”: it is, for instance, “a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places”, as Locke put it’ (36–7). (The Locke quote is from *An Essay Concerning Human Understanding*, Book II, Chap. 27, section 9.) If the animal qualifies as a person, then it turns out ‘that some human people *are* animals [that is, are numerically identical to animals], even if others are not’ (Olson 2007: 37).

There are not two collocated persons, even if Locke’s definition gets it basically right. Recall the original motivations for thinking that a person and her body are not identical. In fanciful thought experiments in which our thought lives are transferred to a different body, we identify ourselves with the thinking body that received the transplanted thoughts, not the body that lacks sentience. And in real-life cases in which someone enters a persistent vegetative state, we lament the loss of that person, even while her body lives. These cases lead us to believe that even though the person and the animal share most properties and activities (thinking, talking, working, tasting wine), the person and her body differ in

persistence properties. If the person and the body differ in some properties, then they are not identical, per the indiscernibility of identicals—or its contrapositive, to be precise. They are discernible because the person could continue to exist in a new body, but the body could not; the body could continue to exist in a persistent vegetative state, but the person might not. There is a set of properties that characterize a person, and there is a set of properties that characterize a human animal. The sets overlap, but not completely. These sets of properties give us reason to posit the existence of (1) a person and (2) a human animal. But there is only one set of *person* properties here, not two. So we have no reason to posit the existence of two persons. Yes, the animal can think, but it can also survive the loss of those thoughts, and so it does not have all of the characteristics of a person. It is not a person.

Does this mean that we must abandon the Lockean definition of ‘person’? No. We keep the spirit of the definition with a key tweak. The Lockean characterization must be had *essentially* for the thinking, intelligent thing to be a person. (Baker 2000’s account of personhood, for example, makes this tweak.)⁹

10. The Personal-Pronoun Reference Problem

Another facet of the too-many-thinkers problem is that of which of the two thinkers the pronouns ‘I’ and ‘me’ refer to. Steinvör Árnadóttir states the problem: ‘One problem is that the referent of “I” is usually taken to be the thinker or speaker of the thought or utterance, but if the animal and person are distinct, and both have the capacity for thought, then we get a competition for the reference of “I”, resulting in a number of problems to do with meaning, truth, and the individuation of thoughts’ (2013: 576).

The dominant answer on the part of neo-Lockeans has been to say that ‘I’ refers to the person only and then to give some account of why that is. Sydney Shoemaker, for example, argues that animals cannot think (2008). Harold Noonan argues that ‘person’ should be defined as the object of first-person reference, in which case ‘I’ always refers to the person (1998; 2001; 2010: 93–4). In this case, there is no competition for the reference of ‘I’ as long as we use the pronoun correctly.

But it seems that in ordinary usage, we use ‘I’ to refer to ourselves as persons sometimes, as animals sometimes, and both at once much of the time. I am optimistic that we can make sense of this and that the question of which one ‘I’ refers does not require a univocal answer across uses. Consider the following sentences, in which it seems clear that on the neo-Lockean account, the referent is not always the same.

⁹ A tempting alternative account, offered by Alan Sidelle, has it that the property of *being a person* is a different property than the property of *being sortally a person*. On Sidelle’s account, both the person and the animal have the property of *being a person* but only the person has the property of *being sortally a person*. Likewise, he argues that both the animal and the person have the property of *being an animal*, but only the animal has the property of *being an animal sortally*. This is akin to how ‘being red is distinct from being essentially red’ (1). See his manuscript ‘Coincidence: thick and thin, top-down and bottom-up’.

(1) Referent is the human animal: In my mother's womb, I [human animal] did not kick or move around when noisy trains approached the station, leading my father to worry that I [human animal] was deaf.

When the human animal and the person together utter the above sentence, both use 'I' to refer to the human animal. The person is not saying falsely that the person was stoic about train noise but rather saying truly that the human animal was. (How one thinker can use a personal pronoun to refer to another thinker is discussed in the next section.)

(2) Referent is the person: If my cerebrum were transplanted, I [person] would be surprised to wake up in a different body.

Once again, the person and the animal use the pronoun univocally, in this case, both use 'I' to refer to the person.

But I suspect that these are the minority of cases, and that in the majority of uses of 'I', the proper referent could be both the animal and the person. Such cases would include

(3) I enjoy eating chocolate.

(4) I went to the office today.

In these cases, it is reasonable to think that we are making claims about both persons and bodies.

The above sentences show that it would make sense to interpret 'I' as ambiguous among an animal, a person, and both at once.¹⁰ We can disambiguate based on context; for example, the body-switching context in sentence (2) suggests that the referent is the person.¹¹ So there *is* competition for the reference of 'I'. But why is such competition thought to be unhealthy? One problem is that the use of 'I' is thought to be immune to reference error. If there are two things jockeying for the reference of 'I', then it would be easy to misattribute the pronoun 'I'. A second problem is that 'I' is considered to be a singular pronoun. How could it refer to both a person and an animal at the same time? Let us consider these problems in turn.

10a. The Impossibility of a First-Person Reference Mistake

The first problem is that first-person reference mistakes are thought to be impossible, or nearly so. If they are impossible, then we should reject a theory that allows that a *person* could use 'I' to refer to a *human being* rather than itself, or that a *human being* could use 'I' to refer to a *person*. When the person and the animal together think, 'I did not kick or move around when noisy trains approached,' the 'I' refers to the human animal, not the person. The person has misused the pronoun 'I' to refer to something other than itself.

¹⁰ See Mills (1993) for another case of ambiguity with the reference of 'I'.

¹¹ It suggests that the referent is the person if we interpret the sentence charitably. I do not mean to suggest that what establishes the referent is a matter of which referent would make the utterance true. Instead, the referent is established by the speaker and context in the typical way. It would be uncharitable to interpret someone who utters sentence (2) as saying that 'If my cerebrum were transplanted, I [my body] would be surprised to wake up in a different body'. But of course, someone might claim just that if he or she were confused.

The solution to the personal-pronoun reference problem lies in looking at the conditions for immunity from reference error. Although the *reasons* for holding that first-person reference mistakes are impossible might support the conclusion that *I* cannot mistakenly refer to *you* in the first-person, those reasons do not support the conclusion that a *person* cannot mistakenly refer to his or her *collocated animal*. On standard theories of first-person reference, why is 'I' believed to refer without fail to the correct person and never to someone other than the speaker? There are two factors: (1) The personal pronoun reflexively refers to the thinker or speaker.¹² But note that if there are two thinkers of these thoughts, then reflexive reference alone will not pick out a unique referent of 'I'. This opens the door for reference error. (2) We each have private, privileged access to our own thoughts and mental experiences (Coliva 2003: 419). Privileged access means that we do not have to do any work to figure out which thinker or speaker we are referring to. I have access to my own thoughts and not yours, nor my fellow bus passengers', nor my family members', so there seems to be only one candidate for the referent of 'I'. That candidate's election is fool-proof.

But the human animal and collocated person have privileged access to one another's thoughts in virtue of those thoughts being numerically identical. The person's thoughts are not private from the animal and vice-versa. Therefore, the second condition for immunity to reference error is not met in the case of the human animal and person. We should not expect the person and animal that share thoughts to be immune from first-person reference error, but this does not impugn sure-fire first-person reference in the common cases of non-identical persons or non-identical animals.

10b. Must the Referent of the First-Person Pronoun Be Singular?

I have proposed that 'I' can refer to the person, to the body, or to both at once. But why should we think that the first-person could refer to more than one thing? Is it not essential to the first-person that it refers to a single thing?

First, let us keep in mind that for most people, the use of 'I' is not metaphysically informed. The English personal pronouns developed before Locke made popular the idea that *human being* and *person* do not refer to the same thing and before perdurantism had been conceived as an account of the way persons exist. We should not expect our pronouns to be free from metaphysical error.

Secondly, even if neo-Lockeanism or perdurantism is true, each view countenances a unity, which makes the single pronoun appropriate. Neo-Lockeans posit a person and an animal, but these two things have between them *a single consciousness*. Perdurantists explain the existence of *a single person over time* in virtue of the time slices that make up that person. *E pluribus unum*. Because there are numerically distinct entities working in concert, we can use the singular pronoun appropriately. We see precedent for this in English when we treat as singular 'jury', 'committee' or 'team'. Each of these collective nouns

¹² Carol Rovane suggested that the referent of 'I' is 'the thinker of this thought' (cited in Coliva 2003); Marga Reimer (2009) suggested that the referent of 'I' is 'the speaker of this utterance'. Noonan suggested that 'I' refers to the person who uses it (1998, 2001, and 2010: 93–4).

has numerically distinct members, but they work as a unit, and their unity is *much, much* looser than the unity of the person and his or her body or the unity of the person and his or her time slices. In sum, the first-person pronoun could refer appropriately to more than one thing working in concert.

What about a case in which context suggests that an utterance is intended to refer to both the person and the body together, but the utterance is true of the one thinker and false of the other? For example, what if someone, intending to refer to both the person and the body as a unity, claimed, 'I [the person and the body together] will be buried in the family plot'? The utterance, unless it is a strange plan to be buried alive, would apparently be false of the person but true of the body.

But on my view, this is not a case of two thoughts, one true and one false, but rather one thought that is false. It is false that both the person and the body (as a unity) will be buried. Further, the case where the utterance is meant to refer to both thinkers would be an atypical case, in which the utterance is based on a confusion. Most people do not expect their consciousness, which is central to their personhood, to be buried underground. A claim that begins 'I will be buried...' would more often be one in which 'I' refers to the body.

11. The Epistemic Objection

The epistemic objection is this: if there are two thinkers thinking my thoughts, it seems that I could not know which thinker I am (Olson 2007: 36; Olson 2010: section 6; Zimmerman 2003: 497; Árnadóttir 2013: 576). Olson puts it this way: 'the Psychological Approach raises the questions, "What am I? Am I a human person, who persists by virtue of psychological continuity? Or am I an animal?" And here too there seem to be no grounds on which to answer these questions' (2010: section 6, paragraph 4). The epistemic problem is that the person and the animal both think 'I am a person', and both have the same evidence for this (Árnadóttir 2013: 576). The person is right, but the animal is wrong. Because the prior probability of being the thinker that correctly judges itself to be a person is only fifty percent, our confidence that we are persons should be low, so we do not know whether we are persons after all. But the neo-Lockean position is premised on knowing that we are persons.

The epistemic objection makes two mistakes. One mistake is the claim that, according to the neo-Lockean view, if the person believes 'I am a person' then the animal will believe, too, that it (the animal) is a person. I argued above that 'I' is ambiguous. In the context of defining a person as a being that is a thinker *essentially*, the animal will often refer to the person when the animal utters, 'I am a person'. Consider the thought-transplant case that motivates the psychological view. If your thoughts are transferred to a new body, then it is natural to conclude that you would wake up in a new body. This motivates the belief 'I am a person', a thing that thinks essentially. It does *not* motivate the belief that the body thinks essentially, for the original body is left devoid of the ability to think. In this context, the natural disambiguation of 'I' is that it refers not to the body but to the essential thinker. Both the person and the body think, 'I [the person] would wake up in a new body,' and both conclude that 'I [the person] am a person'. In contrast, the person refers to the animal when the person thinks, 'I could survive brain death with the aid of life support'.

Olson's question, 'What am I?' has no disambiguating context, so the question is ambiguous. Philosophers of language take 'I' to refer to the thinker or the speaker (see the previous footnote). In the absence of context, we can rephrase Olson's objection this way: 'What is the thinker of these thoughts? Is the thinker a human person, who persists by virtue of psychological continuity? Or is the thinker an animal? And here there seem to be no grounds on which to answer these questions.' In this paper, I have given grounds—in the form of the supervenience argument—for concluding that these thoughts belong both to a person and to a human being. The supervenience argument suggests that the epistemic problem is *not* unanswerable; the answer is just not straightforward because the question presupposes a falsely exclusive dichotomy of person or animal. The bearer of these thoughts is not *one of* a person or animal; instead, the thoughts belong at the same time to a person and to an animal.

It is true that the person and the animal share a first-person experience of the world, so the thinkers cannot discriminate phenomenologically between themselves. This brings us to the second mistake of the epistemic objection: the inference from that fact of inner indiscernibility to the sceptical conclusion that, on a neo-Lockean view, you could never know that you are a person. By analogy, imagine that highway A and highway B run together for 100 miles and that you are driving on the stretch where they are merged. Which highway are you driving on? Both, and they are indiscernible from where you sit. But it would be wrong to conclude that you did not know whether you were driving on highway A or highway B. You know that you are driving on highway A and on highway B. It follows that you know that you are driving on highway A. Likewise, if you know that these thoughts are had by a person and by a human animal, then you know that these thoughts are had by a person. In other words, you know that you are a person.

12. Conclusion

I have argued that in the sense in which there are two thinkers, two non-identical things that think, this is innocuous and there is not too much thinking but only one train of thought. The non-summativ thinking shared between persons and their bodies is a natural extension of partial overlap cases where what supervenes on a shared base is non-summativ. In response to the personhood objection, I have argued that the human body is not a person, even though it thinks, because it lacks the persistence properties of a person. I have argued that we should expect personal-pronoun reference to be vulnerable to error on a neo-Lockean or perdurantist account and that 'I' can appropriately refer to thinkers that share a thought life. In response to the epistemic problem, I have suggested that the animal will not inevitably believe that it is a person, and I have argued that our nature as persons is knowable.

Eric Olson called the too-many-thinkers problem 'the principal support for animalism'. He continued, 'If it fails...then it is hard to see how there could be any reason to suppose that we are animals' (2004: 267). If that is right, then once we understand the relationship between persons and human beings in

the proper context of supervenience and part sharing, we should start looking for new reasons to be an animalist.¹³

Virginia Commonwealth University, USA

¹³ For suggestions that improved this paper, I am grateful to Sara Bernstein, Andrew Cullison, Janice Dowell, Matt Duncan, Maya Eddon, Neil Feit, David Hershenov, Sheiva Kleinschmidt, Karen Lewis, Trenton Merricks, Eugene Mills, Alan Sidelle, Donald Smith, Peter Sutton and Dale Tuggy. Thanks to Ann Cudd and Louise Antony for the opportunity to workshop this paper at the Mentoring Project for Pre-tenure Women Faculty in Philosophy, and to audiences at the University of Virginia, the SCP 2013 Eastern Regional Conference, and SUNY Fredonia.

Works Cited

- Árnadóttir, S. T. (2013) 'Bodily Thought and the Corpse Problem', *European Journal of Philosophy*, 21: 575–92.
- Árnadóttir, S. T. (2010) 'Functionalism and Thinking Animals', *Philosophical Studies*, 147: 347–54.
- Baker, L. R. (2000) *Persons and Bodies: A Constitution View*. New York: Cambridge University Press.
- Clark, A. and D. Chalmers (1998) 'The Extended Mind', *Analysis* 58/1: 7–19.
- Coliva, A. (2003) 'The First Person: Error through Misidentification, the Split between Speaker's and Semantic Reference, and the Real Guarantee', *Journal of Philosophy*, 100/8: 416–31.
- Hawley, K. (2010) 'Temporal Parts', *The Stanford Encyclopedia of Philosophy* (Winter 2010 Edition), Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/win2010/entries/temporal-parts/>>.
- Kovacs, D. M. (2010) 'Is There a Conservative Solution to the Many Thinkers Problem?' *Ratio*, 23/3: 275–90.
- Lewis, D. (1986) *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- Lim, J. (2012) 'Are We Essentially Animals?' Paper presented at the American Philosophical Association Eastern Division One Hundred Ninth Annual Meeting, 29 December, Atlanta, Georgia.
- Mackie, D. 1999. 'Animalism versus Lockeanism: No Contest', *Philosophical Quarterly*, 49/196: 369–76.
- McLaughlin, B. and K. Bennett (2011) 'Supervenience', *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/win2011/entries/supervenience/>>.
- McMahan, J. (2002) 'Section 5.5 Mind, Brain, and Organism', in *The Ethics of Killing: Problems at the Margins of Life*, 88–94. Oxford: OUP.
- Merricks, T. (2001) *Objects and Persons*. Oxford: OUP.
- Mills, E. (1993) 'Dividing without Reducing: Bodily Fission and Personal Identity', *Mind*, 102/405: 37–51.
- Noonan, H. W. (1998) 'Animalism versus Lockeanism: A Current Controversy', *Philosophical Quarterly*, 48/192: 302–18.
- Noonan, H. W. (1999) 'Identity, Constitution, and Microphysical Supervenience', *Proceedings of the Aristotelian Society, New Series*, 99: 273–88.
- Noonan, H. W. (2001) 'Animalism versus Lockeanism: Reply to Mackie', *Philosophical Quarterly*, 51/202: 83–90.

- Noonan, H. W. (2010) 'The Thinking Animal Problem and Personal Pronoun Revision', *Analysis*, 70/1: 93–8.
- Olson, E. (1995) 'Why I Have No Hands', *Theoria*, 61/2: 182–97.
- Olson, E. (2001) 'Review of Persons and Bodies: A Constitution View, by Lynne Rudder Baker', *Mind*, 110/438: 427–30.
- Olson, E. (2002) 'Thinking Animals and the Reference of "I"', *Philosophical Topics*, 30/1: 189–207.
- Olson, E. (2004) 'Animalism and the Corpse Problem', *Australasian Journal of Philosophy*, 82/2: 265–74.
- Olson, E. (2007) *What Are We?* Oxford: OUP.
- Olson, E. (2010) 'Personal Identity', *The Stanford Encyclopedia of Philosophy* (Winter 2010 Edition), Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/win2010/entries/identity-personal/>>.
- Parfit, D. (2012) 'We Are Not Human Beings', *Philosophy*, 87/1: 5–28.
- Rea, M. (1997) 'Supervenience and Co-location', *American Philosophical Quarterly*, 34/3: 367–75.
- Reimer, M. (2009) 'Reference', *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/spr2010/entries/reference/>>.
- Rovane, C. (1993) 'Self-Reference: The Radicalization of Locke', *Journal of Philosophy*, 90/2: 73–97.
- Shoemaker, S. (1999) 'Critical Notice of Olson, *The Human Animal*', *Nous* 33/3: 496–504.
- Shoemaker, S. (2008) 'Persons, Animals, and Identity', *Synthese* 162/3: 313–24.
- Sidelle, A. 'Coincidence: Thick and Thin, Top-Down and Bottom-Up [Or: How Much Do Coincident Entities Differ, and Why?]', Unpublished manuscript (July 2013 version).
- Sider, T. (1996) 'All the World's a Stage', *Australasian Journal of Philosophy*, 74/3: 433–53.
- Snowdon, P. F. (1990) 'Persons, Animals, and Ourselves', in C. Gill (ed.) *The Person and the Human Mind*. Oxford: OUP.
- Thomson, J. J. (2008) 'People and Their Bodies', in T. Sider, J. Hawthorne and D. W. Zimmerman (eds.) *Contemporary Debates in Metaphysics*, 155–76. Malden, Massachusetts: Blackwell Publishing. First published 1997 in J. Dancy (ed.) *Reading Parfit*. Oxford: Blackwell Publishing.
- Van Inwagen, P. (1997) 'The Doctrine of Arbitrary Undetached Parts', in M. Rea (ed.) *Material Constitution*, 191–208. Lanham, Maryland: Rowman & Littlefield. First published 1981 in *Pacific Philosophical Quarterly*, 62: 123–37.
- Zimmerman, D. (1995) 'Theories of Masses and Problems of Constitution', *Philosophical Review*, 104/1: 53–110.

Zimmerman, D. (2003) 'Material People', in M. J. Loux and D. W. Zimmerman (eds.) *Oxford Handbook of Metaphysics*, 491–526. Oxford: OUP.