
AN UNSTABLE ELIMINATIVISM

BY

JOHN W. CARROLL AND WILLIAM R. CARTER

Abstract: Do folk objects exist? A selective eliminativism says that some do and some don't. One eye-catching version of this says that tennis players exist, although tennis balls and tennis rackets do not. In our paper, we focus critical attention upon a defense of such a view presented by Trenton Merricks in his book, *Objects and Persons*. We will show that the argument advanced by Merricks shows either too much or too little.

In his book *Objects and Persons*, Trenton Merricks has reoriented and fine-tuned an argument from the philosophy of mind to support a selective eliminativism about macroscopic objects.¹ The argument turns on a rejection of systematic causal overdetermination and the conviction that microscopic things do the causal work that is attributed to a great many (though not all) macroscopic things. We will argue that Merricks' argument fails to establish his selective eliminativism.

I

On the folk account of the matter, it happens that a certain baseball strikes and shatters a certain window. The impact of the ball causes the window to break. That is opposed by the *No Overdetermination Argument* (NOA-1):

- (1) The baseball (if it exists) is causally irrelevant to whether its constituent atoms, acting in concert, cause the shattering of the window.
- (2) The shattering of the window is caused by those atoms, acting in concert.
- (3) The shattering of the window is not overdetermined; therefore,
- (4) The baseball (if it exists) does not cause the shattering of the window.

Pacific Philosophical Quarterly 86 (2005) 1–17

© 2005 University of Southern California and Blackwell Publishing Ltd. Published by Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

The inference from (1)–(3) to (4) is underwritten by:

Causal Principle:

Suppose O: is an object. The *x*s are objects. O is causally irrelevant to whether the *x*s, acting in concert, cause a certain effect E (i.e. O is not one of the *x*s, O is not a partial cause of E alongside the *x*s, none of the *x*s cause O to cause E, and O does not cause any of the *x*s to cause E). The *x*s, acting in concert, do cause E. And E is not overdetermined. It follows from all this that O does not cause E (Merricks, 2001, p. 58).

The argument can be extended in opposition to any attribution of any causal role to the ball in question. So, from this generalization of (4) and the principle we will call *Merricks' Dictum*² (“for *macrophysical objects*, to be is to have causal powers” (p. 81)), Merricks concludes that the baseball does not exist.

For the most part, our critical attention is focused upon the idea that NOA–1 allows for an ontology that includes some, though not all, folk objects. Merricks' eliminativism is selective in that he allows that persons exist (and that persons are macrophysical human organisms) while denying that baseballs and statues and most other macrophysical objects exist. Our misgivings about this selectivity arise from the fact that another version of Merricks' argument (NOA–2) speaks against human existence:

- (1) Edna (if she exists) is causally irrelevant to whether her constituent atoms, acting in concert, cause the shattering of the window.³
- (2) The shattering of the window is caused by those atoms, acting in concert.
- (3) The shattering of the window is not overdetermined; therefore,
- (4) Edna (if she exists) does not cause the shattering of the window.

By Merricks' lights, NOA–1 is sound, and the natural generalization of its conclusion together with Merricks' Dictum establishes that the baseball does not exist. But, also by his lights, NOA–2 is not sound. This prompts a crucial question: What is the difference between Edna and the baseball that accounts for the fact that NOA–2 is unsound, though NOA–1 is sound?

II

Merricks' answer to the crucial question is basically that Edna is, and the baseball is not, causally relevant. She is so in virtue of being conscious. The first step toward establishing her causal relevance lies with what Merricks calls a *Step One Argument* (SOA–1).

- (1) The intrinsic property of consciousness is such that (a) Edna's being conscious is not, of metaphysical necessity, implied by the

existence and intrinsic properties of, and spatiotemporal and causal interrelations among, Edna's constituent atoms; and (b) Edna caused the window to shatter in virtue of (existing and) being conscious.

- (2) If (1), then the fact that Edna caused the window to shatter in virtue of being conscious does not all by itself offer any reason for judging that Edna's constituent atoms cause such events; therefore,
- (3) The fact that Edna caused the window to shatter in virtue of having consciousness does not all by itself offer any reason for judging that Edna's constituent atoms caused the window to shatter in virtue of their spatiotemporal and causal interrelations (cf., Merricks, 2001, pp. 89 and 104).

Now, admittedly, (1a) of SOA-1 will look pretty implausible to anyone at all sympathetic to the physicalist slant of current work in the philosophy of mind. However, Merricks offers an anti-supervenience argument for (1a). The Finger Argument, as we'll call it, is meant to offer a *reductio* of:

Consciousness (C):

Necessarily, if some atoms $A_1 \dots A_n$ compose a conscious object, then any atoms intrinsically like $A_1 \dots A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1 \dots A_n$, compose a conscious object (Merricks, 2001, p. 94).

The Finger Argument relies on the claim that when a person loses a certain finger the atoms that remain, at the instant after this loss, may be both intrinsically and relationally just as they were shortly before the finger's loss. If (C) were true, these atoms would compose a conscious being prior to the finger's removal. But that is absurd, so (C) is false and (1a) of SOA-1 is true. Furthermore, Merricks assures us that SOAs will not serve to exempt baseballs from NOAs (p. 105):

The key here is that few properties can undergird both (1a) and (1b) (or an appropriate analogue of (1b)).

Merricks can think of no properties that baseballs would have (were such things to exist) that satisfy the conditions posed by (1a) and (1b) (p. 107).

III

There are numerous questions prompted by this use Merricks makes of SOA-1 to exempt us from eliminativism. We will, however, restrict our immediate attention to whether SOA-1 (together with the supporting Finger Argument) is an effective method for exempting *just* human

organisms – without exempting the macrophysical objects like baseballs and statues that Merricks rejects.⁴

We will begin to do so by considering a new SOA that involves a property other than consciousness. Consider SOA–2:

- (1) The intrinsic property of life (being alive) is such that (a) Edna's being alive is not, of metaphysical necessity, implied by the existence and intrinsic properties of, and spatiotemporal and causal interrelations among, Edna's constituent atoms; and (b) Edna causes event E to occur in virtue of being alive.
- (2) If (1), then the fact that Edna causes certain events to occur in virtue of having life does not all by itself offer any reason for judging that Edna's constituent atoms caused E; therefore:
- (3) The fact that Edna causes E in virtue of having life does not all by itself offer any reason for judging that Edna's constituent atoms cause E.

Once again, (1a) may look pretty implausible. What is noteworthy is that the Finger Argument also may be invoked in opposition to:

Life (L):

Necessarily, if some atoms $A_1 \dots A_n$ compose a living object, then any atoms intrinsically like $A_1 \dots A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1 \dots A_n$, compose a living object.

Suppose that one of Edna's fingers is removed. Presumably there aren't, just prior to the removal, two living individuals, Edna and Edna-minus say, occupying Edna's chair.⁵ (L) is accordingly false, and (1a) of SOA–2 is true. Thus it appears that the prospects for SOA–2 are no less promising than those of SOA–1. This conclusion does not appear to be a problem for Merricks until we notice that we have started down a road that exempts from his eliminativism all living things and, as we will see, other things too.

Being alive is not the only property that provides living things with SOA-exemptions. Consider:

Tree (T):

Necessarily, if some atoms $A_1 \dots A_n$ compose an oak tree, then any atoms intrinsically like $A_1 \dots A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1 \dots A_n$, compose an oak tree.

That is opposed by the Branch Argument, structurally similar to the Finger Argument, which focuses attention on a case wherein a single

branch is removed from a certain tree.⁶ If the Finger Argument is a *reductio* of (L), then the Branch Argument is no less a *reductio* of (T). Where ‘Ed’ refers to the putative tree outside a certain window, we might argue as follows (SOA–3):

- (1) The intrinsic property of being an oak tree is such that (a) Ed’s being an oak tree is not, of metaphysical necessity, implied by the existence and intrinsic properties of, and spatiotemporal and causal interrelations among, Ed’s constituent atoms; and (b) Ed causes things in virtue of being an oak tree.
- (2) If (1), then the fact that Ed causes various events in virtue of being a tree does not (all by itself) offer any reason for judging that Ed’s constituent atoms cause such events; therefore,
- (3) The fact that Ed causes various events in virtue of being a tree does not (“all by itself”) offer any reason for judging that Ed’s constituent atoms cause such events.

SOA–2 and SOA–3 apparently leave us with grave questions concerning the first premise of eliminativist argument NOA–3:

- (1) Ed (if it exists) is causally irrelevant to whether its constituent atoms, acting in concert, cause the cracking of the sidewalk.
- (2) The cracking of the sidewalk is caused by those atoms, working in concert.
- (3) The cracking of the sidewalk is not overdetermined; therefore,
- (4) Ed (if it exists) does not cause the cracking of the sidewalk.

If a rejection of (C) establishes a place in the world for human persons, as Merricks would have it, then a rejection of (L) and a rejection of (T) both similarly speak for the existence of trees. Merricks tells us that he doesn’t know “whether there are trees and ants” (pp. 114–115), but – as we see it – the case for exempting such things (from eliminativism) is no less credible than the case Merricks offers for existing human persons.

And indeed why stop with living things? If reflection upon the removal of fingers speaks against (C), similar reflection upon the removal of a single plank from a ship opposes:

Ship (S):

Necessarily, if some atoms $A_1 \dots A_n$ compose a ship, then any atoms intrinsically like $A_1 \dots A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1 \dots A_n$, compose a ship.

Misgivings about (S) naturally extend to NOA-challenges to the existence of ships as surely as misgivings about (C) extend to NOA-challenges to

human persons. The instability of Merricks' position, as we submit it is, consists of the fact that the sort of argument that is said to speak for personal exemption from eliminativism plausibly extends to ships and shoes and sitars – entities that certainly do not find a place in the World According to Merricks. We see no reason at all to believe that NOAs serve to banish ships and baseballs but not human persons. The case for exemption in the ship case is no less credible than is the case for our exemption.

Merricks argues that quidditative properties (being this ship, being this statue) do not provide ingredients for a Step One defense of the reality of ships and statues (p. 105). We do not contest this. Our claim is rather that *being a ship* (being a statue, being a baseball, etc.) serves the cause of the realist quite nicely – no less so, certainly, than the property of consciousness serves the cause of realism concerning persons. We agree wholeheartedly with Merricks when he says:

There is not a mighty host of conscious, reflective, pain- and pleasure-feeling objects now sitting in my chair, now wearing my shirt, now thinking about metaphysics (p. 95).

It is equally true that there is not a mighty host of oak trees nested within the tree that is (according to the folk) outside the window, not a mighty host of ships similarly associated with the ship (supposedly) in the harbor, and so on. Conceivably, it will be denied that an object can cause things in virtue of being a ship (or a tree or a statue or a . . .). But, if humans can cause things in virtue of consciousness, then we see no reason to deny that ships can cause things in virtue of their shipness. Example: The arrival of a ship caused joy on the part of the 200 marooned sailors and also brought about their survival. To be perfectly clear about this, let's suppose that the ship is the only potential vehicle of rescue anywhere nearby and that any vessel smaller than a ship (even if one were around) could not have rescued all of the seaman before the deadly hurricane struck. Then it seems that the shipness of that thing on the horizon is causally implicated in the explanation of the sailors' emotions and in the subsequent rescue. Why contest that? Yes, that it appeared to the sailors that there was a ship was a more proximate cause of their joy than there being the real ship, but that there was a ship was still a cause of the sailors' being subject to that appearance and so was also still a cause of their joy. Yes, a mock ship in place of the ship would also have caused the joy but that doesn't speak to whether the actual ship in fact caused their joy. In any case, considerations about there appearing to be a ship or what a mock ship would have done don't even begin to challenge our contention that shipness was causally relevant to the rescue; if there hadn't been a ship, then the sailors would not all have been rescued.

There is a potential objection to this that should be addressed, one suggested to us by what is at least a curiosity in the wording of Premise (1) of SOA–1. Notice that this premise begins, ‘The intrinsic property of consciousness is such that’. Perhaps, Merricks thinks that, in addition to satisfying conditions (1a) and (1b), a property needs to be intrinsic in order to exempt an object that exemplifies it from elimination. The objection emerges from the conviction that artifactual properties are relational and not intrinsic.⁷ If that is right we can’t run an SOA-like argument for shiphood, which leaves us without (contrary to what we’ve just claimed) any basis for judging that alleged ships are indeed relevantly independent of their microscopic parts. Some reflections on this.

It is argued that artifactual properties do not supervene upon their molecular or microscopic components. Consider the following:

Being a desk is a perfectly respectable supervenient characteristic, although its supervenience base is complex. Imagine a “molecular duplicate” of the desk existing in a world containing no intelligent life; the molecules that make up the desk’s twin have come together as they have purely by chance. Someone might deny – plausibly – that this duplicate is a desk. It does not follow that being a desk is not a supervenient characteristic, however, but only that, if being a desk is a supervenient characteristic, its supervenience base is “broad” (Heil, 1992, p. 137).

Precisely the same line can be taken for statues and ships and other artifacts. What are the implications of this for our conjecture that a SOA can be run both for consciousness and shiphood? If artifactual properties such as shiphood fail to be intrinsic such properties may not be suitable candidates for SOAs. But why believe that shiphood is not an intrinsic property? Suppose the reply is “Because shiphood does not ‘narrowly’ supervene on the microscopic components of alleged ships”. The general principle behind this reply says that failures of narrow supervenience for property ϕ speak against the intrinsicness of ϕ . But apply this principle to the property of consciousness. Merricks argues that narrow supervenience fails for consciousness; so if the general principle is correct we must conclude that consciousness is not intrinsic, whereupon a SOA cannot be run for consciousness. Merricks faces a dilemma here. Either failures of narrow ϕ supervenience do or do not speak against ϕ being an intrinsic property. In the former case, Merricks’ argument against the narrow supervenience of consciousness has the corollary that consciousness is not an intrinsic property, and so not a property that can establish the metaphysical independence of human persons. In the latter case, there is no reason to deny that shiphood and other artifactual properties are intrinsic. SOAs can then be run for such properties, which suggests that the case for independent ships and statues is no less credible than the case for independent human beings.⁸

IV

To this point, our approach has been to construct arguments analogous to SOA-1 that seem just as plausible. Moral: Merricks can't use SOA-1 to exempt us from eliminativism, unless he is prepared to exempt lots of other things. An important question remains: Do the SOAs *really* exempt all of us – and all the trees and all the ships – from the threat of the corresponding NOAs or is there a problem with the SOAs? If the latter is the case, then the NOAs would still be a formidable threat to our common-sense ontology, with us no less threatened than the trees. Our feeling is that we are not out of the woods, at least not yet.

The conclusion of SOA-1, line (3), says that Edna's causing the window to shatter in virtue of being conscious does not all by itself offer any reason for judging that Edna's constituent atoms caused the window to shatter. This clearly leaves open that there could be other reasons that together with Edna's causing the window to shatter do provide excellent reason for judging that Edna's constituent atoms also cause the shattering. Merricks agrees (p. 109), and insists that her atoms *do* cause the window to break. Relying on (3) of SOA-1, Merricks argues for (*):

(*)

Suppose a human person causes an effect E in virtue of (existing and) having a conscious mental property. Then there is no reason to think that that same effect is caused by *any* atoms in such a way that the person is causally irrelevant to those atoms causing that effect (pp. 109–110).⁹

We might inquire whether we are justified in believing that mental properties of human persons cause things to happen. Obviously there is no such causation, if there are no human persons. Are we justified in believing that human persons exist? Merricks advances (p. 88) the *Epistemological Argument* (EA-1):

... one's prima facie justification for believing that baseballs exist is undermined once one realizes that all of one's sensory experience regarding alleged baseballs would be the same whether caused by baseballs or merely by atoms arranged baseballwise (p. 88).

Making some simple substitutions, it is easy for us to formulate an obvious parallel (EA-2):

One's prima facie justification for believing that humans exist is undermined once one realizes that all of one's sensory experience regarding alleged humans would be the same whether caused by humans or merely by atoms arranged humanwise.

We wonder how EA-2 could be unsound if EA-1 is sound. Merricks thinks that the epistemological difference lies in the nature of our evidence for the existence of humans versus the existence of baseballs:

One's evidence for one's existence, and so for the existence of at least one human organism, is not straightforwardly sensory or even obviously causal (p. 88).

The 'and so' part of this is decidedly problematic, because the following argument is invalid:

- (1) We know that we exist.
- (2) We are human beings, so
- (3) We know that human beings exist.

This is a valid argument only if the conclusion is read as saying simply that a certain x is such that x is human and we know of x that it exists. Unfortunately, that does not establish that we know that there are humans (that humans exist). We believe that prospects for introspective knowledge of human existence are no more promising than are prospects for introspective knowledge of baseballs and statues. We are not convinced that Merricks has justification for postulating the existence of humans.

Perhaps, however, introspection speaks for the existence of a person. Merricks assumes that when a person deliberately brings about a certain event this person is causally relevant to the causal role played by her atoms in the event's production. The personal agent and her atoms are part of a single causal chain that results in the event in question, thus no overdetermination. According to Merricks,

The real oddity would be if some atoms caused an effect, an effect that was caused by a person's having a mental property, from *outside the causal chain* that included the person's having that mental property (p. 109).

Consider Edna's atoms as they are at the time she decides to throw the baseball. Certainly, these atoms do cause the window to shatter. The causal chain of microscopic events involving these atoms, the chain beginning with a batter striking the ball and ending with Edna's throw into the window, doesn't have a gap in it at the time of the decision. (We suspect that Merricks would agree with this much.) But, it is not very plausible to suppose that Edna's decision to throw the ball causes the very activity of the atoms that occurs *at the time the decision is made*. So, if there is a single causal chain, as Merricks insists, then, at the time of her decision, Edna and her atoms must be partial causes that work together to bring about the shattering. Merricks is apparently prepared to hold that the state of Edna's atoms at the time of her decision wouldn't have caused the window to break without the help of her simultaneous decision.

As odd as this sounds to us, we are inclined to wonder why a similar (partial cause) line can't be taken when we turn from human persons to trees and

statues? The likely reply is that trees and statues don't make decisions; it is our conscious decision-making that ensures that we are partial causes and that gives us immunity to eliminativism. The problem with this reply is that Merricks never reveals any relevant difference between human decision-making and the macro-level properties of all the other macrophysical folk objects.

Merricks reports that:

... I am happy to agree that (*) has the virtue of being falsifiable. For (*) could – at least in principle – be undermined by empirical evidence for the claim that every physical effect has a microphysical cause to which non-microphysical entities are causally irrelevant. Such vulnerability to empirical refutation is something I share with my opponents here ...

Understatement for effect: Microphysicists have not yet causally explained every physical event. How one bets on whether they will ultimately do so should turn on, among other things, the arguments. Because of the argument of this chapter up to this point, I conclude that (*) and its implications for the completability of microphysics are true (p. 111).¹⁰

Though we find what Merricks says here remarkable, we are pleased to find that he appreciates fully that the issue here is an empirical one. We find what he says remarkable because of how absolutely convinced the majority of philosophers of mind are that the empirical evidence will turn out just the other way.

We will resist out and out siding with these physicalist philosophers of mind. We do think, however, that the prevailing physicalist orientation is ultimately grounded in cognizance of a certain plausible proposition about the success of physics, one that should not be dismissed lightly: To believe that Edna and her atoms, at the time of her decision, work alongside each other as partial causes of the window's breaking is tantamount to believing that physics is misguided in its attempts to formulate the laws of nature that describe the temporal progression of microphysical states. At present, to formulate and test hypotheses about what these laws are, physicists conduct experiments in familiar laboratory settings. We see no reason to believe that, if they were able to study the behavior of atoms in brains instead, then new and radically different hypotheses would emerge and be confirmed. The physicists don't seem worried that the experiments they conduct on mindless things won't reveal all the basic workings of the microphysical – the physicists don't seem worried that there aren't any conscious beings inside their supercolliders!

We admit that science is nowhere close to running a crucial test of the empirical issue that divides Merricks and the physicalist philosophers of mind. That test would be to apply the presently best confirmed hypotheses of what the laws of microphysics are – hypotheses confirmed with evidence from what goes on in standard experiments, and not from what goes on in brains – to the microphysical state of a conscious human brain in order to make predictions about how its microphysical states would

evolve. If those predictions were right, the physicalist philosophers of mind would have been right to think that the mental and the physical don't work alongside each other as partial causes. If those predictions were wrong, then the physicalists would have been wrong and Merricks would have been right. Unfortunately, the complexity of the human brain is so staggering (and the potential consequences of tampering with one so grave) that scientists are nowhere close to being able to predict the properties of brain particles based on prior microphysical states and the best confirmed statements of the temporal progression laws.

This is what leaves open the epistemic possibility that when microphysical particles are arranged in certain ways there might emerge non-redundant causal powers. The very same epistemic possibility, however, is left open about other complex systems *whether they are conscious or not*. It is just as epistemically possible that quasars, fine wines, and maybe even baseballs and statues are partial causes working alongside their atoms to bring about certain physical effects. We find it enormously unlikely that the specific complexity associated with conscious human beings is the unique source for an object partially causing an effect its atoms also partially cause. Merricks is gambling that this is true. We are very eager to put our money down on that not being the case.¹¹



Our charge of instability concerning Merricks' position does not rest on this. Instability is in plain view when we consider *Micro Exclusion Arguments* such as (MEA-1):

Assume for *reductio* that some physical event has mental cause M. But every physical event (that has a cause) has *microphysical causes* (or a microphysical cause). So every physical event that has M as a mental cause has microphysical causes. M is causally irrelevant to whether those microphysical causes cause that physical event. As a result, every physical event caused by M is overdetermined. But there is no such overdetermination. Therefore, what we assumed for *reductio* is false. No physical events have mental causes (p. 141).

Merricks rejects MEA-1. Skepticism concerning MEA-1 naturally carries over to MEA-2:

Assume for *reductio* that some physical event has physical cause P, where P is (or involves) a physical property of a certain human. But every physical event (that has a cause) has microphysical causes (or a microphysical cause). So every physical event that has P as a physical cause has microphysical causes. P is causally irrelevant to whether those microphysical causes cause that physical event. As a result, every physical event caused by P is overdetermined. But there is no such overdetermination. Therefore, what we assumed for *reductio* is false. No physical events have physical causes consisting of (or involving) physical properties of humans.

This reasoning is, by our lights, no more (or less) plausible than MEA–1. We do not know if Merricks would agree. But suppose he does agree. Since he rejects MEA–1, Merricks then will reject MEA–2. But now consider MEA–3, an argument similar to MEA–2 concerning an alleged physical cause P that is some physical property of a certain statue. Presumably MEA–3 should be rejected in the event that MEA–2 is rejected. Objections to MEA–2 will inevitably carry over to MEA–3. The argument for denying that any physical event can have a physical cause that is a physical property of a statue carries no more weight than the corresponding argument for denying that any physical event can have a physical cause that is a physical property of a human. We must reject the former argument in the event that we reject the latter. It emerges from this that Merricks has reason not to reject MEA–2. Merricks can hardly reject MEA–3 given that that rejection would strike at the very heart of his NOA assault on statue realism.

If all of this is correct, then it emerges that Merricks should both defend MEA–2 and also reject (as he does) MEA–1. He should be prepared to argue that only the mental properties of human beings are implicated in the causation of physical events. We can think of no philosophical thesis more deserving of an incredulous stare than this.

VI

To this point, we have considered a serious issue that any selective eliminativism must face no matter how it tries to exempt the selected folk entities from eliminativism. If the eliminative argument works against Gs, one will wonder whether it works against Fs. If some consideration is offered to exempt the Fs, one should immediately wonder whether this consideration also exempts the Gs. Just so, we have highlighted the ineffectiveness of the SOA defense to establish a selective place in the true ontology for us humans. We will now consider a somewhat simpler and potentially more serious problem for this defense that has to do with whether it really exempts the Fs, all of us humans, in the first place.

To begin, there are many human organisms recognized by our common-sense ontology who have not been exempted. We have in mind severely brain-damaged human fetuses and newborns that may never have had even the capacity for consciousness. About these human organisms, the SOA defense does not get off the ground.

Taking this a bit further, it seems clear that each of us conscious humans has at some time been unconscious or at least not conscious. If Edna's exemption from NOA–2 depends upon the fact that she is conscious, then each of us has at times been in a position where an analogue of that argument speaks against *our* existence. What would Merricks say?

We are not much inclined to assume that Merricks thinks that we come into existence when we are first conscious, go out of existence when we lose consciousness, and then come back into existence when we regain consciousness. We are even less inclined to assume that he thinks that *a new entity* comes into existence each time that, according to the folk way of thinking, *we* become conscious. Instead, we suspect that Merricks believes that having a capacity for consciousness at a time is enough to avoid the NOA threat.

We find this a puzzling stance for Merricks to take. First of all, it leads to problematic judgments about a simple case. Consider Earl, who has lived a normal life but has just been in a horrendous accident. Earl is now in a permanent coma without even the capacity for consciousness and so, according to the stance under consideration, Earl does not any longer exist. We find this consequence unintuitive in and of itself. But there is also a contradiction lurking: According to Merricks, Earl was the human organism driving the car prior to the accident. *Prima facie*, that human organism does still exist. So, it looks as if Merricks is also committed to the claim that Earl still exists! (Our guess is that Merricks will bite the necessary bullets and accept (i) that whatever it is that is lying in the hospital bed is not Earl and (ii) that whatever that is lying in that bed it is not the human organism that was driving the car. This avoids the contradiction but has little else to recommend it.) Second of all, merely having a capacity for consciousness doesn't thwart the NOA threat to unconscious persons. Since only consciousness has officially been recognized as a property of macro-objects that makes for nonredundant causation, we don't see what it is about Edna at a time when she is unconscious that differentiates her from the baseball considered at that same time. When it comes to metaphysical independence with respect to constituent simples, surely she is, at that time, in precisely the same sort of situation that the baseball is. Think of it this way: We can run a NOA argument about Edna lying unconscious in the hall just as well as we could about the baseball. Edna (if she exists) is causally irrelevant to her atoms causing the hallway to be blocked. Yet the blocked hallway is caused by her atoms acting in concert and this is not a case of overdetermination. Therefore, Edna (even if she does exist) does not cause the hallway to be blocked. Generalize as Merricks does about nonliving things (see pp. 80–81): If unconscious things exist, they cause nothing. So, unconscious things lack causal powers. Therefore, by Merrick's Dictum, there are no unconscious people.

VII

If SOAs do not undermine the threat of the corresponding NOAs, what does? This is not an easy question. Nor is it one we will attempt to resolve

fully here. But we would like to conclude by calling attention to one problematic feature of NOA–1. In doing so, we are going to have to be a little more careful about the ontological commitments of the key causal claims. We will also have to be more careful about just what overdetermination *is* and why it is *objectionable*.

It appears that a ball strikes and breaks the window, sending the atoms that composed the window this way and that way. Suppose that is so. Notice that none of the atoms that compose the ball is such that *it* causes the window to break; presumably, the existence of any one atom makes no difference. We agree with Merricks that it is *the atoms* (plural) that cause the window to break. We suspect that ‘the atoms’ is employed by Merricks as a plural referring expression (see van Inwagen, 1990, p. 23) and is not meant to refer to a single physical object; e.g., the collection of the atoms. There better not be a premise of NOA–1 that commits Merricks to the existence of a composite object, a collection of atoms, in addition to the individual atoms. So it is true that (i) the baseball caused the window to break, that (ii) its atoms caused the window to break, and that (iii) no individual atom broke the window.

We now want to raise the question of whether (i) and (ii) signal the presence of any objectionable overdetermination given that (iii) also holds. As we see it, (i) and (ii) clearly do not commit us to the claim that, whenever a macroscopic object causes some effect, there are in fact two objects that together overdetermine that effect. Strictly speaking, on a perfectly natural common-sense ontology, there are each of the individual atoms and there is the baseball, and the baseball causes something that none of the atoms do. The baseball causes the window to break; none of the atoms cause the window to break. There is the *appearance* of overdetermination or, it might be better to say, a *kind* of overdetermination stemming from the truth of ‘The baseball caused the window to break’ and the truth of ‘The baseball’s atoms caused the window to break’. But there is no messy ontology that underlies these causal truths, thus, no “ugly picture” (p. 67). There is also no other object that does the work of the baseball, thus no “redundancy” (p. 67, again).

Where exactly does NOA–1 go wrong? The answer depends on exactly what it is best to say about the “overdetermination” stemming from the fact that baseball caused the window to break and the fact that its atoms caused the window to break. Do these facts just create the appearance of overdetermination where really there is none? If you think so, then you should also think that NOA–1 is invalid. The causal principle that underwrites the inference from (1)–(3) to (4) is false: The baseball is an object, its atoms are objects, the baseball is causally irrelevant to whether the atoms cause the window to break, the atoms do cause the window to break, and the window’s breaking is not overdetermined, but it is not true that the baseball doesn’t cause the window to break. Instead, and we

believe this is marginally a more natural way of describing the situation, you might think that the fact that baseball caused the window to break and the fact that its atoms caused the window to break do constitute a kind of overdetermination. If you do, then you should also think that Premise (3) of NOA–1 is false; the shattering is overdetermined but it is not objectionable overdetermination – more cautiously, Merricks hasn't provided any reason to think it is objectionable.¹²

But what about the state of affairs of *the atoms' acting in concert*? Surely, that caused that window to break as did the baseball, and isn't this enough to establish the threat of a genuinely ugly overdetermination since arguably that state of affairs is not identical to any simultaneous state of the baseball? We agree that the atoms' working together caused the window to break.¹³ And, the baseball does too. But, when we consider carefully what ontology this commits us to, we do not discover any obviously problematic overdetermination. As we see it, to say that the atoms' acting in concert caused the window to break is just to say that *the window broke because the atoms acted in concert*. Apparently, all the ontology that commits us to is the window and the atoms. As we've argued, no atoms caused the window to break. The baseball does, but – since none of its atoms do – there is no overdetermination that Merricks has given us any reason to reject.¹⁴

Department of Philosophy and Religion
North Carolina State University

NOTES

¹ The argument is the Exclusion Argument. As far as we know, within the philosophy of mind literature, this argument has not explicitly been used in support of any form of eliminativism. As it has been presented, it is naturally construed as an argument for an identity theory about mental properties, though it is often just taken as a challenge to establishing their causal relevance. An early statement of the argument can be found in Norman Malcolm (1958). Jaegwon Kim (1993 and 1998) is responsible for prompting a lot of the more recent discussion. Our understanding of the argument owes much to Stephen Yablo (1992) and David Robb (1997).

² So dubbed for its similarity to Alexander's Dictum (see Merricks, 2001, p. 65 and Kim, 1993, p. 348). There is at least a small problem with Merricks' application of his dictum, which we will ignore in the text. The natural and seemingly warranted generalization of (4) is that the baseball does not *in fact* cause anything, never has and never will. This does not entail the claim that the baseball does not have any causal *powers*, because it is possible that it have causal powers that are never actually manifested.

³ We assume that Edna, if she exists, has thrown the baseball into the window.

⁴ Not wanting to tackle too much, we will restrict our attention throughout our paper to the specific manner in which Merricks tries to draw the line between humans and the other folk objects. So, for example, though we will have a lot to say about the Finger Argument and Merricks' claims about what sorts of properties are causally relevant, we will not

consider other reasons one might conceivably have for thinking that only a few properties could undergird both (1a) and (1b) of SOA-1.

⁵ Apparently, Merricks would not find anything counterintuitive about two living things, Edna and Edna-minus sitting in the chair. Compare what he says about the two objects on the pedestal both having at least mass *M* (p. 106). We do find this counterintuitive, at least as counterintuitive as there being two conscious beings prior to the removal of the finger. In any case, we think that in our next two deployments of analogues of the finger argument, even Merricks would have to agree that there are not two trees just prior to the removal of the branch or two ships prior to the removal of the plank.

⁶ Merricks (1998, p. 66) apparently accepts the Branch Argument.

⁷ See Baker (2000) pp. 30–31; especially see note 10.

⁸ On p. 104, Merricks tells us that at most one of the following theses are true: “first, many conscious mental properties are intrinsic; second, whether atoms compose a conscious object supervenes on microphysical doings”. Merricks clearly accepts the former claim and rejects the latter. Failures of supervenience do not, for Merricks, establish that we are dealing with non-intrinsic properties! Thus the fact that shiphood fails to (narrowly) supervene on microscopic entities does not establish that shiphood fails to be intrinsic. A SOA can indeed then be run for shiphood (and other artifactual properties).

⁹ This is edited slightly. Merricks includes discussion of a possible exception and then addresses the possible exception. The possible exception does not play a role in our criticism. It is left out to simplify the discussion.

¹⁰ We have substituted ‘(*)’ where Merricks has ‘(5)’.

¹¹ Michael Rea (2001, pp. 138–139) argues forcefully that whatever empirical evidence there is that convinces Merricks that baseballs are causally redundant should also convince him that we are causally redundant. Rea would apparently just as eagerly put his money down where we would.

¹² As Merricks recognizes, there is a good deal to be said about the line between objectionable and non-objectionable versions of overdetermination. We agree with the following: “As overdetermination is ordinarily conceived, . . . overdetermining causes are thought of as both independent and sufficient for their effects” (Marcus, 2001, p. 75). Given this conception of overdetermination, some overdetermination claims look entirely unproblematic.

¹³ Well, we are tempted to be pedantic and point out that the atoms – since they are not agents – don’t literally *act in concert* or *work together*. But, since Merricks could raise the same point about *the atoms’ having such and such momentum*, we’ve resisted this temptation in the text.

¹⁴ Thanks to Ron Endicott, Mark Heller, and some anonymous referees for useful comments on earlier versions of this paper.

REFERENCES

- Baker, Lynne Rudder. (2000). *Persons and Bodies*. Cambridge: Cambridge University Press.
- Heil, John. (1992). *The Nature of True Minds*. Cambridge: Cambridge University Press.
- Kim, Jaegwon. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Kim, Jaegwon. (1998). *Mind in a Physical World*. Cambridge: MIT Press.
- Malcolm, Norman. (1958). “The Conceivability of Mechanism”, *Philosophical Review* 77, pp. 45–72.
- Marcus, Eric. (2001). “Mental Causation: Unnaturalized but not Unnatural”, *Philosophy and Phenomenological Research* 63, pp. 57–83.
- Merricks, Trenton. (2001). *Objects and Persons*. Oxford: Clarendon Press.

- Merricks, Trenton. (1998). "Against the Doctrine of Microphysical Supervenience", *Mind* 107, pp. 59–71.
- Rea, Michael C. (2001). "How to Be an Eleatic Monist", *Philosophical Perspectives* 15, pp. 129–151.
- Robb, David. (1997). "The Properties of Mental Causation", *Philosophical Quarterly* 47, pp. 178–194.
- Van Inwagen, Peter. (1990). *Material Beings*. Ithaca, NY, Cornell University Press.
- Yablo, Stephen. (1992). "Mental Causation", *Philosophical Review* 101, pp. 245–280.