

Introduction

If persons and animals are atom for atom the same, then if one can think so can the other. Neither is missing any parts that the other uses to generate thought. What has been overlooked in discussions of the *Problem of Too Many Thinkers* is that conflicts in interests will arise due to the person and organism having different persistence conditions. One may die or go out of existence before the other, thus resulting in their having different interests at the end of their lives. Since they physically overlap, the autonomy exercised by one could come at the expense of the other in a way quite unlike how you and I can thwart each other's autonomy. Your choosing to do something to your body doesn't involve doing it to mine and preventing me from refraining to do the same to my body.

Examples may be helpful. Since the Lockean person would go out of existence with the loss of self-consciousness, the organism could outlive it in some sort of unreflective conscious state. If there was a risky experimental Alzheimer's drug that was more likely to kill the person than preserve his self-conscious capacities, it could be in the person's interest to take it since he would go out of existence anyway. But it might not be in the interest of the organism who would survive in a more minimal conscious state akin to that of a young child. Matters are even worse if there is an epistemic problem and the animal and the person don't know whether each is the animal or the person. One can't make autonomous decisions when one confuses oneself with another. But the problem remains even if there is, as Parfit recommends, some referential machinery (using "inner I" and "outer I") that avoids the epistemic problem. Despite the animal and person knowing to whom they are each referring with first persons pronouns, they can't both simultaneously and autonomously endorse the pursuit of their conflicting interests.

It is easy to generate other conflicts that can't respect the autonomy of both the dying animal and person. The person might not want her resources spent on the minimally conscious organism after the person ceases to exist due to dementia. The person might not be able to target the donation of her vital organs to save a dying loved one after the loss of her personhood without causing the organism's death. If the organism survives the person's demise, the person couldn't have a quick burial required by considerations of religion or dignity. The animal and the person can't both autonomously endorse the same time that their loved ones should begin death behaviors such as mourning and inheriting. So obtaining the patient's informed consent will face considerable obstacles. I don't envy the attending physician on rounds having to teach his residents about "consenting" both the person and animal patients in each bed!

And the problems don't just arise for views that maintain the person is essentially self-conscious. Even if self-consciousness was a mere contingent trait, an essentially minimally conscious thinker would still lack the crucial autonomy to control his bodily integrity if he overlapped a thinking organism. Imagine a future where those with intact minds but pathological bodies could have their cerebrum transplanted to a healthy body or replaced with an inorganic body as transhumanists hope. Since the organism would be either left behind in a mindless state or destroyed, he would be better off enjoying a mental life with a diseased body as long as he could.

It might be thought these problems wouldn't arise if the person and the animal had the same interests. But why should the animal have the person's interests? If they are physically indistinguishable, then it would be surprising that the person had prudential concern and the animal didn't. Identity should matter for both or neither. If it matters for both, then there will be above conflicts. If identity doesn't matter then both the organism and the person should care about being psychologically connected and continuous to a future being regardless if they are identical to that being. However, the latter is implausible. Our attitudes to future pains after amnesia or strokes that leave us with childlike minds don't reveal the lack of concern that the identity doesn't matter hypothesis suggests. It is also odd biologically to believe all the other organ systems are functioning properly when contributing to the organism's survival but its cognitive system is not so directed. Our attitudes to the death of the minimally minded or mindless young don't reveal the indifference that one would suspect if they lost out on a future to which they were not psychologically connected.

So given the unwelcome consequences of the spatial coincidence of the animal and person, it might be maintained that we persons consist of those organism parts that are directly involved with thought. McMahan and Parfit think this may be just parts of the brain. Our organisms think derivatively in virtue of we persons thinking non-derivatively. To help the reader get a grasp on this idea of a derivative thinker, McMahan offers the analogy of the horn and the car of which it is a part. There might be two noisy entities, the horn and the car, but there is really only one noisemaker, the horn. Likewise, there is really only one 'thought maker,' the small person embedded within the animal.

I'll argue that this won't work. McMahan and Parfit's embodied mind view won't be able to keep the organism from genuinely thinking. They'll instead be forced to defend a sparse ontology and eliminate organisms and any other objects that overlap the brain or minimal thinker. I will spend most of this paper arguing that if a sparse ontology is required to make autonomy and informed consent possible, that a better case can be made for a sparse ontology that preserves organisms instead of roughly brain-size persons.

Thinking Substances and Sparse Ontologies

My contention is that McMahan's account of little persons that aren't spatially coincident with a thinking organism merely delays the return of the problem of spatially coincident thinkers. Some philosophers believe that the human animal could survive being pared down to the size of the brain – the animal would then be composed of every part of the brain and nothing else. So there would again be a problem of spatially coincident thinkers. If readers believe that the minimal thinking being is just a part of the brain such as the cerebrum and thus smaller than the smallest, maimed brain-size animal, the problem can still be reproduced with the brain and its minimal thinking part, rather than the animal and the brain. Brains can change their size.

The (undetached) cerebrum actually needs a good part of the organism for thought to be produced. In fact, I would contend it will need a portion of the organism that includes more than just the brain to bring in nutrients and remove wastes. Nevertheless, the production of thought

might not need all of the organism (hair, nails, toes etc.) However, whatever parts of the organism are irrelevant to thought production, their complement is probably large enough that the organism could survive a reduction in size to that thinking proper part. It would then be thinking non-derivatively if the spatially coincident person is. And I don't see why it would later cease to genuinely think if the amputated parts were restored. If the maimed animal earlier non-derivatively thinks, it is hard to follow the embodied mind view and claim that later when the animal became larger it no longer is really a thinker. It is not plausible that it ceases to strictly think and have its own interests but only thinks in some harmless derivative sense. Even after the organism increased in size beyond its thought allegedly generating parts and became a derivative thinker, it isn't clear that such derivative thought wouldn't really be genuine thought. It is typically believed that the organism digests in virtue of its digestive system non-derivatively digesting but we don't then claim that the organism doesn't really engage in digestion or a digestive disease isn't a disease of the organism. So why then deny that derivatively thinking is really thinking and such thinkers have their own interests? So the problem of too many thinkers isn't avoided by claiming the person is a spatial part of the organism.

It thus seems that the embodied mind view of McMahan and Parfit will require that the world doesn't include any organisms that could be reduced in size to that of the person. That means a sparse ontology which just has thinkers as the only composite or a very gerrymandered sparse ontology that allows composites besides thinkers but nothing that can be reduced to the size of a thinker. The latter is rather unprincipled so we will concentrate on the former answer to the special composition question.

Substances have typically been construed as capable, in some sense, of existing independently. Animals meet that criterion. It isn't clear that brains or their parts do. They won't be involved in thought production unless they have a body or mechanical substitute. (I would contend that the brain in the vat doesn't think, rather the brain/vat composite is the thinker). There is no reason to think a detached cerebrum can think when it is removed and prior to being successfully transplanted, even if it will be involved in thought production post-transplant. So a cerebrum is not as good of a candidate as the animal to be a thinker. Even an undamaged cerebrum won't think at least without a functioning brainstem, according to experts on brain death.

Brains and cerebra don't maintain themselves the way animals and organisms do, which is one reason to think they are not substances and perhaps not natural kinds. The animal or organism takes in matter, metabolizes it, builds up itself, removes waste products, uses some parts of the body to defend and repair others etc. The cerebrum or brain does little or none of this. It is utterly dependent upon the organism for its new parts, energy, removal of spent parts, repairs etc. That suggests they lack the independence constitutive of a substance. The lack of self-standing, autonomous, self-maintenance makes the brain and any of its parts too dependent to be a substance or natural kind and thus not as good a candidate for being the person as the animal. The organism is not dependent upon the environment in the same way as the brain is

upon the body to maintain, grow, replace, and remove its parts. In fact, the organism stands in relation to the environment as an occupant to niche rather than part to whole.

Sparse Ontologies, Development, and the Natural World

A sparse ontology of organisms better fits into our picture of the natural world than a sparse ontology of thinkers without organisms. Ontogenetically and phylogenetically it makes sense to posit the existence of mindless organisms that develop or evolve mental properties than to have no such entities and suddenly consciousness emerges and composites arise mid-pregnancy in the case of an individual thinker, or millennia ago in the case of thinking species. There is a nice evolutionary story of organisms arising from mindless to minded creatures. One can imagine minds providing certain survival and reproductive benefits. So the animalist's organism fits in nicely as a natural kind into evolutionary theory. The embodied mind's person doesn't fit as well. On the sparse embodied mind account, there are no animals that develop mental properties but thinkers just pop into existence when atoms are arranged brain-wise.

A similar natural story is told for the development of the fetal human being who acquires a mind with sufficient brain development. There is something living that develops into a thinker but this doesn't exist in the sparse embodied mind ontology. It also makes no sense in this ontology to say mindless creatures have interests in their health, survival and development. One of those interests that must be denied is that of a fetus in developing a normal healthy brain and becoming conscious. If the mindless animal has such interests in surviving and development, I would think that it was that animal which became conscious, not its cerebrum.

It is admittedly controversial to ascribe interests and well-being to the mindless. Nonetheless, if one doesn't, it is awfully hard to explain the harm of coma (or benefit of coming out of the coma) for harms (and benefits) are typically understood as comparing the well-being of two states, i.e., the well-being of the actual (the mindless state) and the counterfactual well-being (that of being conscious). So harm requires that mindless beings have well-being (perhaps zero) and interests.

Individuation, Development and Brain Transplant Thought Experiments

Our prudential concern towards our adult conscious animal in the future, including those times when it is without any psychological connections to the present, or even the same cerebrum playing a role subserving our future mental life, suggests that we human people are animals essentially. I maintain that once it is recognized that we could survive certain brain injuries and part replacements, we can resist the intuitive pull of two famous thought experiments that have provided considerable support the embodied mind account. The first involves your cerebrum being destroyed and replaced by a new cerebrum. The second thought experiment involves you swapping cerebrum with another person. Most people judge it to be that we wouldn't survive in the first hypothetical scenario but would do so in the second, though in a different body. I shall try to instead elicit intuitions that in neither scenario do we cease to exist or obtain a new body.

Consider that we have prudence-like concern for the stroke victim that would result from damage to our brain reducing its capacities to realizing mere sentience. Many philosophers believe this shows that it is mere consciousness or sentience, not self-consciousness that is essential to our persistence. I think instead that our prudential responses in such scenarios should actually be construed as showing that it is the criterion of biological identity across time that reveals our persistence conditions. Ask yourself whether your concern for your post-injury self with just a rudimentary mind really is due to your possessing the *same organ* that underlies consciousness or is it rather that it is just the *same animal* that is conscious? I think it is the latter and this can be seen by pondering the following twist that makes use of the well-known plasticity of the brain. Consider whether your reaction to the prospect of coming out of a stroke-induced coma with pain and pleasure sectors intact but no cognitive capabilities above this will be different if such sentience is a result of different parts of your cerebrum being rewired during the coma to realize pain and pleasure when you awaken. I suspect that most readers would have prudential concern despite different parts of the brain contributing to such sensations.

If you would have prudential concern for the same animal with different physical structures supporting sentience, then why should you react differently to your animal getting an entirely new functioning cerebrum in the thought experiment in which your original cerebrum is destroyed and a new one imparted? Readers might respond that it matters that the different anatomical structures, without which there would be no sentience, are in the *same* cerebrum. If so, consider a second case where, early in someone's life, in the absence of injury and before a web of beliefs and desires arises, different parts of a developing brain play a role in receiving and processing painful and pleasurable signals. Imagine one is in the brainstem, as Shewmon showed is possible and the other is in the cerebrum. Would it be correct to say there were two thinking beings *of the same kind as the reader* in one body? I suspect readers would say it is not.

Moreover, I doubt readers would assert that there is a new thinking being, one of the same kind as they that is produced by fusion if there is the later development of a self-conscious person who provides the respective pain or pleasure reports when either the sector in the brainstem or cerebrum is "stimulated. And for all we know, this is roughly what happens in normal child development. The initially physically dispersed realization and thus psychologically unrelated fragmented mental states of the baby are only later psychologically united as the older child obtains reflective access to the different states. The child can come to say that "I am in pain now and earlier had pleasant experiences," reflectively linking what before had been experienced without the capacity for reflection upon those experiences. We wouldn't maintain that the conscious states prior to the emergence of the unifying self-consciousness capacities didn't belong to the child. Embodied mind theorists, however, are committed to such fusions of thinkers in what may be normal development! Even if such conjectured development is not how we actually develop, our reactions to such a counterfactual assumption about ourselves does illuminate what we take ourselves to be: living human animals, rather than brain-unified thinkers. So there don't seem to be any reasons to identify ourselves with the feeling parts of the

central nervous system rather than holding that these pains and pleasures would be mine because they are subserved by parts caught up in the same biological life and belong to the same animal.

Individuating Minds and Lives

I suspect that advocates of the embodied mind view may be misled by the truth that thought could continue if the animal is reduced in size into accepting the falsehood that such removals show that it is only some of the parts of the animal that produce thought. The mistake is not to appreciate that what earlier made those amputated toes and fingers into parts of the thinking animal are the same life processes that integrate the neurological parts that McMahan, Parfit, Hudson et al think produces thought. The animal needs to be alive to think. Following van Inwagen, let's give the label *Life* to the event consisting of the biological activities which distinguishes a living human animal from a dead one. *Life* contributes to thought. And *Life* is dispersed throughout the body. Since processes don't think, the thinker is the combined matter caught up in *Life* that makes thought possible. The fact that the event of someone's biological life could configure less material than it does is irrelevant. While it is true that *Life* can involve less matter, i.e., someone can become smaller, that doesn't mean that the life event which makes thought possible was not earlier an event of a larger substance. Since one's thoughts depend upon *Life*, wherever that event is located, so is the thinker of those thoughts to be found. It would be blatantly false to say that the life processes are found only in the central nervous system. We must recognize there are organ systems essential to *Life* that extend beyond the central nervous system, the latter system contributing to thought in virtue of the former providing it with the biochemical necessities for cognitive activities. These biochemical processes are constitutive of thought, not just causally downstream from the production of thought. There is not thought if the brain is not assimilating oxygen, releasing energy, removing waste etc. So it is *Life* that makes thought possible, not a part of it. And the same life that assimilates, maintains and removes the matter necessary for neurological function, also renders toes and fingers part of the living animal.

Another difficulty with positing a minimal thinker has been explored by Eric Olson. He thinks the real problem with brain-size persons is that little sense can be made of the idea of "direct involvement in a being's thinking" that motivates the position. Olson wonders why if the respiratory and circulatory systems are not *directly* involved with thought, we should consider the oxygenated blood vessels in the brain to be so? Olson suggests that someone might maintain that the thought is really produced by the firing of neurons. However, Olson points out that not every part of the neuron is similarly involved in the sending of electrical or chemical messages to other neurons. Some serve other tasks like maintaining structural integrity of the cell or removal of its wastes. This, Olson claims, ought to make "the thinking minimalist uneasy". Moreover, the neurons won't fire without these tasks being performed. Olson cautions that trying to determine what is *directly involved* in the production of thought is as hopeless as trying to determine which of the many workers, suppliers, managers, tools and materials is directly involved with the factory production of a knife, or which parts of the body are directly involved with walking. He insists that the problem is not even one of vagueness - it is not that we have a clear application

and then boundary cases. Instead, the fault lies in the notion of *directly involved* being unprincipled.

Lives, Fission, and The Only A and B Rule

Embodied mind theorists don't fare well with fission scenarios. Consider the splitting of the cerebral hemispheres and a double transplant. The typical response is that there are two new persons. However, if only the left hemisphere was successfully transplanted, the right hemisphere destroyed in the process of dividing and removing the two hemispheres, the original person A would be identical to person B with the left hemisphere. And if the right hemisphere had been split off from the left hemisphere and successfully transplanted while the left was destroyed, then the original person A would be identical to person C with the right hemisphere. It is commonly held that a person can survive the loss of a hemisphere due to stroke or cancer etc.

However, Parfit and McMahan's appeal to brain size persons runs afoul of the rationale behind "the only A and B rule." That rule does not allow that our identity can be determined by whether there are two or more equally good candidates as there would be in cases of fission. The rule restricts questions of whether A is identical to B to the internal relationship between A and B, the existence of a C being irrelevant. The rationale for the rule is that there should not be unexplained existences where entities owe their existence to other beings despite the absence of a causal connection between them. The problem with Parfit's cerebral fission and transplant case is that the person in body B would not be there if it wasn't for the existence of the person in body C likewise being psychologically continuous with Adam. So the person in Body B owes his existence to the person in body C, and vice versa, but there are no causal connections between person in body B and the person in body C despite the existence of each playing a role in the creation or sustaining of the other.

Organisms avoid violating the only A and B rule because when they split the facts relevant to their persistence have nothing to do with whether there are one or two candidates for being the original organism. For example, the cell has died when crucial life functions cease in the preparations immediately prior to fission, regardless of whether there are one or two resulting living cells. When the round worm is cut in half, the movements of the two halves mislead, there are not *immediately* two living organisms. The two halves have to reconstitute themselves from parts of organisms into organisms. They are akin to plant cuttings that are not immediately transformed from plant parts into plants. They must grow roots and do the other things characteristic of a self-maintaining organism. So if we choose between sparse ontologies, better to accept the one that doesn't violate the rationale behind the only A and B rule.

Cognitive Science Objections to the Embodied Mind Account

The animalist position that thought is produced by activity of the organism rather than its brain. Consciousness is not something that happens inside us, it is not something that the brain secretes, but something we make or achieve. As Alva Noë provocatively suggests "Consciousness is more like dancing than digesting." This is supported by a revealing

experiment of Paul Bach-y-Rita that enabled the blind to have a vision-like experience. The experiment involved placing a camera that caused vibrations on the torsos of the blind which enabled them to perceive objects, even becoming capable of swatting moving ping pong balls. The vibrations weren't processed by the so-called somatosensory cortex as body touches but as the visual field in front of them. While color is missing in the tactile-vision substitution system (TVSS), spatial information is provided in a very visual way. As you move closer to an object, its apparent tactile/visual size increases just as if you were seeing them. As you turn to the left, objects in 'view' swing to the right in your tactile-visual field, just as they would if you were seeing them. The TVSS perception is like seeing because "they are similar ways of *exploring the environment*: they are governed by similar sensorimotor constraints, draw on similar sensorimotor skills, and are directed toward similar visual properties."

Noë explains that this perceptual plasticity without neural plasticity serves to undermine the dogma that our consciousness is a neural correlate. Noë surmises that "what makes experience the kind of experience it is – is not the neural activity in our brains on its own; it is, rather, our ongoing dynamic relation to objects... We *see* with the Bach-y-Rita system because the relationship that system sets up and maintains between the perceiver and the object is... the sort of relation we bear to things when we see them". What is important and what serves the dynamic sensorimotor interpretation is that transition to quasi-visual perception depends on the subject's exercising active control of the camera. If the camera is stationary, or if someone else controls it while the subject passively receives tactile inputs from the camera, subjects report only tactile sensations.

It shouldn't be countered that dreams show that consciousness is not a dynamic bodily production but something that just transpires inside us. First, it doesn't follow that every experience can also be dreamt. Secondly, Noë stresses that normal perceptual experiences have a stability that dreams lack for it is the world rather than our creative imagination that provides the details. Noë emphasizes that the world is given to perception as available. "What we experience *now* goes beyond what we represent now in consciousness. Detail, three dimensionality, color, are present in experience not as represented but as available." The content is there only potentially. It might be enacted because of the perceiver's embedding in the world and the perceiver's possession of skills...needed to assemble it". The world is present in experience virtually the way information from a remote computer is. Experiential presence is virtual *all the way in*. This availability of perception explains why dreams lack a stability of details. The upshot is that a neurological duplicate brain just appeared, it would not have our perceptual experience. The previous duplicate state wouldn't cause a duplicate of our brain's next state which depends upon the environment that it is embodied in.

Conjoined Twins

The brain's role in producing thought is not the only weapon the embodied mind theorists use against animalists. They appeal to dicephalic conjoined twins to argue that persons aren't organisms. They focus on two headed twins sharing a single body. The thoughts involving either

of the two heads are as cut off from each other as the reader's thought is from mine. The animalist has to claim this is just one person cut off himself. That will strike many readers as an implausible interpretation. They will find it more intuitive to treat each head as belonging to a different person. The possessor of the mental life supported by the other head would seem to have no prudential reason to care about a mental life whose contents are as causally uninvolved in the mental contents of two people with different bodies. McMahan is led by this to claim: 'because there is no reason to suppose that the dicephalic twins are a different kind of entity from ourselves, or that a different account of personal identity applies to them, we should further conclude that we are not animals either.'

Ironically, the conjoined twin scenario provides the animalist with the resources to explain away much of the initial appeal of an approach that insists that mental contents cut off from each other must belong to different thinkers. Since the two heads belong to the same animal, then if the animal can think, it can think with both brains. We would describe this thinking animal as one individual whose mind is divided, its thoughts cut off from each other. So, surprisingly, the very scenario which the brain-size persons view provides in support of the claim that the person is a substance distinct from the animal guarantees the existence of a creature with a divided mental life! Thus the advocates of the brain-size persons view can't reject the animalist account of the dicephalus on the grounds that its positing a single thinker with a divided mind is implausible. Unless supporters of the brain-size person view have other arguments that can deny genuine thought to the human animal, they will have to admit that their theory posits a thinking being cut off from its own thought, thus undercutting the initial appeal of their theory when contrasted with the animalists in their handling of the two-headed animal.

Admittedly, it is hard to believe that the dicephalic twins are just a single person and even harder to believe that one should be concerned for a mental life from which one's present mental contents are permanently cut off. But it may help to think of our attitude to our dream life. Imagine that every night when you go to sleep, the tortured grimaces, groans, and electrical readings suggest terrible nightmares have ensued but you don't recall them or suffer any visible effects. Nor does the sleeping person undergoing such nightmares entertain during his dream any memories from your waking life. Wouldn't you still be prudentially interested in some treatment that avoided such terrible pains even though you have no waking awareness of the torment? I suspect many readers would have such prudential concern and this concern elicits the assumption of their being identical to the sleeper. Likewise, the dicephalic subject should be concerned with thoughts not psychologically continuous with half of his other thoughts.

Conclusion

If the human person is identical to the organism and there are no other entities overlapping the organism that are large enough to produce thought, then end of life decisions will be made by only one thinking being. So if you want to confront death as an autonomous subject, then you had better hope that the sparse ontology of animalism provides the right account of personal identity. Or you should become a soul theorist.