Merrick’s Identification of the Person and Organism

Introduction

Trenton Merricks argues for the eliminativism of every kind of composite object except for one on the basis of some familiar and some original arguments (2001).The older arguments arise from the well known objections to spatially coincident entities. If there are lumps co-located with statues, there is the problem of accounting for their different modal properties given their identical microphysical structures. Moreover, there would be a needless multiplication of objects as far as causal explanations go. Whatever the statue causes, the lump arranged statuewise would seem to as well. Merricks also rehearses some familiar arguments against person/organism spatial coincidence: there would then be two thinkers where we prefer just one and consequently epistemic problems would arise as each thinker had no reason to believe she was the organism rather than the person.

Merricks also provides a fascinating, original argument against there even existing just one composite object that is not spatially coincident with another. This argument is based on the overdetermination that would arise if the composite object has causal powers for it would seem not to cause anything that wasn't already being caused by the microphysical objects composing it. The "atoms arranged statuewise" seem to be causally able to produce anything that the statue does. Merricks doubts that there is pervasive overdetermination involving *every* composite object and its microphysical components. If there is only microphysical causation, that would mean any existing composite object would be epiphenomenal. Since Merricks insists that every existing physical object should have causal powers, he concludes that we ought to eliminate all but one kind of composite object on the grounds that they don't do any causal work. His exception is thinking beings. They possess nonredundant causal powers. The person's conscious powers don't supervene on the properties and relations of their microphysical parts and so their effects

aren't overdetermined. He believes persons exist and they are organisms though not necessarily organisms. He is open to the possibility that a person could exist without being biologically alive if its cognitive capacities were preserved when all of its organic parts were gradually replaced with inorganic parts.

While Merricks believes that we persons are organisms, his arguments eliminate mindless organisms like protozoa and plants, as well as deny that any of us were ever once mindless early embryos. However, he mentions that he is open to the possibility that unthinking organisms may avoid the eliminativism of the overdetermination argument. I will argue that they do, in fact, avoid elimination by Merricks's overdetermination argument but that their existence is not something that he can be so nonchalant about; He can't calmly add mindless organisms to his ontology without making amendments elsewhere in his theory. Readers shall see that because of the existence of mindless organisms Merricks cannot maintain that identity is what matters (2001: 125), organisms can be transplanted if their brains are (2001: 52) and that it is possible for living persons to undergo part replacement and become inorganic (2001: 86-87), while also insisting that there is not a human person co-located with a distinct human organism. The root of the problem is that persons can survive if just the parts of the brain needed for realizing thought do and these parts are not needed for organisms to survive. The result is that organisms and persons can be separated and found in different places and can each survive events the other can't. An obvious consequence of this is the non-identity of persons and organisms.

If my above claims survive scrutiny, I see Merricks as then having a pair of options, neither of which will be welcome to him. His first option is to accept co-location and any causal overdetermination it brings. And if he accepts co-location in the case of persons and organisms, then it might become harder to resist it elsewhere in the cases of statues and lumps, flags and

cloths and so on. However, I will suggest that Merricks still has some grounds to resist widespread co-location and pervasive causal overdetermination and thus can sustain much of his eliminativist project. I will add that the co-location of persons and organisms may not be as bad as he thinks for he exaggerates the epistemic problems of self-reference and thus the counterintuitiveness of the two co-located thinkers. And I will also point out that there are bizarre cases of conjoined twins that are best described as two organisms possessing the same cerebrum and thus sharing the same thoughts. Although these thinking entities overlap rather than are spatially coincident, they provide analogues of the problems of co-located thinkers that cannot be avoided by those, like Merricks, who believe that organisms are identical to persons.

Merricks's second option is to abandon his adherence to the three related positions that organisms and persons can be transplanted, that persons can possibly survive inorganic part replacement, and that it is identity that matters to us in our survival. Abandoning these three positions would involve embracing a position more similar to van Inwagen (1990) and Olson's (1997) full fledged animalist views. If persons not only are identified with organisms but are understood to be essentially alive, then they wouldn't be able to survive inorganic part replacement or be transplanted if their cerebrum is. However, Merricks will then have to explain away our intuitions about being transplanted when our cerebrum has been and our surviving inorganic part replacement. He most likely will argue that we are misled by the belief that identity matters. To do so, he must defend the claim that what matters is not that we survive but our psychology does. But the costs of this move are considerable because it's difficult to believe that the continuation of our psychology in someone else offsets our becoming mindless or ceasing to exist.

Causal Overdetermination and Eliminativism

Let's begin with Merricks's argument against the existence of most composite entities and his explanation of why the sentient are not likewise eliminated. What he finds problematic about macroscopic objects such as baseballs is that they would appear to be causally inefficacious. They fall prey to what he calls the "Overdetermination Argument" (2001: 56). The first premise of this argument is that if a baseball existed, it would be causally irrelevant to its atoms acting in concert as they break the window. ("Atoms" is just his placeholder for the building blocks of composite material things. It may stand for electrons and quarks or some yet unknown subatomic particle.) The shattering of the window is an event composed of many micro events of the atoms causing the scattering of the parts of the window. The second premise is that the atoms acting in concert break the window. The third premise is that the shattering of the window is not overdetermined. Merricks concludes that if the baseball exists, then it does not cause the shattering of the window. The moral he draws is that there are no such things as baseballs.

Merricks's reasoning is that if there were to be things like baseballs, they would have to be causally efficacious, and to allot them causal powers would be to accept overdetermination. And such causal redundancy wouldn't be limited to baseballs; there would be countless analogues amongst the objects of folk ontology: bats would cause the same effects that their atoms arranged batwise did, rocks would cause the same events that their atoms arranged rockwise do, and so on. Since Merricks doesn't think there is such pervasive overdetermination, and given that there is independent evidence for microphysical atoms causing things to happen when they are not arranged baseballwise, that leaves the composite object epiphenomenal. Skeptical that physical objects can be epiphenomenal, Merricks endorses eliminativism. He

maintains that it is better to eliminate such objects than to accept either such pervasive overdeterminism or epiphenomenal objects.

We persons escape the eliminativist moral of the Overdetermination Argument because our minds, though casually efficacious, do not overdetermine our actions. Merricks is not a substance dualist. He believes we are composed only of atoms (2001: 85). Nevertheless, we are neither identical to the sum of those atoms nor do our mental attributes supervene upon their intrinsic properties and causal and spatial relationships to each other. Merricks rejects the following account of microphysical supervenience: "Necessarily, if some atoms $A_1...A_n$ compose a conscious object, then any atoms intrinsically like $A_1...A_n$, interrelated by all the same spatiotemporal and causal interrelations as $A_1...A_n$, compose a conscious object. (Merricks: 2001, 94). Since consciousness does not supervene upon such microscopic atoms and their properties and relations, then we shouldn't expect whatever a mind causes to be also caused by certain atoms acting in concert.

The basis for Merricks's claim that consciousness does not supervene upon microphysical states is that the atoms composing a person at one moment may have been in the exact same states vis a vis each other a moment earlier without composing a person. Composing a conscious person at T_2 is an aggregate of atoms, call it "A-," that existed in the exact same physical state previously at T_1 embedded in an aggregate one atom larger. Call the larger aggregate "A." But the atoms of A-, even with their intrinsic properties and spatio-temporal and causal interrelations remaining unchanged from T_1 to T_2, didn't compose a person at T_1. If they had, then there would have been two thinking persons where commonsense tells us there was just one. In fact, there would be far more than two people under the reader's clothes since the same kind of step from A to A- can be repeated many times. If being a person with a consciousness

supervened upon the atoms of A- when they were so arranged with the powers that they possessed, then anytime they were so arranged it would be *necessary* that they composed a person. So if consciousness is an intrinsic property - and Merricks emphasizes that he is not talking about contentful states - then consciousness does not supervene and microphysical supervenience is false. Moreover, if conscious states are causally efficacious, then it would appear that conscious states cause physical events that no microphysical states do. This doesn't mean that physical states don't have a role in causal processes, they are obviously involved in neural firings and muscular contraction. But at the start of the chain is a conscious cause that can't be accounted for by an arrangement of microphysical objects. That is, there would be times that the atoms of A- don't cause a certain subsequent physical state that a conscious being would initiate when composed of atoms intrinsically like those of A- and interrelated by all the same spatial and causal interrelationships. Merricks interprets these possibilities as evidence for downward, nonredundant mental causation.[1]

[1] A reply to Merricks, one that I put forth quite tentatively, is to claim that what makes something a part of an organism rather than a foreign body is that it is *necessarily* caught up in certain causal life processes such as metabolizing food, maintaining homeostasis, excreting waste, assimilating oxygen, repairing damage, fighting infection and so on. If an object is not involved in such processes, it is not part of the organism. So it may be necessary that if a microscopic part of the organism is removed, doing so will immediately have an effect on some other atoms in the organism, changing how they causally interact to each other. For instance, remove an electron from a membrane and you instantly change its charge and permeability. So it may be that any atom that can be destroyed or removed without affecting some of those remaining was never part of the same life and thus organism as the latter. Therefore the existence of a conscious organism would be necessarily dependent upon the arrangement and powers of its microphysical parts. If this means overdetermination, then so be it, for it makes no sense to eliminate us. However, I will proceed in the rest of the article as if Merricks's thesis escapes this criticism.

A Biological Form of Downward and Non-Redundant Causation

What I'm interested in is showing that the same argument that Merricks's uses to protect us from eliminativism provides grounds for resisting the elimination of all living beings, even those that are devoid of cognitive capacities. Organisms have various metabolic and homeostatic properties. For example, an organism's internal monitors may recognize the need to break down more of the sugar in the bloodstream. The organism that has such a property of sugar regulation is composed by atoms that do not *necessarily* entail the existence of such an entity whenever they are so arranged. They could have been encased moments earlier in a slightly larger organism. They wouldn't have then composed a metabolizing organism within a larger metabolizing organism. So the property of being a sugar-regulating organism doesn't supervene on certain microphysical atoms and the causal and spatial relations between them. Since maintaining sugar levels is a causally efficacious property of unthinking organisms, it would appear that they have nonredundant properties - at least if the analogous argument that organisms have the causally nonredundant property of being conscious is sound. The organism's property of sugar regulation doesn't overdetermine chemical changes, causing the same effects that the atoms arranged organism-wise do. Those very atoms when embedded within an aggregate just a few atoms larger did not maintain and monitor sugar levels in every part of the larger organism except in the region where those few extra atoms were found. If atoms so arranged did of necessity compose one organism, and were part of a larger organism, there would then be two overlapping sugar-regulating organisms only a few atoms different in size, where folk ontology and biological theory tells us there should be just one. What has been said about the abilities of organisms to metabolize sugar can be extended to other life processes including maintaining

homeostasis, assimilating oxygen, excreting waste etc. They don't supervene on their microphysical components, if they did there would be organisms embedded within organisms

Merricks would seem not to mind this conclusion. He acknowledges in another place (2001: 114-15) that he does not know whether or not there are any mindless organisms but sees no threat to his general eliminativist project in admitting their existence. His thinking may be that he can add a new kind of object and still reject any co-location and deny existence to artifacts and other things thought to populate the nonliving natural world like mountains and rivers. However, my contention is that matters are actually not as favorable for Merricks's project as they may first look. The same argument that will preserve the thinking person now prevents the elimination of mindless organisms. So there was a mindless organism that existed before thought emerged in the course of human development. It is hard to imagine that the organism goes out of existence when the property of personhood is instantiated. No problem, you might say, Merricks can just identify the two. The person existed earlier without having a mind and later acquired one. However, we shall see in the next section that the problem with that move is that the same intuitions that suggest the transplant of the organism will support the transplant of a person that is not an organism thus ruling out the identification of persons and organisms.

Cerebrum Transplants: Leaving the Organism Behind

Merricks believes that if he couldn't account for our intuition that we move from one body to the another when our brain is transplanted then it would be better to recognize the brain's existence and to identify each of us persons with our brain. He writes: "I agree that the view that we are brains is more plausible ceteris paribus, than any view that denies either that one 'goes with the brain' or that each human person is a thinking being" (2001: 136). Merricks doesn't want to grant that the organism and the brain both exist for then there would be too many

thinkers. He would rather eliminate the brain and preserve the organism. But he can endorse the converse, although that would involve giving up some commonsense beliefs that one can see and hold persons (2001: 136). However, Merricks is not worried that he has to adopt the view that the only composite human entities that exist are human brains and that we are identical to them because he believes he can capture the intuitions that motivate the view that we are thinkers and can be transplanted. He explains:

> Suppose my 'brain' is put into a new 'body.' I believe I go with my 'brain.' Yet I also, believe that I am organism, not a brain. I think the best way to render these beliefs consistent is to interpret a case of 'brain transplant' in the following way (van Inwagen 1990: 169-81). I am whittled down to brain size. This is the logical extreme of amputation and the resultant shrinking of the amputee. When put in a new body, I grow rapidly as new parts are added to me. This is the logical extreme of receiving transplants and the resultant growth of the recipient (2001: 52).

Van Inwagen (1990: 172-81), and apparently Merricks, think a whole brain transplant is the transplant of a very small organism because the controls of vital biological processes are there in the brainstem even if they are not attached to any organs that they can control.[2]

The problem for Merricks is that one can make a similar transplant argument for just a a part of the brain, the cerebrum, rather than the whole brain. And if organisms can exist without their cerebrum, then the organism is left behind when the person is transplanted, so identifying the person with the cerebrum will not be identifying the organism with the person. The organism cannot become pared down to the size of the transplanted part of the brain for the detached

[2] A similar view of the pared down (brain size) organism can be found in Olson 1997: 44-46, 132-34.

cerebrum is an organ and not an organism 1997: 115). Given that we saw in the earlier quote that Merricks believes it is prima facie more plausible that we are the transplanted brain then to deny that one goes with the transplanted brain of a thinking being, analogous reasoning suggests that we are the cerebrum and not the organism.

Merricks might respond in either of two ways: His first option is just to treat the identification of persons with cerebrums in the same manner in which he said he could accept identifying the person with the whole brain. He has written that that if compelled to adopt the claim that we were brains and not organisms:

> that gestures more at friendly amendment than at wholesale refutation. For it grants that we exist and think. It grants that we are composite macrophysical objects. And it poses no threat to my arguments against statues nor to our surviving those arguments by, for example, non-redundantly causing things in virtue of our mental properties. (2001:135)

Merricks didn't think he had to accept the identification of the brain with the person because organisms can become whittled down to the size of the whole brain. But since organisms can't become cerebrum-size, he doesn't here have the comparable option. However, identifying the person with part of the brain might not appear to him to be any more of a wholesale refutation than was the prospect of identifying the person with the whole brain.

The second move Merricks might make is to assert not that person was always identical to the transplanted cerebrum but that the person was an organism which has ceased to be alive when whittled down to the size of the cerebrum. Since Merricks believes we are only contingently organisms there would be no more reason to say during the cerebrum transplant that the person in transit won't be the organism than there would be to say that a fifteen year old

adolescent won't be a middle aged man in thirty-five years. Just as normal aging doesn't cause the adolescent to go out of existence or be left behind as a distinct entity, the transplant of the organism's cerebrum does not destroy or leave the organism behind and separated from the transplanted person. The adolescent moves into the future, though without the pubescent and other properties that characterize adolescents. Likewise, the organism could be transplanted but loses the properties of being alive that characterized it. "Organism" would be in Wiggins's language, a phase sortal (Wiggins 2001: 30-33). It is a term that picks us out in virtue of a property we have at one time rather than necessarily always have. Elsewhere, Merricks admits the possibility that we could become inorganic and thus no longer alive. He writes "our currently being organisms might be consistent with our later undergoing gradual replacement of our cells by circuits until we are not longer biological entities" (2001: 86-87). So this possibility would also supply a reason to say "organism" was just a phase sortal.

However, neither option is open to Merricks because we have independent reasons, as we saw in the last section, for granting the existence of organisms in the absence of thought. Merricks could have taken the first of the above options only if there didn't exist an organism that the cerebrum was a part of. And the second response would have been possible only if the organism doesn't survive the loss of the cerebrum in a mindless state but rather persists as a transplanted person who no longer engages in biological functions like homeostasis and metabolism. But if a mindless but living organism is left behind when the cerebrum is transplanted, then that means there were two distinct beings before the transplant. For surely a new organism didn't just come into existence replacing the one which had just lost its cerebrum. So if persons are not organisms, then they would be spatially located with organisms - or if they were cerebrum-size, then they would be parts of organisms. And the result would be two

thinkers, the organism and the person, where we would like there to be just one.[3] To avoid this, Merricks needs to deny that transplanted cerebrums can support thought. Then persons and organisms couldn't come apart and thus the argument for their being non-identical would be weakened.

It might be claimed that personhood is realized in more than just the cerebrum. So transplanting a cerebrum isn't a threat to Merricks's position even if organisms can exist before they can think. Since the brainstem is needed for there to be any conscious life, Merricks can argue that the transplant of just the cerebrum is insufficient to transplant a person. Only a whole brain, and thus a whittled down organism, can think and be transplanted. My response is that the brainstem is not required in the same way as the cerebrum is needed to realize thought and preserve identity. The brainstem is more like the power source for the computer. Different electrical outlets and batteries can all serve equally well so we wouldn't say the computer and its files aren't the same as a result of a change in power source. The cerebrum is special in that the person's biography and capacity for rational thought is realized there. Surgeons poking around in someone's cerebrum can elicit memories or desires but nothing comparable occurs with probing the brainstem. If a scientist

[3] Merricks states that if folk ontology was correct and brains did exist then "I think the brain is the best candidate for being the thinker. The organism seems to think only in a derivative sense (2001: 136 nt. 10)." I have doubts that this account of derivative thought as presented by Merricks will avoid the problems of too many thinkers. (See McMahan 2002 and Persons 1999 for support of this derivative thought approach and Hershenov 2005 for criticism.) This move isn't as harmless as saying cars are noisy because their horns are. It is true that there aren't two noisemakers there. But if the person thinks "I am essentially a person" or "I am the size of my cerebrum" then if the organism thought such thoughts derivatively it would be having false thoughts. Since a single thought can't express a single proposition that is both true and false at the same time, then there must be two thoughts where we would want just one. More later will be said about what other tinkering would be needed for the idea of a derivative thinker to work.)

replaced one's cerebrum with a duplicate it seems as if the resulting person would also be a duplicate. If a person's cerebrum is rewired, it would completely change his mental life. The brainstem, on the other hand, makes a very different kind of contribution to our mentality. It is more of an all or nothing contribution to consciousness. If it were rewired but still worked, this wouldn't be a threat that we've been replaced as in the cerebrum "scramble." We need a brainstem to be conscious but a mechanical surrogate would do since what is distinctive about our mental life seems to be realized in the upper brain.

However, if I'm wrong, even if personhood depends intimately upon other brain structures than just the cerebrum, my main point about persons being transplanted without being living beings can still go through. Additional brain structures can be transplanted with the person's cerebrum and the person still won't be alive but will exist during the transplant procedure when the cerebrum and those attached structures have been removed from one body and not yet put in another. As long as the transplanted parts of the brain aren't sufficient to compose a whittled down organism, then the person could be transplanted and separated from the organism left behind or destroyed. As we have mentioned, some philosophers, like van Inwagen (1990: 172-81) and Olson (1997: 44-46, 132-34), think a whole brain transplant is the transplant of a very small organism because the controls of vital biological processes are there in the brainstem even if they are not attached to any organs that they can control. So to avoid this threat of a pared down organism being transplanted thus undermining the claim of person/organism nonidentity, we must imagine that if parts of the brainstem are transplanted with the cerebrum, only the capacities of the brainstem necessary for thought remain intact while those structures required for integration of the bodily life processes are irreparably damaged and thus lost.

Merricks might then claim that since he isn't committed to our being essentially organisms, he can accept that such a transplant would be of a person that used to be a living organism, rather than spatially coincident with one. Merricks might claim that my modified brain transplant thought experiment only works against those like Olson and van Inwagen who claim that we are *essentially* organisms. He, on the other hand, doesn't have to accept that the remains of an organism would be left behind if the person with the damaged brainstem ceased to be a living organism during the transplant procedure. Now I think it is very odd to maintain that organisms can survive the loss of all their capacities to metabolize food, maintain homeostasis, assimilate oxygen etc. when they are whittled down and transplanted (Olson 1997: 118). But if organisms are only contingently organisms, as adolescents are only contingently adolescents, then survival through such a change is not objectionable given *those* assumptions. However, even if these costs are acceptable, I will use another claim of Merricks's about our not being essentially organisms to undermine the proposal that the person transplanted with the impaired brainstem is the very same being that was earlier a living organism.

Merricks presents another thought experiment to support his claim that we might be just contingently organisms that involves the person's organic parts being replaced with functionally equivalent inorganic parts. He writes:

> It is not because I endorse a biological criterion of personal identity that I can identify each person with an organism. Rather I make the identification because I oppose co-location. There is no co-location. So there is exactly *one* thing where we truly believe there to be a human person and a human organism. Obviously this implies the person is identical with the organism. But it does not – at least not obviously – imply that we have biological persistence conditions. For example,

> our currently being organisms might be consistent with our later undergoing gradual replacement of our cells by circuits until we are no longer biological entities (2001: 86-87).

This possible scenario involves the person ceasing to be composed of any organic parts like cells. Since the thought experiment assumes that the person's mental capacities have not been altered, there is a strong intuitive pull to declare that the person survives the procedure. Support for this speculation is drawn from the actual fact that throughout the life of an organic person all the cells except those of the brain are replaced, and even those brain cells eventually come to be composed of entirely new atoms. This complete natural change in material composition doesn't lead us to believe that the person has been replaced. I suspect the basis for our belief in the person's persistence is the presence of the *same* cognitive functions. So if the person can survive such organic part replacement in the regular course of events, then there is reason to believe that he would survive the acquisition of inorganic parts if they too were functionally equivalent to their predecessors.

We can now fulfill our earlier promise. We considered the possibility that Merricks would claim a cerebrum was insufficient for thought, the brainstem's reticular formation, responsible for awareness, was needed. So even if the brainstem that was included in the transplant of the cerebrum was no longer capable of being the biological autonomic control center for an organism, Merricks could claim that the organism wasn't destroyed by the transplant but persisted without being alive any longer. "Organism" would be construed as a phase sortal. But now that we have seen that Merricks allows for the possibility of a person surviving inorganic part replacement, we can imagine the entire organic brainstem being left behind, preserving the organism in one place and the person in the other. This would occur when

an inorganic brainstem substitute would be (gradually) attached to an inorganic substitute cerebrum while the organic brainstem was left intact governing the autonomic processes of the organism. The inorganic and now much smaller person would be separated from the organism that existed both before and after the loss of its thinking capabilities. That means prior to the inorganic part replacement the person and the organism were co-located.[4] So Merricks can not pull off identifying the organism and person and maintain that we are contingently organisms since the being that is the organism appears to again to be separable from the entity that is the person.

The Asymmetry Problem and Whole Brain Transplants

I want to raise one other problem for Merricks's account of brain transplants. I don't think he – nor van Inwagen, whom he credits as being the source of the idea - should claim that the transplantation of the whole brain (including the brainstem) is the transplant of an organism. There is an asymmetry problem in the account that Merricks and van Inwagen offer. This is also a problem that plagues the contemporary criterion for the death of the organism being the death of the whole brain and brainstem (Hershenov 2002; McMahan 2002: 426-33). Van Inwagen believes the organism comes into existence a few weeks after fertilization (1990: 154). But there is no brain at that time. It strikes me as very odd that the brain is essential at one time and not at another. If the human organism could have existed once without a brain as an embryo, though dependent upon its mother's body, I don't see why the human organism could not exist later in life in an Intensive Care Unit (ICU) in a "brain dead" state though dependent upon machines. It

[4] Or if we're just located where the brain (or part of it) is, there would still be co-location between the person and the brain (or its part) since the person can survive inorganic part replacement but the brain can't. Thus if persons can become inorganic in composition they can't be identified with the brain or any part of it.

shouldn't matter that the support is artificial rather than organic since in either case the embryo is a distinct object from its support.

Van Inwagen's position, which Merricks relies upon, is summarized by Olson as "biology tells us that the result of cutting away a man's head is not the headless living organism, but a mere heap of flesh, a headless corpse (van Inwagen 1991: 173-177). The heap is composed, for the time being, of living cells; but those cells are unable to coordinate their activities in the way that the parts of a living organism coordinate theirs" (Olson 1997: 132). The neurologist Alan Shewmon offers what seems almost like a tailor-made response to those who approach the brain dead body as if it was just a heap of flesh. He writes: "If anything, the idea that the non-brain body is a mere 'collection of organs' in a bag of skin seems to entail a throwback to a primitive atomism that should find no place in the dynamical-systems-enlightened biology of the 1990s and twenty-first century" (2001: 473). Shewmon argues that there is integrated functioning characteristic of organism in those clinically diagnosed as brain dead. He insists that the role of the brain has been overstated. The brain is a modulator and enhancer rather than an integrator. He concludes that somatic integration is not localized to any single 'critical' organ but is a holistic phenomenon involving mutual interaction of all the parts (Shewmon 2001: 473).

Shewmon describes how the brain dead patient's body is involved in "the homeostasis of countless variety of mutually interacting chemicals, macromolecules, and physiological parameters, through the functions of especially liver, kidneys, cardiovascular and endocrine systems, but also of other organs and tissues (e.g., intestines, bone and skin in calcium metabolism; cardiac atrial natriuretic factor affecting the renal secretion of rennin, which regulates blood pressure by acting on vascular smooth muscle" (2001: 467). He also points out how the brain dead eliminate, detoxify and recycle cellular wastes throughout the body. While

the brain dead can not keep their body temperature at normal, the mere addition of blankets can keep it a few degrees below normal; but blankets obviously won't accomplish this with a corpse (2001: 471). The brain dead fight infection and their wounds heal. When incisions in the brain dead are made to obtain transplantable organs there is a rise in blood pressure and heart rate. Perhaps most startling is that a brain dead pregnant woman can carry a child to term. Shewmon rebukes those who speak of the brain dead woman gestating a baby as a mere "human incubator" for that phrase does an "injustice to the complex, teleological, organism-level physiological changes of pregnancy (weight gain, internal redistribution of blood flow favoring the uterus, immunologic tolerance towards the fetus, etc.) which occur despite the absence of brain function."

So if the organism could exist without a functioning whole brain and brainstem, I don't see how Merricks and van Inwagen can claim that a whole brain transplant is the transplant of an organism. Or they, at least, have no grounds for saying that the whole brain is a better candidate than the headless body for being the original organism that was once a brainless embryo.[5] Thus not only have we seen arguments for the transplant of the person and not the organism when an organ (the cerebrum) is transplanted, but it is difficult for those who believe persons are organisms to insist that the transplant of the person occurs when the whole brain and thus allegedly the organism is transplanted. There are good reasons to believe that even whole brain transplants may not be transplants of the *organism* and that makes it much harder to for those who believe we are organisms to capture the transplant intuition that a person goes where the

[5] Hershenov (2002) considers the possibility that the detached whole brain and brainstem and the brainless body to both be organisms. This may actually be a case of induced fission. Shewmon also seems open to that possibility (2001: 474, nt. 6). I think it is best to say that the detached head is an organism that budded from the original (and surviving) organism rather than is one of two fissioned descendants, neither of which is identical to the original..

physical basis for their thought does. If a human organism didn't once need a brain, then it is hard to believe that the brain can become an essential part of it that will determine if and where it exists in the future. So while transplanting your whole brain seems likely to be moving you (a person) from one body to another, there isn't an overwhelming case that can be made that has the organism moved. The organism has stayed behind, a living though very dependent organism, in a state much like it was in as a very young embryo.

Merricks's Alternatives

It appears that if organisms and persons are not identical, then they are co-located. If co-location is possible, it would seem to bring us overdetermination since the person and organism would each seem to be causing what the other does. So what can Merricks do? One option is to accept co-location and allow that persons are spatially coincident with organisms. Merricks would understandably be reluctant to do this. For then why not elsewhere? But maybe he can eliminate his artifacts and other nonliving and nonthinking entities and just grant co-location and "downward overdeterminism" where the only overdeterminism is "from above." That is, person *and* organism co-location is acceptable, while person and microphysical overdetermination as well as organism and microphysical overdetermination are not. Merricks can avoid statue/lump spatial coincidence since both the statue and the lump would suffer microphysical overdetermination and thus there are grounds for eliminating both.

It may also be that person/organism co-location is not as bad as Merricks envisions. It may not involve, as Merricks seems to hold, two thinkers each ignorant of whether she is the organism or the person (2001: 50). It may be that both the organism and the person refer to the person by the first person pronoun. Merricks seems about to recognize this in his response to Noonan, but then he makes a claim that I don't think follows. He imagines an objector saying

"'I' can only have a person as a referent" (2001: 50). But his response is "that only makes things worse because one can't tell whether your I thoughts refer to you or the other…" However, if "I" can only refer to the person, then the organism can't possibly be confused about self-reference for it can't refer to itself. To ask which of two beings one is, is already to refer to oneself, it is just not to know what type of being one is. The organism can't do this if the person is the only entity that can refer to itself. So there can't be the errors Merricks's hypothesizes. His mistake is not to realize that the content and referent of "I" for the organism is the same as for the person. The organism is always thinking of the person when the co-located person uses "I". One could say as Baker does (2002: 42-43), that the organism derivatively thinks the person's thoughts, referring to the person when the person does. So if the organism can't refer by the first person pronoun to itself, then it can't wonder which being it is. Since direct self-reference is impossible (the organism can refer to itself only in the third person when the person entertains thoughts about the person's body), there is no problem of referential error. The organism can't even pose the question of whether it is the person or the organism. If a metaphysically confused person wonders whether or not she is a person, the organism wonders whether that person is a person.

Perhaps also mitigating the problem of too many thinkers is the recognition that those who believe we are organisms will have to suffer similar problems in a bizarre case of conjoined twins where there are two organisms sharing a cerebrum. They have different brainstems and share no organs or limbs beneath their shared cerebrum.[6] If the human organism is the subject of thought, then there will be two thinkers. And since the conjoined thinking organisms share a

[6] For a description see *Ultrasound in Obstetrics and Gynecology*. 18, 2001, pp. 289-290. A picture of a related kind of conjoined twins (Craniothoracopagus), one cerebrum and possibly two organisms, can be found at the following website: http://www.conjoined-twins.i-p.com/

cerebrum they would appear to have the same problems of self-reference that the spatially coincident organism and person will have as envisioned by Merricks. Whatever thought one has in virtue of its cerebrum, the other will have as well. If the two conjoined organisms can each succeed in referring to themselves by the first-person pronoun, so can the person and the organism in the spatially coincident case. If the conjoined twins cannot know whether the first person pronoun refers to the individual on the left or the right but that must be accepted, there is less reason to object in the case of person/organism spatial coincidence to their being unable to know which kind of entity each is. So it seems that those who identify persons and organisms, in part to avoid the problems of spatially coincident thinkers, will still have to allow that there are scenarios where there occur the same problems of self-reference, though less often.[7] Thus either each organism can't know which one it is, or the conjoined organisms constitute a single person and only that person can engage in direct self-reference as we speculated in the discussion of Merricks's remarks about Noonan. The two conjoined organisms may think derivatively, each referring by "I" to the person when the person does.

Merricks's second option is to become a full-fledged animalist like van Inwagen and Olson.[8] That involves giving up the notion that we can survive brain transplants and complete inorganic part replacement. But then one has to explain away what appears to be the acquisition of a new body in the scenarios involving cerebrum transplants and inorganic part replacement. The standard move is to claim that identity doesn't matter. But Merricks is on record as saying it

[7] And we shouldn't forget that there could be worlds where every person is such a twin so I don't see the opponent of co-location obtaining much support from the rareness of the phenomena. Metaphysics is concerned with necessary truths, so being free of a problem in the actual world but not others isn't a great plus for a theory.

[8] The animalist believes that we are essentially animals. The capacity for thought is irrelevant to our persistence. We can survive the loss of cognition but not the loss of involvement in life processes.

does, thus he must abandon that claim (2001: 125; 1999: 993-995). The standard argument for rejecting the claim that identity doesn't matter involves the case of cerebral fissioning. This story is probably is well-known to readers so I will be brief recapitulating it (Parfit 1983: 253-66; Olson 1997: 46-72). If a person can survive in a diminished state with one cerebral hemisphere after say a stroke destroys the other, then that surviving hemisphere could be transplanted and she would go with it if she would go with an intact (undivided) cerebrum. But if the cerebral hemispheres are divided and transplanted into different bodies, it is arbitrary to think she is one and not the other resulting person. And she can't be both post-fission persons on pain of violating the transitivity of identity since they are not identical to each other. So it would appear persons don't survive cerebral division and double transplantation. But it doesn't seem as bad as death. In fact, Parfit finds it hard to see why it would be worse than one hemisphere surviving. He actually calls it a "double success." He concludes that what matters to us is that our psychology goes on, not that we do. The animalist, most famously Olson, then claims that we don't switch bodies when there's the transplant of even an undivided cerebrum into a single skull, but are misled into thinking we do. What matters to us has switched bodies, but we haven't. So just as what matters to us survives in a brain split and double transfer, though we do not, Olson argues that in the whole cerebrum transplant what matters to us is transplanted and thus continues though we do not. A similar moral will be extended to the case of inorganic body replacements. What matters to us is continuing on, i.e., our physically based desires and beliefs etc., but we do not.

This second option preserves the identification of organisms and persons and can still rule out other composite objects since those would have redundant causal powers as a result of their macro properties supervening on the properties of their microphysical parts. It will allow

non-thinking organisms to exist while granting that the overdetermination argument still rules out the possibility of statues and mountains etc. However, it isn't easy to give up the claim that identity doesn't matter. I haven't been able to do it. I don't think cerebral fission and a double transplant is as good as survival. I would care less about the persons, each with one of my cerebral hemispheres, than I would in the stroke case in which only one of my hemisphere survives and I continue in a diminished state due to the loss of the other cerebral hemisphere. The difference is that it is *ME* in the latter scenario. I believe that this suggests that identity matters more and I don't have reason to give up the view that in transplant scenarios my concern is tracking identity. Moreover, it is very hard to believe that a new entity, a person, came into existence when the cerebrum of the earlier person was removed in the transplant scenario. It will have apparently all the same psychology but it will be wrong to believe it did any of the things it seems to recall doing.

I think readers can see that neither of the two options will be easy to defend. And both involve major changes in Merricks's account of persons and organisms.

References

Baker, Lynne 2002. On Making Things Up: Constitution and its Critics, *Philosophical Topics* 30: 32-50.

Hershenov, David 2002. Olson's Embryo Problem, *Australasian Journal of Philosophy* 80: 502-511.

Hershenov, David 2005. Persons as Parts of Organisms, *Theoria* 71: 29-37.

McMahan, Jeff 2002. *The Ethics of Killing: Problems at the Margins of Life,* Oxford: Oxford University Press.

Merricks. Trenton 1999. Endurance, Psychological Continuity, and the Importance of Peson Identity, *Philosophy and Phenomenological Research* 59: 983-997.

Merricks, Trenton 2001. *Objects and Persons,* Oxford: Oxford University Press.

Olson, Eric. 1997. *The Human Animal: Identity without Psychology,* Oxford University Press.

Perssons, Ingmar, 1999. Our Identity and the Separability of Persons and Organisms, *Dialogue* 38: 519-33.

Shewmon, Alan 1998. Brain Stem Death, 'Brain Death' and Death: A Critical Reevaluation of the Purported Equivalence, *Issues in Law and Medicine* 14: 125-45.

Shewmon, Alan. 2001. The Brain and Somatic Integration: Insights into the Standard Biological Rationale for Equating 'Brain Death' with Death, *ournal of Medicine and Philosophy.* 26:

Van Inwagen, Peter 1990. *Material Beings,* Ithaca: Cornell University Press.

Wiggins, David 2001. *Sameness and Substance Renewed,* Oxford: Oxford University Press.