

The Many-Subjects Argument against Physicalism

Brian Cutter

University of Notre Dame

1. Introduction

There is exactly one conscious being in your vicinity—you. But if physicalism is true, there are many conscious beings in your vicinity, each with experiences similar to yours. Therefore, physicalism is false.

This is the basic form of the *many-subjects argument* against physicalism, which I attempt to clarify and defend in this paper. It comes in several forms. One version poses the threat of conscious *parts*. Its key premise is that if physicalism is true, then not only you, but also certain parts of your body, are conscious. These parts might include your brain, your head, all-of-you-except-your-left-pinky, all-of-your-brain-except-one-neuron, or your left cerebral hemisphere. A second version poses the threat of conscious *coinciders*, things that currently occupy the same spatial region as you and share all your matter. Its key premise is that if physicalism is true, then not only you, but also some things that currently coincide with you, are conscious. These might include your body or the aggregate of particles of which you are currently composed. A third version poses the threat of conscious *person candidates*, the many human-body-shaped material objects in your vicinity that mostly overlap one another, differing only in a few peripheral atoms. This version claims that if physicalism is true, then there are many conscious person-candidates in your vicinity.

As we'll see, there are ways for the physicalist to avoid the result that there are many subjects (i.e., many experiencers or phenomenally conscious entities) in your vicinity. But all such strategies are costly, and most have a consequence almost as bad: that there are many things

in your vicinity in a state *very much like* consciousness, a state with similar normative significance.

The many-subjects puzzles discussed in this paper can be seen as instances of a more general set of puzzles. Similar arguments purport to show that there are many clouds, many tables, or many houses where we ordinarily think there is one.¹ But no one should respond to the latter arguments by rejecting physicalism about cloudhood, tablehood, or household, so it's reasonable to wonder how parallel arguments could shake our faith in a physicalist view of consciousness. The answer is that many-subjects puzzles are importantly different from these other puzzles. There are plausible responses to the latter whose analogues for the case of conscious subjects are extremely implausible, as I'll argue below. The difference is largely due to the normative significance of consciousness. If we say about conscious subjects what we ought to say about clouds, tables, and houses, we are saddled with unacceptable revisionary normative conclusions.

The many-subjects argument poses a threat to physicalism that is largely independent of the more familiar suite of "epistemic arguments" against physicalism (the knowledge argument, the conceivability argument, etc.). The most important difference, in my view, is that the many-subjects argument has a broader target. It's widely held that epistemic arguments, if they work at all, only work against "standard physicalism," not "Russellian physicalism."² In contrast, the many-subjects argument, if successful, would establish the falsity of physicalism in both its standard and Russellian forms. Indeed, the many-subjects argument has extra force against Russellian physicalism. Some of its premises are motivated by intuitively plausible claims about the nature of phenomenal properties, such as that phenomenal properties are intrinsic,

¹ See Unger (1980) and Geach (1980).

² See, e.g., Stoljar (2001) and Chalmers (2003).

categorical, or normatively significant. In practice, these claims are more likely to be accepted by Russellian physicalists, partly because they fit especially well with the distinctive commitments of Russellian physicalism (e.g., that consciousness is absent from the scientific image precisely because the scientific image leaves out the intrinsic, categorical qualities of matter), and partly because Russellian physicalists tend to be more sympathetic to the idea that we can know a great deal about the nature of experience through introspection or *a priori* reflection.³

If the many-subjects argument reduces our confidence in physicalism, it should raise our confidence in (property) dualism, the view that phenomenal truths and physical truths are co-fundamental and ontologically distinct. Dualism rejects the physicalist claim that phenomenal truths are grounded in, reducible to, or analyzable in terms of physical truths (understood broadly to include not just the structural/relational truths revealed by the physical sciences, but also the underlying quidditative truths that figure in Russellian physicalist accounts of consciousness). The argument might also be used to support some forms of idealism, but I will not discuss idealism here. Throughout, I'll show how the various many-subjects puzzles can be handled by the two main forms of dualism: substance dualism and the dual-aspect theory. Substance dualism holds that, in addition to physical individuals, there are non-physical concrete individuals, and the latter are the fundamental bearers of phenomenal properties. These non-physical individuals are sometimes called "souls," though I will call them "immaterial subjects" to avoid the religious connotations of "soul." The dual-aspect theory combines property dualism with substance monism: every concrete individual is physical (roughly, wholly decomposable into fundamental physical entities), but some concrete individuals have, in addition to their physical properties, irreducible phenomenal properties. On this view, the fundamental bearer of your phenomenal

³ Cf. Goff (2017), Strawson (2006).

properties will be some macroscopic physical thing, such as a human animal that is exhaustively composed of physical particles.

I will argue that both forms of dualism can avoid commitment to many subjects without the costs that physicalists must incur to do so. Of course, neither form of dualism automatically avoids the problem of too many subjects, since one can easily describe psychophysical laws that would generate too many subjects. But one can also describe psychophysical laws that don't, as I'll show below. Since it is reasonable to hold that there is only one subject in your vicinity, it is reasonable for the dualist to conjecture that laws of the latter kind prevail in our world (assuming such laws aren't much more complex or otherwise objectionable by comparison). In my view, the many-subjects argument does not strongly favor either form of dualism over the other.⁴ However, my tentative preference is for substance dualism, partly because I think its solutions to many-subjects puzzles are, on balance, more satisfying, and partly because I think substance dualism is more intrinsically plausible than the dual-aspect theory. (After all, it's natural to suppose that fundamental properties attach to fundamental individuals, and a simple immaterial subject is a more plausible candidate for a fundamental individual than a complex physical organism or brain composed of trillions of particles.⁵) But it is not my goal in this paper to justify this tentative preference, but instead to argue, in non-partisan fashion, that the many-subjects argument lends support to some or other kind of dualism.

The plan: §2 poses the threat of conscious parts, §3 the threat of conscious coinciders, and §4 the threat of conscious person-candidates. In §5, I consider how the arguments fare under various restrictive ontological schemes.

2. The Threat of Conscious Parts

⁴ Contra Zimmerman (2010) and Unger (2006).

⁵ Cf. Cutter (forthcoming)

The first argument raises the threat of conscious parts. Focusing on the brain as a candidate conscious part, we can formulate the argument as follows:

The Brain Argument

Uniqueness: There is exactly one conscious being in your vicinity—you.

Distinctness: You are not your brain.

Conscious Part: If physicalism is true, then your brain is conscious.

C. Therefore, physicalism is false.

Variants of this argument substitute other parts for your brain. An important variant, which I will consider below, focuses on large gerrymandered parts of you, such as your “finger-complement,” (something composed of all your matter except for that in your left pinky), or large gerrymandered parts of your brain, such as a “neuron complement” (all of your brain except for some arbitrarily chosen neuron). What matters for this argument is, roughly, that the part be chosen in such a way that an intrinsic duplicate of the part could support consciousness on its own.⁶

A brief point of clarification: some philosophers draw a distinction between being conscious in a non-derivative sense and being conscious in a derivative sense (alternatively, between fundamental and derivative bearers of consciousness and other mental properties). If one finds this distinction intelligible and helpful, read “conscious” in this argument and those below to mean consciousness in the non-derivative sense. Thus, for example, Uniqueness should be understood as entailing that you are something that is conscious in a non-derivative sense, or

⁶ Merricks (2001) defends a closely related argument against a doctrine of “microphysical supervenience.” Parfit (2012) defends and Olson (2007) develops (without endorsing) a related argument against animalism (and for the view that you are you brain).

as Chisholm (1989: 125) puts it: “If I happen to be feeling sad, then, surely, there is no *other* thing that is doing my feeling sad for me.”

Let us consider each premise in turn. Uniqueness is a piece of common sense, at least if “your vicinity” is understood narrowly enough (e.g., “in your chair,” not “within a one-mile radius.”). You can know with something approaching Cartesian certainty that *you* are conscious, so there must be at least one conscious being in your vicinity. And it is arguably a Moorean fact (“one of those things that we know better than we know the premises of any philosophical argument to the contrary” (Lewis 1996: 549)) that there aren’t multiple conscious beings in your vicinity. More cautiously: there aren’t multiple such beings with rich, complex experiences of the kind you presently enjoy. This weaker claim, which would suffice for the purposes of the argument given unproblematic adjustments to other premises, is compatible with panpsychist views according to which you are made up of many subatomic parts, each with simple experiences very different from your own.

A further argument for Uniqueness is that denying it leads to implausible revisionary normative conclusions. A key assumption of this argument is that consciousness is directly significant to *welfare*. Subjects of consciousness are welfare subjects—things can go well or badly for them—and phenomenal states directly contribute to a subject’s welfare or level of well-being. For example, pleasant and unpleasant experiences can directly make one better or worse off. Since it matters morally how our actions affect the welfare of others, and how many are thus affected, it matters morally how many subjects will (say) experience pain as a result of our actions.

One might think that, while accepting a multiplicity of subjects will force us to revise our estimates of the number of subjects whose welfare is affected by our actions, it won’t affect the

comparative evaluation of actions, and therefore won't lead to revisionary conclusions about what to do. Suppose that commonsense counting says that action A alleviates the pain of one subject and action B alleviates the pain of two subjects, so B looks preferable. If instead there are two subjects (say, an animal and a brain) wherever we previously thought there was one, then A really alleviates the pain of *two* and B alleviates the pain of *four*, but B still comes out as the morally preferred option. Generalizing, it may seem that revisionary counting will never lead to revisionary verdicts about what we should do.

The generalization is too quick. First, there is no guarantee that the revised counting schemes will simply scale all the numbers by a common factor. With some many-subject challenges, like those invoking large gerrymandered parts (finger-complements, neuron-complements, and the like), or the many person-candidates to be considered below (§4), larger people will tend to have vastly more of the relevant parts or person-candidates than smaller people for reasons of basic combinatorics. Suppose that, by the lights of common sense, action A relieves the suffering of one large person and action B relieves the (equally intense) suffering of one small person. A Uniqueness denier will say that A relieves the suffering of *vastly* more subjects than does B. As Jonathan Simon (2017a: 452) observes, this would seem to invite a revisionist moral conclusion given the following plausible principle:

Hedonic Beneficence: If option ϕ relieves the pain of n experiencers, option ψ equally relieves the pain of m experiencers, and $n < m$, then other things equal one ought to do ψ rather than ϕ .

Moreover, even if all the numbers are scaled by a common factor (which might be the case if brains were the only relevant conscious parts), we still get normative revisionism once we move away from the most simplistic forms of consequentialism. Commonsense morality and

many deontologists accept *prerogatives* and *constraints*. Prerogatives make it permissible to perform an action that leads to a suboptimal outcome. Constraints make it impermissible to perform an action with an optimal outcome.⁷ It's widely held that prerogatives and constraints are typically defeasible when the consequentialist stakes are sufficiently high. For example, it's not permissible to break a promise if keeping the promise has marginally worse consequences than breaking it, but perhaps it is permissible (even obligatory) if keeping the promise has vastly worse consequences. Likewise, perhaps it's not obligatory to allow myself to be tortured to prevent two others from being tortured, but it may be obligatory if it's the only way to prevent a million others from being tortured. Wherever the defeasibility thresholds are set, a revisionary view of the number of subjects affected by an act can lead to a revisionary view about whether the threshold is met, and thus a revisionary verdict about what we may or must do.⁸

A Uniqueness-denier might respond that the many subjects share token experiences, and that what matters morally is the number of experiences (e.g., pains and pleasures), not the number of subjects. This response suggests a view on which subjects are mere "value receptacles," morally significant only for the valuable experiences they host.⁹ I find this moral claim implausible. If A experiences an intense pain, wouldn't it be worse to make B experience the same token pain? But in any case, we can give parallel arguments that, if physicalism is true, there are many experiences where we ordinarily think there is one. On one view, experiences are instantiations of phenomenal properties by subjects at times, individuated by the phenomenal properties, subjects, and times involved. (So A's having Q at t = B's having G at t* iff A = B, F

⁷ For discussion, see Kagan (1998: ch. 3, 5).

⁸ For a more comprehensive survey of the potential revisionary normative consequences of denying Uniqueness, see Crummett (2022: sect. 5). Further arguments for Uniqueness include Olson's (2007) epistemic argument (roughly, if Uniqueness were false, we couldn't know which thing we are), Builes and Hare's (2023) whale argument (roughly, uniqueness-denying metaphysical views falsely predict that I should be part of a whale), and Unger's (2006) free-will argument (roughly, Uniqueness is entailed by our freedom of choice). Cf. Hudson (2001: 39-44).

⁹ Cf. Singer (1993: 121), Yetter-Chappell (2015).

= G, and $t = t^*$.¹⁰) On this view, distinct subjects entails distinct experiences (independently of physicalism). On another view, an experience is a concrete, localized going-on in your brain, a complex physical occurrence composed of many smaller occurrences—individual neural firings, neurotransmitter releases, chemical reactions, and ultimately fundamental microphysical occurrences like particle motions. For any such event e that a physicalist might reasonably identify with my token pain, we can consider a slightly smaller event (e.g., e -minus-the-movement-of-some-particle). Arguments analogous to those given below for Conscious Part would suggest that this “ e -minus” is also a pain, yielding multiple distinct pains.

One could instead deny Uniqueness by claiming that, although there is only one (non-derivatively) conscious being in your vicinity, that conscious being isn't you. Rather, it's a proper part of you: your brain. At best you are conscious in a derivative sense, in virtue of having a conscious brain as a part.¹¹ This response rejects the assumption that you are conscious in a non-derivative sense. But this assumption, while attractive, is not sacrosanct. (In my view, substance dualists should be open to a parallel view on which the immaterial subject that is non-derivatively conscious is a proper part of you, alongside your body, while you are only conscious in a derivative sense, in virtue of the consciousness of your immaterial part.) Those who accept that the brain is (non-derivatively) conscious should consider a variant of the argument that invokes your “neuron-complement.” Below I'll suggest that if physicalism is true, then your neuron-complement is (non-derivatively) conscious. So even if your brain is (non-derivatively) conscious and you aren't, we still have too many (non-derivatively) conscious beings.

¹⁰ Cf. Kim (1976).

¹¹ See Lee (2017) for the closely related view that the metaphysical subject of your experiences is not you, but a part of you (e.g., parts of your brain, perhaps different parts for different experiences).

One might think there is something methodologically backwards about arguing for a descriptive metaphysical conclusion (here, that there is one subject in your vicinity, and—ultimately—that physicalism is false) on the grounds that supposing otherwise has counterintuitive normative implications. This is a bit like arguing for the descriptive conclusion that recycling helps the environment on the grounds that if it didn't, it wouldn't be true that we ought to recycle, and this would conflict with our prior normative beliefs. Clearly the descriptive question about recycling's impact on the environment should be settled independently of our prior normative beliefs about the goodness of recycling, and our normative beliefs should be adjusted to fit our descriptive conclusions. So too, one might think, the question of how many subjects are in my vicinity, or the question of whether physicalism is true, should be settled independently of our normative beliefs (say, about whether we ought to give preference to the large over the small).

While I agree that normative-to-descriptive reasoning is often illegitimate, as in the recycling case, I reject the tempting generalization that it is never legitimate to revise or form descriptive beliefs to achieve coherence with normative beliefs, especially when the normative beliefs are foundational to our normative thought. Philosophy, I think, is largely a matter of achieving reflective equilibrium within one's total system of beliefs. Just as reflective equilibrium does not give absolute priority to particular case judgments over judgments about general principles (or *vice versa*), I do not think descriptive judgments must be given absolute priority over normative judgments in philosophy. I'm sympathetic to Lewis's (1973: 88) suggestion that the job of philosophy is to conservatively adjust and expand our prior beliefs into an orderly system. If one metaphysical view coheres more readily into an orderly system with our foundational normative beliefs than its alternatives, this can be a legitimate (defeasible)

reason to accept that metaphysical view. For example, I think it can be reasonable to reject the metaphysical view that all that exists are Locke's "masses of atoms"—material bodies that have all their parts essentially—partly on the grounds that this view coheres very badly with deep-seated normative beliefs tied to identity over time (e.g., that I am bound by promises made last month, or responsible for actions performed last month).

Let's turn to the Distinctness premise: you are not identical to your brain. Distinctness can be motivated by Leibniz's law arguments. Olson (2007: 85) argues that you and your brain have different persistence conditions. Your brain can, and likely will, exist for some time after your death, after it ceases to function (perhaps it will be preserved in a jar). But you wouldn't exist at that time. (This assumes no afterlife. If you could exist in a disembodied state in the afterlife, we can reverse the argument: you could exist after your brain ceases to exist.) Therefore, you aren't your brain, for your persistence conditions are different. We can also give Leibniz's law arguments involving size and shape properties. Your brain is about four inches tall and three pounds, but you aren't.

Granted, this would be an odd argument from a substance dualist who thinks you are a heightless, weightless immaterial subject (presumably the driving concern is that you are taller and heavier than your brain). But this argument is available to many dual-aspect theorists, like animalist dual-aspect theorists (i.e., a dual-aspect theorist who holds that you are an animal). In my view, the size/shape objection is inconclusive. A proponent of the "brain view" (the view that you are your brain) can respond that, when we say that you weigh more than 100 pounds, this is a case of derivative predication, or perhaps a case of semantic flexibility in personal pronouns, as when we say "I'm touching the curb" when my car is, or "I'm covered in mustard" when my

clothes are.¹² Those who take the brain view seriously should consider a variant of the argument that replaces “your brain” with “your neuron-complement.” The distinctness premise in this version of the argument is unassailable, for it’s obvious that you aren’t identical to your-brain-minus-one-neuron.

Finally, let’s consider the Conscious Part premise: if physicalism is true, then your brain is conscious. Conscious Part follows from two plausible claims:

Intrinsic Duplication: If physicalism is true, then your brain is an intrinsic duplicate of some (possible) conscious being.

Intrinsicity: Intrinsic duplicates are alike with respect to whether they are conscious.

In support of Intrinsic Duplication, consider Brainy, your brain-in-vat counterpart on Twin Earth. Brainy is an intrinsic physical duplicate of your brain, which is artificially stimulated by a computer. If physicalism is true, it’s plausible that Brainy is conscious. For it’s plausible that something is conscious in this scenario, and given physicalism, there is no immaterial entity that could be the subject of this consciousness. The subject of consciousness would have to be a physical entity, and Brainy would seem the most natural candidate. Furthermore, if physicalism is true, Brainy is an intrinsic duplicate of your brain. Brainy perfectly resembles your brain in all intrinsic *physical* respects, and given physicalism, there are no other respects. (A dual-aspect theorist will say that there are phenomenal respects of similarity over and above physical respects, and she can say that your brain is nothing like Brainy in phenomenal respects. Perhaps Brainy has rich technicolor phenomenology while your brain has no phenomenal properties whatsoever, since your phenomenology attaches to a larger material substance, such as a human animal. This is why Intrinsic Duplication takes the form of a conditional.) For the “neuron-

¹² Cf. Parfit (2012: 20-1).

complement” version of the argument, let Brainy have one fewer neuron than your brain, and have the computer make compensating adjustments elsewhere in the brain so Brainy remains intrinsically like your neuron-complement.

Intrinsicality is also very plausible. Hawthorne (2004: 352) describes a roughly equivalent claim as obvious.

If a spatio-temporal region wholly contains me and I am conscious and some other spatio-temporal region does not contain a conscious being, then it is altogether obvious that the two regions are not duplicates.

Intrinsicality follows from the widely held view that consciousness is an intrinsic property, since intrinsic duplicates share all their (non-haccetistic/qualitative) intrinsic properties.¹³ Moreover, while some philosophers would deny Intrinsicality, the argument can be revised to accommodate prominent views on which consciousness is extrinsic, such as externalist representationalist views that make consciousness depend on our historical extrinsic connections to environmental properties (Dretske 1995, Tye 1995). To accommodate these theories, we could weaken Intrinsicality to the claim that intrinsic duplicates *that also share the relevant extrinsic historical properties* are alike with respect to consciousness. Intrinsic Duplication could then be strengthened to say that if physicalism is true, there is a possible conscious being that is an intrinsic duplicate of your brain *and* shares with your brain relevant historical properties (like having states that historically causally covary with colors and shapes). In support of this strengthened claim, we’d consider recently-envatted Brainy, an envatted brain that came from a living organism with the right historical/evolutionary connection to external properties.

¹³ Cf. Mørch (2019: 134), Merricks (2001), Pautz (2013).

Instead of turning to standard forms of phenomenal externalism, a physicalist might reject Intrinsicity by claiming, with Ted Sider (2003), that consciousness is a *maximal* property, not an intrinsic property. Roughly, *F* is a maximal property iff (sufficiently large) proper parts of *F*s aren't *F*s. Following Sider's stock example, consider the property of being a house. Here is an argument that being a house is not intrinsic. All-of-my-house-except-this-window is not a house. (Otherwise there would be a multitude of houses in the vicinity of my house.) But this large part of my house is an intrinsic duplicate of some (possible) house. (Consider a house just like mine, but with one window removed.) Since intrinsic duplicates agree on their intrinsic properties, it follows that being a house is not intrinsic. Rather, it is maximal. For Sider, to be a house is (i) to have a certain intrinsic property—being a house*—and (ii) to not be a proper part of any house*. Similarly, Sider thinks, to be conscious is (i) to have a certain intrinsic property, which he calls *consciousness**, and (ii) to not be a (sufficiently large) part of anything else that has *consciousness**. That is, consciousness is equivalent to a conjunctive property with an intrinsic component and a (negative) extrinsic component. For Sider, the human animal, its brain, finger-complement, and neuron-complement are all *conscious**. But only the animal is *conscious*, because only the animal fails to be a part of another *conscious** thing. (Likewise, *mutatis mutandis* for more specific phenomenal properties like feeling pain and having a reddish experience.)

This view can be motivated by a metasemantic principle of charity. Roughly, our statements about consciousness, specifically our statements about how many conscious subjects there are in a given situation, come out true if “conscious” is interpreted as expressing a maximal property rather than an intrinsic property. A suitable charity principle might then recommend the conclusion that “conscious” expresses a maximal property.

I think the maximality response is a wretched subterfuge, to borrow Kant's phrase. Even if phenomenal properties like being in excruciating pain have an intrinsic component (excruciating-pain*) and an extrinsic component (not being a proper part of anything in excruciating pain*), it is the intrinsic component that is normatively significant. In particular, it is the intrinsic component that directly contributes to one's welfare. Being intrinsically just like someone in excruciating pain is just as bad as being in excruciating pain.¹⁴ The mere failure of an individual to be a proper part of something in the relevant intrinsic state doesn't make things dramatically worse for that individual. Even if metasemantic principles imply that an individual doesn't count as being in the extension of "is in excruciating pain," that's cold comfort for that individual if it's intrinsically just like someone in excruciating pain. So, even if the maximality response can secure the result that only one thing in your vicinity counts as conscious, it still leads to the kind of normative revisionism that comes with the denial of Uniqueness.

The substance dualist can easily avoid this threat of too many subjects. He can say that the only thing that is (non-derivatively) conscious is the simple immaterial subject. He might say that, for some intrinsic physical property F, both your body's having F and your brain's having F are causally sufficient for the (single) immaterial subject to be conscious. In typical circumstances, both conditions are met—a benign case of overdetermination. Another option for the substance dualist is to say that one's consciousness is only directly causally sensitive to the intrinsic physical states of one's brain (without thereby having to say that the brain is conscious, or that you are your brain).

The dual-aspect theorist can also avoid the threat of conscious parts. She can say that, while consciousness is itself an intrinsic property of a physical individual, the psychophysical

¹⁴ Cf. Crummett (2022: 321).

laws responsible for the generation of consciousness operate on maximal physical properties, not intrinsic physical properties.¹⁵ For a given intrinsic property F, we can define a corresponding maximal property: the property of being an F that isn't part of any other F. The animal and the brain (and finger-complement, neuron-complement, etc.) have the relevant intrinsic physical property F. But only the animal has the corresponding maximal property. A dual-aspect theory that holds that the physical inputs to the psychophysical laws are maximal physical properties can deliver the result that, in typical circumstances, the animal (and not the brain) is the unique bearer of conscious, but in envatment cases, where a suitably active brain is disembodied, consciousness attaches to the brain (or something coincident with the brain). Of course, a physicalist can equally avoid the threat of conscious parts by *identifying* consciousness with this same maximal physical property, as we've seen. The key difference is that the dual-aspect theory just described needn't give up on the plausible idea that consciousness is *itself* an intrinsic property, nor does it lead to implausible revisionary normative consequences mentioned above. On this dual-aspect view, when you are in excruciating pain, there aren't two things that are intrinsically just like something in excruciating pain.

It may seem that the Brain Argument is vulnerable to a parody argument. One can give a similar argument that, if physicalism about (say) tablehood is true, then there are many tables where we ordinarily think there is one. (Replace "conscious being" with "table," "consciousness" with "table," "you" with "this table," and so forth. In place of your brain, we consider a large proper part of the table that is intrinsically just like some possible table, like table-minus-a-chunk-in-the-corner.) This argument must be unsound because dualism about tablehood is

¹⁵ Cf. Morch's (2019: sect. 4.3) suggestion that the information-integration theory of consciousness might be reconciled with the intrinsicality of consciousness if we say that the extrinsic physical property of max- Φ is merely the causal basis for consciousness, while consciousness is itself intrinsic and irreducible to max- Φ .

absurd, so I had better find a disanalogy. The key disanalogy, in my view, is that the maximality response succeeds in this case. Being a table is a maximal property, not an intrinsic property. Indeed, I think this is a conceptual truth. It is *a priori* that if something intrinsically suited to be a table is seamlessly embedded in a slightly larger thing that is intrinsically suited to be a table, then the first thing is not a table. (Given the epistemic gap between physical/spatial notions and phenomenal notions, the corresponding claim about consciousness is certainly not a conceptual truth, *pace* Sider (2003: 144-5).) As with the maximality response to the Brain Argument, this view implies that there are many things in the vicinity of my table with a property that has roughly the same welfare significance as being a table. That is, being a table* makes roughly the same contribution to something's welfare as does being a table. But this doesn't lead to normative revisionism, since being a table has no direct normative significance (a thing is not made better or worse off in virtue of being a table).¹⁶

Another potential disanalogy worth noting is that, with tables, the nihilist and pluralist options are, if not plausible, at least *more* plausible than nihilism or pluralism about subjects. That is, it is more plausible to hold that, strictly speaking, there are either no tables (Unger 1980) or many (Lewis 1999, Williams 2006), and it is merely a useful fiction to suppose there is one, than to make analogous claims about subjects. In my view, we can know with Cartesian certainty that subject nihilism is false, but we lack similar justification to reject table nihilism. And subject pluralism faces the normative-revisionism objection, which we've seen has no bite against table pluralism.

3. The Threat of Conscious Coinciders

¹⁶ Cf. Simon (2017a: 458)

The next version of the many-subjects argument raises the threat of conscious coinciders. Say that x and y are coincident at t just in case, at t , x and y occupy the same spatial region and share their matter. Many philosophers hold that there can be distinct coincident objects, the canonical example being a statue and the lump of clay that constitutes it. The statue and the lump differ in historical and modal respects. The clay has been around longer than the statue, and they clay would survive squashing while the statue would not. It would therefore seem to follow, by Leibniz's law, that the statue and the clay are distinct.

I assume for the argument below that, if physicalism is true, then you are some physical thing that coincides with your body. (Perhaps you are an animal, or at least something with the same size and shape as an animal.) But this argument could easily be recast to target physicalist views on which you coincide with your brain.

The Body Argument

Uniqueness: There is exactly one conscious being in your vicinity—you.

Distinctness: You are not your body.

Conscious Coincider: If physicalism is true, your body is conscious.

C. Therefore, physicalism is false.¹⁷

Variants of this argument replace “your body” with different candidate coincident objects. For example, on some views you coincide with an “aggregate of particles,” something composed of the very particles that currently compose you, and which is essentially so composed. On more abundant ontologies, there may be things coincident with you for arbitrary intervals (e.g., from

¹⁷ Olson (2004) presents (without endorsing) a closely related “thinking body argument,” though his is an argument against animalism rather than physicalism.

noon yesterday to noon tomorrow), but which don't exist before or after, or which coincide with you as long as you are sitting, but coincide with LeBron James whenever you aren't.¹⁸

The Uniqueness premise is identical to the Uniqueness premise in the Brain Argument. I won't repeat the motivations given above. As before, one could affirm that there is exactly one (non-derivatively) conscious being in your vicinity, but deny that it is you. For example, a four-dimensionalist might say that you are conscious in a derivative way, in virtue of having a non-derivatively conscious person-stage as a temporal part. A four-dimensionalist can say that both you and your body are derivatively conscious, but only one currently existing thing in your vicinity is non-derivatively conscious—namely, a person-stage that is a shared temporal part of both you and your body (where you and your body are distinct but largely overlapping “spacetime worms”). While the four-dimensionalist physicalist may have a good response to the Body Argument, I think four-dimensionalism should be rejected on the basis of Mark Johnston's (2016a, 2016b) “personite” objection.¹⁹ The gist of the personite objection is that four-dimensionalism implies that you contain many “personites”—temporal parts that are, if not quite persons, sufficiently person-like to possess significant moral status (e.g., many personites are intrinsic duplicates of possible persons). Johnston argues that this leads to implausible revisionary normative conclusions (e.g., that it's wrong to accept a significant burden now for a larger benefit in the future, since this would wrong a personite who ceases to exist before the benefits are realized). This objection is structurally similar to the many-subjects argument against physicalism (specifically the threat of conscious parts), and I think it gives us strong reason to reject four-dimensionalism. I will therefore set aside the four-dimensionalist response.

¹⁸ Closely related arguments, though often with somewhat different conclusions, are defended or discussed by Olson (2010), Zimmerman (2003), Briggs and Nolan (2015), Kovaks (2016), and Baker 2000.

¹⁹ See also Taylor (2013) and Olson (2010) for similar objections.

The argument for Distinctness is a Leibniz's law argument similar to the argument for the distinctness of clay and statue. When you die, your body will continue to exist for some time as a corpse (unless you die in an especially dramatic fashion). But your death will be the end of you (barring an afterlife, in which case we can reach the same conclusion from the premise that you'll exist when your body doesn't). So you and your body differ in their temporal properties. You also differ in related modal properties. Your body would continue to exist if your heart were to stop beating now, but you wouldn't.²⁰

It remains to defend the Conscious Coincider premise. Olson (2004: 266) presents (without endorsing) the following argument for a similar claim:

Now your body ought to be able to think. It is physically indistinguishable from you, with the same surroundings and causal history. What could prevent it from using its brain to think?

Substituting consciousness for thought, we might similarly argue that, given physicalism, your body should be conscious, since it is physically indistinguishable from you. However, even on the assumption that you are a material thing that coincides with your body, we can resist the claim that you and your body are physically indistinguishable. The differences between you and your body mentioned above seem to be physical differences. Your body (currently) has the property that it will one day lie underground in a coffin, while you lack this property. I see no good reason not to call this a physical property. In at least one good sense of "physical properties," you and your body don't share all your physical properties.

²⁰ It's well known that modal and temporal counterpart theorists have resources to resist arguments of this kind. I won't relitigate the well-trodden debates about counterpart theory here, though see Hawthorne, Dorr, and Yli-Vakkuri (2021: ch. 10) for objections to which I am largely sympathetic. Note also that modal and temporal counterpart theory is powerless to resist the Distinctness premises of the other arguments in this paper.

But there is a truth in the vicinity of Olson’s physical-indistinguishability claim. If you coincide with your body, you and your body (currently) share all your *temporally intrinsic* and *categorical* physical properties. You physically differ only in your temporally extrinsic and “hypothetical” (or modal) physical properties. Roughly and intuitively, your temporally extrinsic properties are those you have at least partly in virtue of what happens at other times, whereas your temporally intrinsic properties are those you have wholly in virtue of what’s going on right now. When we specify the nature of a temporally intrinsic property—when we say what it is to have the property—we needn’t make reference to other times, to what *will* or *did* happen. Analogously (and somewhat more roughly), hypothetical or modal properties are those you have at least partly in virtue of what happens in other *worlds*, while categorical properties are those you have wholly in virtue of what’s going on in the actual world (cf. Yablo 1992). In specifying the nature of a categorical property—in saying what it is to have a given categorical property—we needn’t invoke modal notions; we needn’t make reference to what *would* or *could*, or *would-probably*, or *must* happen.

Relying on this restricted physical-indistinguishability claim, we can argue for the Conscious Coincider premise from two claims, which are structurally parallel to those used to argue for the Conscious Part premise in the Brain Argument:

Categorical Duplication: If physicalism is true, your body is exactly like some conscious being (namely, you) in categorical and temporally intrinsic respects.

Categoricity: Any two things exactly alike in categorical and temporally intrinsic respects are alike with respect to consciousness.

In support of Categorical Duplication: we’ve seen that, on the assumption that you are a material thing that coincides with a body, you and your body are exactly alike in categorical and

temporally intrinsic physical respects. Given physicalism, exact similarity in categorical and temporally intrinsic *physical* respects amounts to exact similarity in categorical and temporally intrinsic respects *period*. (It's important that Categorical Duplication is conditional on physicalism. A dual-aspect theorist can say that you and your body are alike in all categorical and temporally intrinsic *physical* respects, but not in all categorical and temporally intrinsic respects *period*, because there are non-physical categorical and temporally intrinsic properties—phenomenal properties—that you have and your body lacks. More on this proposal below.)

Categoricity is also very plausible. Consciousness and determinate modes of consciousness, such as experiencing red or feeling pain, seem to be categorical and temporally intrinsic properties. The way it feels, right now, to be you does not constitutively depend on what will happen decades hence, nor does it constitutively depend on your persistence conditions. Phenomenal properties are in this respect like shapes. Any two things that differ *only* in hypothetical and temporally extrinsic respects, like the statue and the lump, must be alike with respect to their shapes.²¹

Contrast phenomenal and shape properties with properties like *being a statue*. Part of what it is to be a statue is to have certain persistence conditions. (A complete answer to “what makes it the case that this is a statue?” can't *just* cite temporally intrinsic and categorical properties, since these are shared by the lump, which is not a statue. A complete answer will have to mention hypothetical features.) But it seems plainly false that part of what it is to feel pain is to have certain persistence conditions. A mere hypothetical difference between two things

²¹ Cf. Cutter (2021: 402-3).

can therefore constitute a difference in whether they are statues, but cannot constitute a difference in whether they feel pain.²²

A physicalist might say that consciousness and its determinates are “mixed” properties. To be in pain is to have such-and-such categorical, temporally intrinsic property (*consciousness*) and so-and-so hypothetical property or temporally extrinsic property (e.g., having animal-like persistence conditions rather than body-like persistence conditions).²³ Like the maximality view considered above, this view might be motivated by a metasemantic charity principle. We are disposed to say things like “there is exactly one conscious being over there,” so charitable interpretations will assign to “conscious” a referent that makes these statements come out true, and only a partly hypothetical property will do the trick.

I respond as before that this is a wretched subterfuge. Even if consciousness is a mixed property, it is the categorical and temporally intrinsic component that is directly significant to welfare. If *x* is in excruciating pain, and *y* is exactly like *x* in categorical and temporally intrinsic respects, then *y* is in a state that is just as bad as being in excruciating pain. The mere fact that an individual has persistence conditions according to which it ceases to exist when vital functioning stops, rather than upon physical disintegration, doesn’t make this state dramatically worse for that individual. Even if metasemantic principles imply that something doesn’t count as being in the extension of “is in excruciating pain,” that’s cold comfort for that individual if it’s categorically just like someone in excruciating pain. So, even if the mixed-property response can

²² For the reasons suggested in the introduction, I think this claim is especially hard for a Russellian physicalist to deny. “Standard” physicalists might deny it, since many standard physicalists think of consciousness as a non-categorical role-functional property. However, functionalist views of consciousness don’t seem to avoid the threat of conscious coinciders, since properties like “having a representation poised to play a certain role” doesn’t seem like a property that divides you and your body.

²³ Hawthorne, Dorr, and Yli-Vakkuri (2021: 320, note 15) suggest a view along these lines as an alternative to Sider’s maximality response to Merricks’ (2001) argument against microphysical supervenience. Shoemaker (1984) defends a similar view: that whether something is *thinking* constitutively depends on its persistence conditions.

secure the result that only one thing in your vicinity counts as conscious, it still leads to the kind of normative revisionism that comes with the denial of Uniqueness.

The substance dualist can easily avoid this threat of too many subjects. Even if the physical nomological basis (call it “P”) for consciousness is a temporally intrinsic and categorical property shared by both a physical animal and a distinct, coinciding body, it doesn’t follow that two immaterial subjects will be conscious as a result. We can say that the animal’s having P and the body’s having P are both nomologically sufficient for a (single) immaterial subject to be conscious: a benign kind of overdetermination.

What about the dual-aspect theorist? Suppose the psychophysical laws say: for any x , if x has physical property P, then x is conscious. If P is a temporally intrinsic and categorical property, then both you and your body will be conscious. But the dual-aspect theorist can reasonably conjecture that P is not a temporally intrinsic and categorical property. The psychophysical laws may be sensitive to hypothetical distinctions (or temporally extrinsic distinctions, though I find that less plausible). In effect, she can conjecture that P, the nomological basis of consciousness, is not a purely categorical property. Perhaps P is a “mixed property,” equivalent to the conjunction of such-and-such categorical property and so-and-so hypothetical property.

Of course, we’ve seen that the physicalist can try to avoid the threat by saying that consciousness *itself* is a mixed physical property. But there are two important differences between the claim that a mixed property is the nomological basis of consciousness and the claim that the relevant mixed property *is* (or grounds) consciousness. The former view does not violate the intuition that consciousness is *itself* categorical—that it doesn’t *constitutively* depend on hypothetical factors. Nor does it lead to implausible normative revisionism, as the mixed

physicalist view does. On the former view, the mixed physical state that serves as the nomological basis of consciousness makes no direct contribution to welfare. Its contribution is purely instrumental. When you are in excruciating pain, there are not two things in a state with the same direct welfare-significance as excruciating pain.

As before, one can give a parody argument that, if physicalism about tablehood is true, then there are many tables where we ordinarily think there is one. (Again, replace “conscious being” with “table,” “consciousness” with “table,” and “you” with “this table.” In place of your body, we might consider the chunk of wood that constitutes the table.) This argument must be unsound because dualism about tablehood is absurd, so I had better find a disanalogy. The key disanalogy, in my view, is that the mixed-property response succeeds in this case. Being a table isn’t a categorical property, since part of what it is to be a table is to have certain persistence conditions. As with the mixed-property response to the Body Argument, this view implies that there are many things in the vicinity of my table with a property that has roughly the same normative significance as being a table (e.g., being a table’ makes roughly the same contribution to something’s welfare as does being a table). But this doesn’t lead to normative revisionism, since being a table has no direct normative significance.

4. The Threat of Conscious Person-Candidates

A final version of the many-subjects argument involves what I’ll call “person candidates.” Near the boundaries of your body are various “questionable parts,” material things for which it is unclear whether they count as parts of you. These might include skin cells in the process of dying and shedding, loose hairs that haven’t yet cleanly separated from the skin, and the countless atoms midway through the process of being incorporated into your body through digestion or respiration, or the process of being expelled from your body through excretion,

perspiration, or metabolism. For each questionable atom, we can find many human-body-shaped collection of atoms that include it, and many that exclude it. Assuming each such collection composes something, there will be many numerically distinct human-body shaped things in your vicinity, a cloud of overlapping bodies differing very slightly in composition. We'll call them "person candidates" (or "candidates" for short). From here we can argue as follows:

The Person-Candidates Argument

Uniqueness: There is only one conscious being in your vicinity—you.

At-Least-One: If physicalism is true, then at least one candidate is conscious.

Physical Egalitarianism: If physicalism is true, and at least one candidate is conscious, then multiple candidates are conscious.

C. Therefore, physicalism is false.²⁴

The motivations for Uniqueness have already been given. The motivation for At-Least-One is that you are conscious, and given physicalism, it seems that there's nothing else for you to be but one of the candidates. You couldn't be an immaterial thing, and it seems there is no material thing distinct from each of the candidates that could plausibly be identified with you. (If you think you are a material thing much smaller than a human body, like a brain, the same points would apply, *mutatis mutandis*, to brain candidates.)

The final premise is Physical Egalitarianism, for which I have four arguments. First: for any chosen candidate C1, there is another candidate C2 that, in *physical* respects, seems equally qualified to be a bearer of consciousness. Given physicalism, whether something is conscious is

²⁴ Closely related arguments are defended by Unger (2006), Zimmerman (2010), and Simon (2017a), though Unger's argument is for substance dualism specifically rather than dualism (or non-physicalism) generally, and Zimmerman's is for the disjunction of substance dualism and the "speculative materialist" view that you are a material thing, but not a "garden variety" object like an animal, body, or brain. Of the three, Simon's argument is most similar to mine, both in its conclusion and in its appeal to normative claims to motivate the premises, though he is more open than I am to property dualist views that deny Uniqueness (2017a: 464).

simply a matter of how it is in physical respects, so if two things are equally qualified in physical respects to be conscious, then either both or neither are conscious.

A second argument for Physical Egalitarianism relies on the attractive idea that, even if we allow that phenomenal states might be fully constituted by, or grounded in, physical states, we should not allow *massive discontinuities* in physical-phenomenal grounding. Really big phenomenal differences, like the difference between being in excruciating pain and feeling nothing at all, or the difference between having vivid technicolor experience and being a zombie, can't be constituted by miniscule physical differences. More generally, if phenomenal states are grounded in some more basic set of states (whether physical, protophenomenal, or whatever), big phenomenal differences of the kind just mentioned can't consist in miniscule differences at the underlying level. This is a "proportionality" principle: differences at the higher level shouldn't be wildly out of proportion with the underlying differences that ground or constitute them. This principle doesn't rule out the possibility of *causal* discontinuities, where two minutely different physical states cause very different phenomenal effects. (A dual aspect theorist may want to accept these here.) Causal discontinuities are a familiar phenomenon: n straws on a camel's back can have a very different effect from $n + 1$ straws. But it seems absurd to suppose that a trivial, minute difference—like the physical difference between one candidate and another—could itself *constitute* a phenomenal difference as radical as the difference between excruciating pain and zombiehood.²⁵ This would be like saying that the difference between a completely healthy, uninjured camel and a paralyzed camel with a shattered spine might simply *consist in* a miniscule microphysical difference.

²⁵ Cf. Merricks (1998: 845).

From physical-phenomenal proportionality, it is a short step to Physical Egalitarianism. If we think that at least one candidate is conscious, this is presumably because we think that at least one candidate must be conscious *in the way that you are*. Thus, at least one candidate, call it C1, undergoes vivid technicolor experience of the kind you presently enjoy. We can now find another candidate, C2, that in physical respects differs only minutely from C1. Suppose for *reductio* that, while C1 undergoes vivid technicolor experience, C2 lacks consciousness altogether. In that case, we have a *really big* phenomenal difference between C1 and C2, a phenomenal difference as large as that between you and a rock. Given physicalism, this phenomenal difference must be grounded in some physical difference, and given physical-phenomenal proportionality, a miniscule physical difference won't do. But there aren't any relevant non-miniscule physical differences between C1 and C2 that could do the trick. Therefore, given physicalism, C2 is conscious as well.

A third argument for Physical Egalitarianism relies on a structurally similar proportionality principle concerning normative properties and their descriptive grounds. This principle—call it descriptive-normative proportionality—says that a huge normative difference, like the difference between a situation that is extremely bad and one that is not bad at all, can't be grounded in or constituted by a miniscule fundamental descriptive difference.²⁶ To illustrate, suppose Peter's life is very good, and suppose that Paul's life is almost exactly identical to Paul's in all fundamental descriptive respects, intrinsic and extrinsic, differing at most in some small microscopic detail. Descriptive-normative proportionality allows that Paul's life might be *very slightly* better or worse than Paul's, but it rules out the possibility of Paul's life being *massively*

²⁶ Pautz (2017: 368) invokes a similar principle (the “small-difference principle”) in his “significance argument” against reductive physicalism. For discussion of how best to precisely formulate descriptive-normative proportionality principles, see Hawthorne (2022).

better or worse. This is immensely plausible. Normative differences shouldn't be wildly out of proportion with the underlying descriptive differences that ground them. Thus, if x is in a state that is extremely bad, and y is almost exactly the same as x in all fundamental descriptive respects, intrinsic and extrinsic, then y can't fail to be in a state that's bad.

From here, there's a plausible argument for Physical Egalitarianism. Suppose you're in excruciating pain. If at least one candidate is conscious, then at least one candidate is in excruciating pain. So, at least one candidate, call it C1, is in a state that has a very high degree of (final) disvalue. (I assume it is very bad to suffer excruciating pain.) Again, we find another candidate C2 that differs at most minutely from C1 in physical respects, and we suppose for reductio that C2 *isn't* conscious. It seems obvious that if C2 isn't conscious—if there's nothing it's like to be C2— then C2 is not in a state with anything like the same degree of (final) disvalue as being in excruciating pain. After all, C2 experiences no pain at all. So there is a big normative difference between C1 and C2. Normative differences are grounded in fundamental descriptive differences, and given physicalism, the only such differences we have to work with are physical differences. Given descriptive-normative proportionality, a miniscule physical difference won't do. But there aren't any relevant non-miniscule differences that could do the trick. Therefore, given physicalism, C2 is conscious.²⁷

My last argument for Physical Egalitarianism has to do with vagueness. Physicalists who reject Physical Egalitarianism are likely to hold that it is *vague* which candidate is the conscious one. In other words, they may say: it is determinately, definitely, non-vaguely the case that one and only one of the candidates is conscious. But it's vague whether Candidate 1 is conscious,

²⁷ A deflationary pluralist like Geoffrey Lee (2013) might say that C2, despite lacking experience, has a state that with the same normative significance as excruciating pain. This view has the virtue of upholding descriptive-normative proportionality, but leads to the kind of normative revisionism that comes with denials of Uniqueness.

vague whether Candidate 2 is conscious, and so on for all the other candidates. Call this the *vagueness thesis*. If one denies Physical Egalitarianism, the vagueness thesis is *prima facie* attractive. If each of the many candidates exists, and none is objectively distinguished by the unique possession of an outstandingly natural non-physical property (as the dual-aspect theorist can say), then it's hard to believe that our pronouns, names, and phenomenal predicates determinately and stably "lock on" to a unique candidate. Nevertheless, we should reject the vagueness thesis, as I shall argue. If the rejection of Physical Egalitarianism leads to the vagueness thesis, this is one more reason to accept Physical Egalitarianism.

I have two arguments against the vagueness thesis. The first targets versions of the thesis that accept an entailment from "it is vague whether A" to "it is indeterminate whether A," here taking "indeterminacy" to mean that there is "no fact of the matter," that A is neither true nor false. (This would include standard forms of supervaluationism, among many other views, but not, say, Williamson's (1994) epistemicism.) This version of the vagueness thesis implies an especially implausible kind of phenomenal indeterminacy. Some philosophers hold that it can never be indeterminate whether something is conscious. (Often, this claim is supported with an inconceivability argument: we cannot positively conceive of a borderline case of consciousness, and inconceivability is a guide to impossibility.²⁸) But we can distinguish more or less radical forms of phenomenal indeterminacy. For a simple creature like a garden snail, one might say that it is indeterminate whether it has a faint flicker of experience or no experience at all.²⁹ This would be a case of *mild* indeterminacy. For a more extreme kind of indeterminacy, suppose (to modify an example from Parfit (1995: 26)) we have an advanced AI that is behaving as though it

²⁸ See Tye (2021: 16) and Simon (2017b) for a related argument. See Cutter (2022) for discussion of inconceivability-to-impossibility/falsity inferences.

²⁹ Schwitzgebel (2020: sect. 4.3).

is in intense pain. We ask, “Is it conscious, and in great pain, or is it merely an insentient machine?” Suppose someone answers: “There’s no fact of the matter. It’s indeterminate whether it is experiencing intense pain or nothing at all. Either way of talking does an equally good job describing the facts.” If this answer is correct, it would be a case of *radical* phenomenal indeterminacy, a case where something is indeterminate between having some rich, intense, vivid form of experience and having no experience at all.

While I doubt that there are any cases of even mild phenomenal indeterminacy, radical phenomenal indeterminacy seems clearly impossible. It could not be indeterminate whether something feels intense pain or nothing at all. Likewise, it could not be indeterminate whether something has rich, vivid color experience of the sort you currently enjoy or no experience at all. The problem with the indeterminacy version of the vagueness thesis is that it implies that radical phenomenal indeterminacy is not only possible, but rampant in actuality. Each of the candidates in your vicinity is a case of radical phenomenal indeterminacy: for each candidate in your vicinity, it is indeterminate whether it has rich, vivid, multisensory experience or no experience at all.

My second objection targets an epistemicist version of the vagueness thesis, specifically one that adopts Williamson’s (1994) epistemic account of vagueness. A key feature of Williamson’s theory is that vagueness is associated with *semantic plasticity*. Where F is a vague predicate, the meaning (intension) of F is highly sensitive to miniscule, even imperceptible, changes in our dispositions to apply the predicate. Part of the explanation for why we cannot know the location of the sharp cutoff for a vague predicate is that the cutoff location could easily have been slightly different, given only minute differences in usage. The epistemicist version of the vagueness response says: perhaps candidate 217 isn’t conscious, so “is conscious” fails to

express a property that candidate 217 has, but we very easily could have expressed a slightly different property with the predicate “is conscious,” a property candidate 217 *does* have, and which you don’t have. (Of course, this other property isn’t consciousness, since candidate 217 isn’t conscious and couldn’t easily have been conscious.) But it is implausible that, without significant or even perceptible change in usage, we could easily have used “is conscious” to pick out a property other than being conscious, a property that you don’t have and wouldn’t have had under the relevant counterfactual circumstances. Relatedly, the epistemicist view implies that, if people’s linguistic usage dispositions had been imperceptibly different, we wouldn’t have had a word for consciousness. This is hard to believe. As Hawthorne, Dorr, and Yli-Vakkuri (2021: 319) observe, it is very strange to suppose that there are

close possibilities where certain conscious people use the word ‘conscious’ to express a property other than consciousness—a property they don’t have. Consciousness, one might suppose, is the sort of property such that if you have it, it is not easily going to escape your notice, and you are going to want to have a word for it.

Of course, in these nearby scenarios, your utterance of “I am conscious” would still have been true, since there would be a coordinated shift in the reference of “I.” Your use of “I” would refer, not to yourself, but to candidate 217. As Hawthorne, Dorr, and Yli-Vakkuri note, this is very strange.

You might have thought that the practice of using ‘I’ as a device of self-reference is extremely entrenched, so that a world would have be drastically different from actuality for someone such as yourself not to refer to themselves by an ordinary, literal use of ‘I’. (For these and other related reasons, Hawthorne, Dorr, and Yli-Vakkuri express sympathy for substance dualism, albeit as their “second-choice view” (326).)

The substance dualist can easily avoid this threat of too many subjects. He can say that the psychophysical laws pair each cloud of overlapping person-candidates with a single immaterial subject. None of the candidates is conscious. All of them have physical states that are causally relevant to the consciousness of the single subject (another case of benign overdetermination).³⁰

The dual-aspect theory can also avoid the threat. Perhaps the best option here is a “selectionist” view: the psychophysical laws somehow arbitrarily select one candidate from each cloud of overlapping candidates to be the bearer of irreducible phenomenal properties. One model involves indeterministic selection. Perhaps there is some physical condition C, common to each of the candidates, such that it’s a law that: if there are overlapping Cs, one is selected at random (with uniform probability?) to be the bearer of phenomenal properties. No doubt further refinements of this crude sketch are possible, and there may be other reasonable selectionist models.³¹ But it shows how dual-aspect theory can, in principle, avoid this threat of too many subjects.

As before, one can give a parody argument that, if physicalism about tablehood is true, then there are many tables where we ordinarily think there is one. Again, this argument must be unsound because dualism about tablehood is absurd, so I had better find a disanalogy. The key disanalogy, in my view, is that, in the table case, we can sensibly say that while each of the table candidates is very much *like* a table, only one candidate is a table, but it is vague which one it is. This is the analogue of denying Physical Egalitarianism (while adding the vagueness thesis). But this view doesn’t lead to the same problems as the denial of Physical Egalitarianism in the

³⁰ Unger (2006) and Zimmerman (2010) suggest a similar overdetermination account.

³¹ Simon (2017a: 464) considers a model on which the psychophysical laws “involve principles of attraction and repulsion, that secure that phenomenal subjects are both unified, and not too close to one another.”

Person-Candidates Argument. For example, this view about tables suggests that there are many things with properties that have roughly the same welfare significance as tablehood, but this doesn't yield normative revisionism because tablehood lacks welfare significance. Relatedly, we can say that the many "almost-tables" lack welfare without rejecting descriptive-normative proportionality.

5. Questioning the Ontological Assumptions of the Many-Subjects Arguments

In this concluding section, I want to consider a final strategy for resisting the many subjects argument: the *restricted-ontology* response. This response aims to sidestep the problem of too many subjects by denying the existence of the problematic entity or entities (brains, neuron-complements, bodies, person-candidates, etc.).³²

The best versions of the restricted-ontology strategy will start with a general ontological view that tells us what sorts of material things exist. It will then appeal to this view to motivate the rejection of this or that problematic entity. Here I'll discuss four ontological views and consider how each might support a restricted-ontology response. *The austere view* says that there are no composite objects, only simple objects. The only physical things that exist are very small things, perhaps elementary particles or spacetime points (or, on some austere physicalist views, one big thing: the cosmos).³³ *The organic view* holds that composition occurs sometimes, but only when the activity of some things constitutes a life. It accepts simples and organisms, but nothing else.³⁴ *The conservative view* accepts more-or-less all the entities accepted by scientifically informed common sense, but not much else. Thus, it accepts "ordinary objects" like brains and bodies, but not "extraordinary" objects like trout-turkeys (creatures composed of the

³² Cf. Olson (2007), Merricks (2001), and van Inwagen (1990).

³³ Sider (2013), Horgan and Potrč (2000).

³⁴ van Inwagen (1990).

front half of a trout and the back half of a turkey), finger complements, or things coincident with your body that would cease to exist were you to dance a jig in Times Square.³⁵ Finally, we have the abundant view, which accepts both ordinary and extraordinary objects. The abundant view embraces unrestricted composition (so we'll have trout-turkeys and finger-complements), as well as multiple things distinct but temporarily coincident with your body.³⁶ All versions of the restricted ontology approach are united in rejecting the abundant view, so we'll focus our attention on the other three views.

The austere view: The austere view, combined with physicalism, straightforwardly avoids commitment to too many subjects. On the austere view, there are no brains, finger-complements, bodies, or person-candidates, so there won't be *conscious* brains, finger-complements, bodies, or person-candidates. The problem with austere physicalism is that it faces the opposite problem of too *few* subjects. There is at least one subject in my vicinity, for I can know with something like Cartesian certainty that I am conscious, and therefore that I exist. If everything is simple, then I am simple. But I am obviously not any simple physical thing. (No elementary particle or spacetime point is me, nor am I the cosmos.) I must therefore be a simple immaterial thing. The austere view thus leads to substance dualism. Let us therefore set aside the austere view.

The conservative view: The conservative view can help with some of the arguments. First, a conservative can reasonably maintain that there are not many distinct human-body-shaped person-candidates. There is just one human animal present. For some peripheral atoms, it may be indeterminate whether those atoms are part of the animal. Thus, it may be indeterminate which set of atoms composes the human animal. But (determinately) there are not trillions of distinct roughly-human-body-shaped things in your vicinity. Since the conservative rejects

³⁵ Korman (2015).

³⁶ Lewis (1986), Hawthorne, Dorr, and Yli-Vakkuri (2021), Fairchild (2019).

unrestricted composition, there is no pressure to accept that, for each collection of atoms, there is something those atoms compose. Thus, the conservative can arguably sidestep the threat of many person-candidates.

The conservative view can also help somewhat with the threat of conscious parts. At least, it sidesteps the finger-complement argument and the neuron-complement argument, since the conservative will deny the existence of gerrymandered objects like finger-complements and neuron-complements. However, it won't directly help with the Brain Argument, since brains belong to commonsense ontology. The conservative physicalist will need another response to the Brain Argument. One possibility is to accept the "brain view," the view that you are your brain. But we've seen that the brain view seems to get your persistence conditions wrong (your brain will likely exist longer than you do). Moreover, conservatism combined with the brain view may also face a threat of conscious parts. The threat won't stem from gerrymandered objects like neuron-complements, but from ordinary parts that may have intrinsic physical features which, given physicalism, should be sufficient for consciousness. Possible candidates include your cerebrum, your cerebral hemispheres, or your cortex. Consider your left cerebral hemisphere. It's plausible that there could be an intrinsic physical duplicate of your left hemisphere in a vat, hooked up to wires and appropriately stimulated to duplicate the activity of your left hemisphere. And it's plausible that this hemisphere in a vat would be conscious. Again, the assumption that consciousness is intrinsic, and so shared by intrinsic duplicates, would support the conclusion that, if physicalism is true, your left hemisphere is conscious.³⁷ Thus, conservative physicalism combined with the brain view may not avoid the threat of conscious parts.

³⁷ Cf. Crummett's (2022: 323-4) argument from split-brain data that individual hemispheres of *non-split* brains might be conscious.

Nor does the conservative view help with the threat of conscious coinciders, since commonsense ontology accepts the existence of bodies. I conclude that conservative ontology is of little help in responding to too-many-subjects challenges to physicalism. It may sidestep the many person-candidates argument, but the conservative physicalist will need other responses to the other challenges.

The organic view: I concede that the organic view likely has the potential to avoid all the too-many-subjects problems. And unlike the nihilist, the organic view doesn't face the problem of too few subjects. The organic view denies the existence of brains, finger-complements, and bodies, so it easily avoids the threat of conscious parts and conscious coinciders. Matters are less clear-cut for the threat of conscious person-candidates. It's not obvious why there wouldn't be multiple distinct human organisms in your vicinity (one can argue that each of the many overlapping human-body shaped pluralities of particles is qualified to compose a living being). But there is at least the option of saying that there is just a single organism here with vague composition (van Inwagen 1990: ch. 17). To the extent that one is sympathetic to physicalism, the organic view's ability to avoid too many (and too few) subjects should be regarded as a significant point in its favor.

Unfortunately, there is not much else to be said in its favor. The three alternative ontological schemes described above seem far preferable. The austere and abundant views have the appeal of simple, elegant, non-arbitrary compositional principles. Conservatism doesn't share this virtue, but at least it appeals to Moorean sensibilities. It's a tricky methodological question how much weight Moorean considerations should be given relative to considerations of simplicity, non-arbitrariness, and so forth, and perhaps it's reasonable to give the former enough weight to make conservatism a viable position. The organic view scores poorly on both

measures. It offends against Moorean sensibilities and lacks the theoretical virtues of simplicity, elegance, and non-arbitrariness. It is a significant advantage of dualism that it can be combined with any of the three most attractive ontological schemes (the austere view, the conservative view, and the abundant view) without generating too many, or too few, subjects.

I conclude that restrictive (non-abundant) ontologies are of little help to the physicalist. The restrictive ontological schemes we've considered either (i) lead to a problem of too few subjects when combined with physicalism (the austere view), (ii) fail to avert the threat of too many subjects (conservatism), or (iii) are independently unattractive (the organic view). I obviously have not surveyed every logically possible restrictive ontological scheme that a physicalist might adopt, but my suspicion is that every such scheme will be impaled on at least one horn of this trilemma.³⁸

References

- Baker, Lynne Rudder (2000). *Persons and Bodies: A Constitution View*. New York: Cambridge University Press.
- Briggs, Ray & Nolan, Daniel (2015). Utility Monsters for the Fission Age. *Pacific Philosophical Quarterly* 96 (2):392-407.
- Builes, David & Hare, Caspar (2023). Why Aren't I Part of a Whale? *Analysis* 83 (2):227-234.
- Chalmers, David (2003). Consciousness and its place in nature. In Stephen P. Stich & Ted A. Warfield (eds.), *Blackwell Guide to the Philosophy of Mind*. Blackwell. pp. 102-142.
- Chappell, Richard Yetter (2015). Value Receptacles. *Noûs* 49 (2):322-332.
- Chisholm, Roderick (1989). *On Metaphysics*. University of Minnesota Press.

³⁸ Thanks to Geoffrey Lee for very helpful comments on a draft of this paper. Thanks to Adam Pautz, Trevor Teitel, Jeff Speaks, Harvey Lederman, Joel Hamkins, Paddy Blanchette, Kris McDaniel, Michael Rota, Tim Pawl, and the audience of a University of St. Thomas department colloquium for helpful discussion.

- Crummett, Dustin (2022). What if We Contain Multiple Morally Relevant Subjects? *Utilitas* 34 (3):317-334.
- Cutter, Brian (2021). Perceptual illusionism. *Analytic Philosophy* 62 (4):396-417.
- Cutter, Brian (2022). The Inconceivability Argument. *Ergo: An Open Access Journal of Philosophy* 9(12): 329-356.
- Cutter, Brian (forthcoming). The AI Ensoulment Hypothesis. *Faith and Philosophy*.
- Dretske, Fred (1995). *Naturalizing the Mind*. MIT Press.
- Fairchild, Maegan (2019). The Barest Flutter of the Smallest Leaf: Understanding Material Plenitude. *Philosophical Review* 128 (2):143-178.
- Geach, Peter (1980). *Reference and Generality*. Ithaca, NY: Cornell University Press.
- Goff, Philip (2017). *Consciousness and Fundamental Reality*. New York, USA: OUP.
- Hawthorne, John (2004). Why Humeans are out of their minds. *Noûs* 38 (2):351-58.
- Hawthorne, John (2022). Moral Vagueness and Epistemicism. *Journal of Ethics and Social Philosophy* 23 (2).
- Hawthorne, John, Cian Dorr, & Juhani Yli-Vakkuri (2021). *The Bounds of Possibility: Puzzles of Modal Variation*. Oxford: Oxford University Press.
- Horgan, Terry & Potrč, Matjaž (2000). Blobjectivism and Indirect Correspondence. *Facta Philosophica* 2 (2):249-270.
- Hudson, Hud (2001). *A materialist metaphysics of the human person*. Ithaca, N.Y.: Cornell University Press.
- Johnston, Mark (2016a). The Personite Problem: Should Practical Reason Be Tabled? *Noûs* 50 (4):617-644.

- Johnston, Mark (2016b). Personites, Maximality And Ontological Trash. *Philosophical Perspectives* 30 (1):198-228.
- Kagan, Shelly (1998). *Normative Ethics*. Routledge.
- Korman, Daniel (2015). *Objects: Nothing out of the Ordinary*. Oxford University Press.
- Kovacs, David Mark (2016). Self-made People. *Mind* 125 (500):1071-1099.
- Lee, Geoffrey (2013). Materialism and the Epistemic Significance of Consciousness. In Uriah Kriegel (ed.), *Current Controversies in Philosophy of Mind*. Routledge. pp. 222.
- Lee, Geoffrey (2017). Selfless experience. *Philosophical Perspectives* 31 (1):207-243.
- Lewis, David K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.
- Lewis, David K. (1986). *On the Plurality of Worlds*. Wiley-Blackwell.
- Lewis, David K. (1996). Elusive knowledge. *Australasian Journal of Philosophy* 74 (4):549 – 567.
- Lewis, David K. (1999). Many, but Almost One. In *Papers in Metaphysics and Epistemology*. Cambridge University Press.
- Merricks, Trenton (1998). On whether being conscious is intrinsic. *Mind* 107 (428):845-846.
- Merricks, Trenton (2001). *Objects and Persons*. New York: Oxford University Press.
- Mørch, Hedda Hassel (2019). Is Consciousness Intrinsic?: A Problem for the Integrated Information Theory. *Journal of Consciousness Studies* 26 (1-2):133-162(30).
- Olson, Eric T. (2004). Animalism and the corpse problem. *Australasian Journal of Philosophy* 82 (2):265-74.
- Olson, Eric (2007). *What Are We? A Study in Personal Ontology*. Oxford University Press.
- Olson, Eric T. (2010). Ethics and the generous ontology. *Theoretical Medicine and Bioethics* 31 (4):259-270.

- Parfit, Derek (1995). The unimportance of identity. In H. Harris (ed.), *Identity*. Oxford University Press. pp. 13-45.
- Parfit, Derek (2012). We Are Not Human Beings. *Philosophy* 87 (1):5-28.
- Pautz, Adam (2013). The real trouble for phenomenal externalists: New empirical evidence. In Richard Brown (ed.), *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. Springer. pp. 237-298.
- Pautz, Adam (2017). The significance argument for the irreducibility of consciousness. *Philosophical Perspectives* 31 (1):349-407.
- Schwitzgebel, Eric (2020). Is There Something it's Like to be a Garden Snail. *Philosophical Topics* 48 (1):39-63.
- Shoemaker, Sydney (1984). Personal Identity: A Materialist's Account. In Shoemaker and Swinburne (eds.), *Personal Identity*, Oxford: Blackwell.
- Sider, Theodore (2003). Maximality and microphysical supervenience. *Philosophy and Phenomenological Research* 66 (1):139-149.
- Sider, Theodore (2013). Against Parthood. *Oxford Studies in Metaphysics* 8:237–293.
- Simon, Jonathan A. (2017a). The hard problem of the many. *Philosophical Perspectives* 31 (1):449-468.
- Simon, Jonathan A. (2017b). Vagueness and zombies: why 'phenomenally conscious' has no borderline cases. *Philosophical Studies* 174 (8):2105-2123.
- Singer, Peter (1993). *Practical Ethics*, 2nd edition. Cambridge University Press.
- Strawson, Galen (2006). Realistic monism: why physicalism entails panpsychism. In A. Freeman (ed.), *Consciousness and its place in nature: does physicalism entail panpsychism?* pp. 3-31.

- Stoljar, Daniel (2001). Two Conceptions of the Physical. *Philosophy and Phenomenological Research* 62 (2):253-281.
- Taylor, A. P. (2013). The Frustrating Problem For Four-Dimensionalism. *Philosophical Studies* 165 (3):1097-1115.
- Tye, Michael (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. MIT Press.
- Tye, Michael (2021). *Vagueness and the Evolution of Consciousness: Through the Looking Glass*. Oxford University Press.
- Unger, Peter (1980). The Problem of the Many. *Midwest Studies in Philosophy* 5 (1):411-468.
- Unger, Peter (2006). *All the Power in the World*. Oxford: Oxford University press.
- van Inwagen, Peter (1990). *Material Beings*. Ithaca, N.Y.: Cornell University Press.
- Williams, Robert (2006). An argument for the many. *Proceedings of the Aristotelian Society* 106 (1):411-419.
- Williamson, Timothy (1994). *Vagueness*. Oxford: Oxford University Press.
- Yablo, Stephen (1992). Mental causation. *Philosophical Review* 101 (2):245-280.
- Zimmerman, Dean (2003). Material people. In Michael J. Loux & Dean W. Zimmerman (eds.), *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press. pp. 491-526.
- Zimmerman, Dean (2010). From Property Dualism to Substance Dualism. *Aristotelian Society Supplementary Volume* 84 (1):119-150.