ANIMALS, PERSONS AND BIOETHICS

I. Introduction

My contention is that considering a person to be co-located with an animal, or one of its spatial or temporal parts, gives rise to a host of problems as a result of there then being too many thinkers. These problems, which Olson (1997, 2007) has emphasized, can be mitigated (somewhat) by a Noonan-style pronoun revisionism (2003, 2009). But doing so will have very unwelcome consequences for bioethics as autonomy, informed consent, advance directives and substituted judgment will be impossible for the human animal. I count it as a point in favor of Olson's answer to the metaphysical question "What are we?" that it avoids such ethical quandaries. But his animalism - with its Parfit-inspired claim that it is not identity that matters in survival but the continuation of our psychology even if someone else is its subject - appears to be at odds with our self-conception and practical concerns.[1] And if the only argument for this thesis is the fission scenario, then the thesis is further undermined as Parfit's account of fission runs afoul of the rationale behind Wiggins's Only a and b rule. What I will very tentatively suggest is that we explore an alternative account of animalism which denies that being identical to a future being is only of derivative importance to us.

II. Pronoun Revision and its Bioethical Problems

Assume that we are essentially thinking beings that are related to but distinct from animals. The problem which arises is that if the person can think, why can't the animal think since it shares the same brain? Accounts that deny thought to human animals (Shoemaker, 1997) or give animals different thinking capacities than persons (Baker, 2001), make a mystery of thought. I don't see any explanation of why the human animal can't do what the human person can. Usually when we point out cognitive differences between two entities, we appeal to a difference in their cognitive hardware, i.e. their brains, or their environments. But the human

animal and the human person have the exact same brain and are even related in the same way to the same environment. So the cognitive differences are unexplained, just brute. The person has capacities the animal doesn't just because it is a person. I think that we should be suspicious of such brute differences in psychology between physically indistinguishable composite entities.[2]

The problem of too many thinkers will not be avoided by claiming that persons are just parts of animals. Although this approach of identifying a person with part of the animal is not as common as that which assumes the spatial coincidence of a numerically distinct person and animal, there are prominent philosophers who defend it. Some claim we are three-dimensional brains or three-dimensional parts of such brains, others claim we are four-dimensional brains, our parts spread out in time as well as space.

> Michael Tye asserts: "On my proposal, then, persons are neither egos nor bundles. They are the subjects of appropriately complex psychological bundles. In actual fact, I claim, these subjects are brains insofar as those brains are in the appropriate physical states." (2006: 143)

> Jeff McMahan states: "One possibility is that the relation between ourselves and our organisms is the relation of part to whole. This suggestion will seem most cogent if we assume that the mind is entirely reducible to certain regions of the brain. If for example, the mind just *is* those regions of the brain in certain functional states, and if I am this mind, then I am, in effect, this functional brain, which is itself a part of this organism; therefore I am a part of my organism. (2002: 92).

> Hud Hudson argues: **"**Human persons …. are those (spatially and temporally gappy) space time worms that are certain proper temporal parts of the brain and central nervous system of living human organisms." (2001: 147)
>
> Ingmar Persson concludes: "In fact, its brain, or even a certain part of the brain seems minimally sufficient for a mind to be realized…The whole organism is not the minimal mind owner…The form of the separability of these subjects from their organisms that I have tried to vindicate is: (S) Predicates of thinking and experiencing are primarily applicable only to the proper parts of our human organisms because these parts are all that is minimally sufficient for their applicability; they are only derivatively applicable to our human organisms in virtue of their having parts." (1999: 522, 524, 527)

What motivates such claims, at least in part, is the belief that the brain is what directly produces our thought. The brain has been described as the "organ of thought" and even said to "secrete thought as the liver secretes bile."[3] If we are the beings that strictly think our thoughts, then we are brains, or the parts of brains that are directly involved in the production of thought. Some have found this view attractive because it seems to avoid the extra thinker posited by theories that allow person and animal to be spatially coincident, that is, atom for atom the same. I will concentrate upon the claims of McMahan since he has applied his views to bioethical issues more than the others. McMahan claims that the problems of persons being co-located with the animals constituting them can be avoided if persons are considered to be just parts of animals. Animals will think only derivatively in virtue of having parts that strictly or nonderivatively think. McMahan believes that attributing thought to the animal is no more problematic than

claiming a car is noisy because its horn is noisy. Just as there aren't two noisemakers in the case of the car and its horn, there aren't two *thought makers* in the situation of the animal and the embedded person. My contention is that this will just move around the lump in the metaphysical carpet. If persons are material entities no larger than brains, then persons and the part of the brain composing them would seem to be distinct since they have different persistence conditions. Brains, commonsense ontology has it, can survive the loss of the capacity of thought.[4] Persons cannot. So it looks as if McMahan will still have to countenance spatially coincident thinkers, though these are not the person and the animal.

Moreover, even if the alleged difference in persistence conditions of brains and persons isn't a problem for McMahan's view of an embedded person, it only delays the return of the problem of spatially coincident thinkers. If it is correct to maintain that the human animal could survive being pared down to the size of the brain - a view held by philosophers such as Olson (1997:45, 133) van Inwagen (1990: 172-181) and Merricks (2001), as well as the neuroscientists Shewmon (1997), Damasio (2005: 228) and Bernat (1998: 19) – the animal would then be composed of every part of the brain and nothing else. It is quite odd that the animal would *then* only be derivatively a thinker. (Not so for Hudson's 4D worm account since there will be stage sharing rather than spatial coincidence though there will be a problem that the worm thinks derivatively in virtue of the stage) This queerness could be reinforced if it were metaphysically possible for a functioning brain-size animal to be made first, and then head, neck, trunk and other appendages added later. It is very hard to explain why at the early stage only the brain-size person would be strictly (nonderivatively) the thinker. It seems that there would be at least two brain-size entities genuinely thinking the qualitatively same thoughts. If the animal then non-derivatively thinks, it is hard to follow McMahan and later claim that when the animal became

larger it no longer is really strictly a thinker but is only a thinker in some derivative sense. **(**Or at least there will be two non derivative thinkers for a time, even if one ceases to be so.)

Incidentally, if readers believe that the minimal thinking being is just part of the brain 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. Your brain could become smaller but still exist if it was reduced to being no larger than the part of it that composed your person. Even if McMahan identifies the person with a brain part, the problem of spatially coincident thinkers would reemerge if the larger brain lost some matter. The brain would *not* then become identical with what before was just one of its parts. Instead, there would then be two thinking beings composed of the very same brain stuff.

I doubt that there is a convincing account of why the (completely or partially) overlapping beings don't all have similar mental lives. Olson draws our attention to four specific problems if there are any such additional thinkers. There is first the *duplication problem* that the thinking animal would seem to meet the same cognitive criterion for being a person – it too is self-conscious, rational, free and responsible etc. Secondly, there is the *trivialization problem* if the animal is denied personhood because it has the wrong persistence conditions. Personhood then becomes insignificant for there would be non-persons that were also self-conscious and moral etc. Third, there would also be an *epistemic problem* for an individual wouldn't have any reason to believe that he was the person rather than the animal. Any reason the person had to think he was the person, so would the overlapping thinking animal sharing his thoughts. Finally, there would be the related *false self-ascription problem*. A truth expressed by one about its essential nature would be a falsehood simultaneously espoused by the other.

To mitigate the problems of too many thinkers, some sort of Noonan-style pronoun revisionism is required. This is true for four-dimensional accounts like Noonan's own in which temporally extended individuals think in virtue of stages thinking, and for spatially coincident animals and persons in constitution approaches like that championed by Baker,[5] as well as in theories like McMahan's in which people are construed as spatial parts of organisms. Noonan suggests that to have thoughts about thoughts is not enough to make an entity a person, rather an individual must have the appropriate psychological persistence conditions. So the referent of the personal pronoun "I" is not the thinking animal, thinking brain, thinking stage or a less than maximal psychologically related sum of stages, but the person with the appropriate psychological persistence conditions. As a result, while there might be many entities thinking one's thoughts, they all refer to the same person. This avoids the *duplication problem* of two persons, one essentially a person and the other contingently, because the non-persons are only conscious of the thoughts belonging to the person, not of their own thoughts qua animal, qua brain or qua stage. Unable to use the first-person pronoun to think about themselves as themselves, they don't meet Locke's criterion for personhood.[6] And since the non-persons can't self-consciously refer to themselves by the first-person pronoun, the *trivialization problem* doesn't arise. There also isn't an *epistemic problem* of a thinking animal, person, brain or stage wondering "Am I the person? Or could I be the animal? Or am I my brain?" This wouldn't occur since they all recognize that the referent of the first person pronoun is the person. For the same reason they avoid the *false self-ascription problem* when they claim "I am essentially a person." The animal and the brain have not said something false and the person something true. Rather they all said something true about the person.

While I am unconvinced that pronoun revision removes the problems posed by overlapping thinkers, I'll relegate my doubts to a footnote and turn to the bioethical dilemmas that arise from a metaphysics adopting pronoun revisionism.[7] (Incidentally, some of the same bioethical problems will arise in the absence of pronoun revisionism due to the *epistemic problem*.) If human animals can't refer to themselves with the first-person pronoun, then how can they be said to autonomously agree to any immediate treatment or make provisions for their future with say a living will? While I don't have a favored theory of autonomy to expound, it would seem safe to say that one couldn't be autonomous if one could not reflect upon one's interests, desires and reasons as *one's* own. Since the autonomy literature often runs parallel to the free will literature, what makes autonomy impossible will, in many cases, also make free will impossible. Without free will there will not be moral responsibility and so our ethics will be turned upside down. To see the importance of self reflection, just consider the 'taking ownership condition' in Fischer's semi-compatibilist account of free will, or those approaches that insist freedom involves identifying with one's higher order values or volitions. And if there isn't autonomy, there won't be informed consent which is so important to the doctor/patient relationship. As the author of a well-received book on the topic has said: 'informed consent is the cutting edge of the patient autonomy movement" (Wear 1998: 2). Patients are recognized as having the right to choose one form of medical care over another, or even to refuse any treatment. According to a classic statement of informed consent, a decision is just that if "It is given by a patient or subject with 1) substantial understanding and 2) in substantial absence of control of others 3) intentionally 4) authorizes a professional (to do intervention I)" (Beauchamp, Faden, King 1986: 278). So if the human animal is not able to think: "*I*, the human animal, am the patient… *I* intentionally, on the basis of *my* understanding the relevant information, freely

authorize *my* doctor to so such and such …to me, the human animal," then it can't be considered to have given informed consent to its medical care. And according to pronoun revision, the human animal is not able to use first person pronouns to do that.

This will also lead to a problem with advance directives.[8] The directive will be made by the person, but dictate the treatment of an individual that remains when the person has been extinguished by disease or injury causing the irreversible loss of self-consciousness. Brock and Buchanan call this *the Slavery Argument* (1990: 158). DeGrazia less dramatically labels it *the Someone Else Problem*. DeGrazia thinks this is only a problem if a merely sentient but unreflective *post-person* pops into existence with the extinction of the person due to the loss of the capacity for self-consciousness. DeGrazia doesn't believe there is a *Someone Else Problem* if animals are spatially coincident with persons. He writes:

> "For on the two substance hypothesis, the human animal, who eventually becomes the demented patient (after the person is extinguished *sic*), was there along. Since he and the person share the hand that writes, the human animal also signed the advance directive, just as the human animal, who shares a functioning brain with a person, was competent to give voluntary informed consent."(169-70)

But DeGrazia is wrong to assert this. Pronoun revisionism means the animal doesn't think of itself as the animal and thus doesn't consider its animal interests.[9] It merely offers a "shadow endorsement" of the person's intentions.

The problem is not just with advance directives. Everyday care of the human animal becomes problematic if the person and the animal both use the first person pronoun to refer to the person. It would also seem that relying upon what is called "substituted judgment" to determine what should be done to the human animal when it cannot make its wishes (qua animal)

known would be impossible. Substituted judgment involves deciding for the patient as he would have when competent and autonomous. But if the animal could never be autonomous, then the only way to decide for its care would be to rely upon the "best interests" doctrine, generally considered a less attractive option (Degrazia: 163; Beauchamp and Childress:135-140).

One might respond that the person and the overlapping animal's interests are the same, so the animal's inability to construe itself as the subject of thought is not a bioethical problem. I think it is more likely that human animals (and the other thinking non-persons) don't realize they have distinct interests because of pronoun revisionism or the *epistemic problem*. Just because they don't protest that their interests are being neglected provides us with no more reason to deny their interests are being ignored here than it would in cases of the brainwashed or constitutionally unreflective. For instance, I don't see why it is not true that embedded perduring objects are having their immediate gratification sacrificed for that of the four-dimensional worm composed of them.[10] Nor do I see why it would be in the three-dimensional or four-dimensional animal's interest to acquire a new body to avoid a non-fatal but physically incapacitating disease. Let's imagine a future in which the person's healthy brain could be transplanted from a sick body into the empty skull of a physically healthy twin or clone animal.[11] While the person might embrace the opportunity to survive with a new body, the animal "seconding" that treatment will be endorsing the end of *its* mental life in the transplant scenario, or the end of its actual biological life in the (the slightly more medically realistic) case of too many inorganic prosthetics replacing its organic body.[12] An even more medically realistic case might be if there is an experimental drug that may prevent the further decline into Alzheimer's disease, but will far more likely kill the users. The person, who goes out of existence anyway with the loss of self-consciousness, might think she has nothing to lose since either the disease or the drug's unwanted side effect

will end her existence. However, it may be in the interest of the animal not to take the drug since it could survive with the minimal sentience of late stage Alzheimer's disease [13]

Moreover, if the person and the animal both support donating organs at their deaths, the possibility of the animal and person's deaths occurring at different times could prevent the full realization of their shared value.[14] This is clearest in the case of the spatially coincident thinking animal and person; but may even arise in the case of the 'little' thinking persons embedded within animals. McMahan insists that organs can only be taken for donation with the person's consent. On his view, the person would die and go out of existence with the loss of the capacity for consciousness brought on by a permanent vegetative state. He believes that the person's interests can be violated posthumously, even though the person doesn't exist then, and a fortiori, has no conscious states. But McMahan claims that the mindless animal can't be harmed and that "there would be no moral objection to killing the unoccupied organism in order to use its organs to save others (2006: 48)."[15] I find it very peculiar that a person could have a non-experiential interest in whether its organs are transplanted after it no longer exists, but the human animal that has become a mindless organism does not have such an interest even though it still exists.

McMahan's argument would only be obviously correct if the organism could never think. I doubt that his argument works even if the organism only thinks, as he claims, in a derivative sense by having a thinking part, the brain. Anyway, we have seen that a parallel problem of two nonderivatively thinking entities will plague McMahan's account even if that second thinker is not the animal. It instead would be the brain that could be reduced to the size of the person. The brain would thus be capable of nonderivatively thinking before the onset of

mindlessness and the person's death. So the brain, as well as the person, could have a non-experiential interest in its organs not being taken prior to its death.

A similar problem will arise if the human animal and person's religious beliefs or conception of dignity demand some immediate posthumous treatment such as next day burial.[16] There is no recourse here to a "conscience clause" that some states have implemented to handle different conceptions of death amongst their citizenry.[17] Such conscience clauses only help a single individual that doesn't overlap another to choose between different criteria for death. They are no help where a single body is shared by two entities that can die or go out of existence at different times. The "remains" of the person who died due to say the onset of a permanent vegetative state would be the living body of a mindless animal. Satisfying the person's wish for a speedy burial would come at the expense of the animal's desire to for its speedy burial not to come until *after* its own death.

If advocates of three-dimensionalism or four-dimensionalism try to avoid dilemmas like these by arguing that the animal is just concerned with the welfare of the person, then they will be positing an animal in the transplant, prosthetic replacement and Alzheimer drug cases that doesn't care about itself since it appears to be nonplussed about the prospects of being either left behind in a mindless state or destroyed. I find it very hard to believe that with creatures so much alike as the human person and the human animal, it is in the interest of one to be identical to a future being but it is not in the interest of the other. It seems that they should either both care about themselves, or it should matter to neither of them that they are each identical to a future being. If the former is chosen, there are the just mentioned moral dilemmas. But if the latter is chosen, I don't see why a reader would then resist the claim there is only one intelligent entity where he is, and it is essentially a living being, rather than essentially a thinking entity. Doing so

would mean that one doesn't have to countenance spatially coincident animals and persons. Moreover, if identity doesn't matter, then the brain transplant and inorganic replacement body thought experiments will fail to provide such compelling support for psychological approaches to identity. For example, if our being identical to a entity is not the basis for a prudence-like concern, then our concern in the transplant scenario for the entity that ends up with our brain would not be evidence for the conclusion that we have switched bodies and thus are essentially thinking persons.

Leaving aside issues of patient autonomy, interests and prudential concern, consider the havoc that too many thinkers bring to families and friends. Do they start to grieve when the person dies (at the onset of a permanent vegetative state) or not until the human animal does, or do they grieve for both but to different extents? The latter would be quite different from the two stages of grieving for a single individual whose mental life might be extinguished before his biological life. Even with pronoun revisionism, family members and dear friends conversed with and were loved by both the non-identical animal and the person.

It thus seems a good number of metaphysical and bioethical problems can be avoided if there aren't any thinkers overlapping animals. So the animalist can argue that the advocates of psychological accounts of identity don't do as well by our practical concerns as they claim because, in all likelihood, they have to accept that other creatures share our thoughts and intend our actions. As a result, they must either tolerate identity not mattering to animals and other person-like non-persons, or suffer the ethical problems if such creatures do have interests of their own.

III. Animal Identity Matters

A metaphysical theory is more attractive the better it coheres with the rest of our beliefs. Unfortunately for the Olson-style animalist, the well-known thought experiments seem to elicit beliefs that we are essentially thinkers. To offset these intuitions, Olson must claim we are misled in the hypothetical cerebrum transplant and inorganic part replacement thought experiments by our concern for psychological continuity. Olson claims identity is really only of derivative value for us. Ordinarily, identity is correlated with what really matters, psychological continuity, but this is merely a contingent correlation and they can come apart in the fission thought experiment where the cerebrum is split and both halves transplanted. Fission cases, Olson alleges, teach us that identity does not matter and so our concern is not metaphysically informative even in the case of the transplant of an *undivided* brain (1997: 56) However, I, like many others, fail to share Olson's intuitions about identity not mattering. I want to survive into the future and find little comfort in a merely qualitatively identical replacement. Identity, as Unger argues (1990), seems to be a precondition for much of what we value, it is not something only of derivative value. I also agree with Baker that "our practices of apologizing, promise keeping, and intending become incoherent if we suppose our interest in identity is really only in psychological continuity (2000: 129).[18] And some of Parfit's own examples (1983: 264) about fission enabling someone torn between two activities to satisfy both seem not to work. He imagines someone who couldn't fulfill both his desire for a career as a novelist and a philosopher, doing so by splitting. But I would expect that the products of fission would likewise each be torn between the two careers.

I think the attitude that identity really does matter is very evident when contemplating one's young son or daughter splitting because concern for the well being of offspring is more

clearly dependent upon their identity being preserved than their psychology continuing. We don't come to love our children in virtue of their psychology and we would continue to show the same great concern if they underwent radical psychological discontinuity. But if they cease to exist via fission, our concern won't transfer undiminished to their successors.

Moreover, I suspect that if the argument about identity not mattering is based on the famous fission scenarios, then it is flawed for the reason Hawley (2005) gives: it leaves unexplained correlations between distinct existences. Each of the fissioned or branching-produced individuals exists only because of the other but they are without causal connections. Hence the appeal of Wiggins's Only A and B rule. That is, whether person A survives as person B should depend only on the relations between A and B and not upon the existence a qualitatively similar individual elsewhere. Hawley tries to explicate the intuition that there is something fishy about positing a no branching, uniqueness type of clause that is found in Parfit and Sydney Shoemaker and Nozick's closest continuer accounts of identity to deal with fission cases. She stresses that there are unexplained correlations where things are dependent upon each other for their existence or nonexistence, but in a noncausal manner. So if person A would be the person in body B if it wasn't for a psychologically similar competitor person in body C, then the person in body C can prevent A from surviving without any causal interaction. And the person in body B would not be that person if it wasn't for the existence of the person in body C likewise being psychologically continuous with person A. So the person in Body B owes its existence to the person in body C, and vice versa, but there are no causal connections between the 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. Contrast that with the counterfactual dependency of you on your mother.

You would not exist if she didn't, but there is a causal story connecting her to you. Not so with persons in the B and C bodies.[19]

It is also worth adding that the animalist's claim is, ironically, bad or, at least, peculiar biology. I would claim that survival is in the interest of a mindless animal just as water and sun is in the interest of a plant. But according to the Olson-style animalist, when animals develop significant cognitive function they aren't nonderivatively concerned for themselves. What they come (or ought) to really care about is their psychology continuing, not themselves as the subject of such thought. I think this is an odd sort of disconnect that animals at one stage in their ontogentic (and phylogenetic?) development have survival as a good (which then must obviously be nonderivative) but come later to care only derivatively about their own interests and persistence.

Readers may be concerned that I am equivocating in my discussion of interests. But I am aware of there being two senses of "interest." I have in mind Tom Regan's example of a kid who is not interested in broccoli but broccoli is in his interest. My thinking is that the kid could come to be interested in his interest, i.e. come to desire and be concerned with the good that broccoli served. Likewise I think that if we could speak of the survival being in the interest of the mindless animal, i.e., that we could ascribe to the mindless a good or even well-being or a sense in which it could flourish without thought, that when it became conscious it could become interested in that interest or good or well-being. But if a Parfit-inspired animalist like Olson is correct, the human animal would never become genuinely interested in it survival (or well-being or good) but really only cares about its psychology continuing. I find it quite bizarre to say that we really only care about our psychology continuing and not that we want to be the subject of that psychology. I can't accept that it is just as good if my psychology is continued by individual

X or individual Y or individual Z as it being continued with me as its subject.[20] So I am hoping that there will be some readers who will find it odd from a biological perspective (in some broad sense) that the mindless animal with survival in its interest never actually comes to be nonderivatively interested in that survival or well-being. It strikes me as making greater biological sense for an animal to have interests and a good when it was mindless that then become explicitly aimed at when the animal develops cognitively. I think it would be natural for the animal to come to desire that interest (survival), just as Regan's kid could come to be interested in its own interest (health).

A similar charge of bad or peculiar biology can be leveled against Olson-style animalism in terms of proper function which will reinforce the above argument, or replace it if it is an error to ascribe interests to mindless animals. On Boorse's (1977) account of health, organ systems are functioning properly when they make their contribution to the organism's survival.[21] But if we read the Parfit-Olson claim about identity not mattering in a normative fashion, then when the animal's cognitive system develops, it is functioning properly when it serves not the animal's survival but that of its psychology, whoever may be its subject. The animal would be malfunctioning if it cared about its own survival in the transplant scenario. This is thus evidence of a rather peculiar biological disconnect between the proper functioning of an animal's cognition and the rest of its organ systems.

Even with a better account of proper function (that doesn't involve crude malfunctioning when saving a stranger's life), the concern for one's own life is still nonderivative, just outweighed, and thus Parfit-Olson account of derivative self concerns is anomalous amongst organ systems. It would be a malfunction on Parfit-Olson's account to prefer A) saving someone's life and surviving with a slight loss of psychological continuity due to an injury that one received during the rescue to B) saving someone's life, though dying in the process and being replaced by a perfect psychological duplicate

IV. An Animalist Alternative: Prudential Concern without Psychological Continuity

So if psychological approaches of identity give us problems stemming from too many thinkers, pronoun revisionism means a lack of animal autonomy, while Olson-style animalism is incompatible with our practical self-understanding, what are the alternatives?[22] What I now want to suggest is that in a brain transplant scenario, where my cerebrum will be destroyed and that of my identical twin placed in the body where my cerebrum had been, I not only would survive as the animal but should *now* care about my future self even though my psychology will then become just like that of my twin.

To soften up the audience, let me remind them, or some of them, that they did or will care that their own mindless embryonic child have a certain future for its own sake. It isn't that such expectant parents only start to care later about a new being (a person) that emerges or care now for a being yet to come into existence. They want their mindless unborn child to become happy and to flourish for its own sake and believe that a brain will obviously be beneficial to the child. Their attitudes seem to be presupposing that identity is a component in what matters to the child since they care about that child for its own sake, even before it is psychologically continuous with anyone. If we can care for the mindless embryo or fetus before it has a psychology on the basis that the later psychology will be good for it, then perhaps adult human animals should have prudential concern for their own future even in the absence of psychological continuity of any sort as in the earlier thought experiment of our acquiring a new upper brain.

I want to try a different strategy to strengthen this position that we should be prudentially concerned about our adult conscious animal even when it is without any psychological connection to the present or even the same brain playing a role subserving our future mental life. Consider that we care prudentially for the stroke victim that we would become even if our mental

capacities are reduced to mere sentience. Many philosophers (Johnson, Lockwood, McMahan, Unger, Salmon etc.) believe this shows that it is mere consciousness or sentience, not self-consciousness that is essential to our persistence.[23] I think instead that such thought experiments can actually be construed as showing that animal identity matters rather than is of just derivative concern. Is our concern for our post-injury self with just a rudimentary mind really due to our possessing the *same organ* that underlies consciousness or is it rather that it is just the same animal that is conscious? Ask yourself if 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? If you would have prudential concern for the same animal with a different physical realization for sentience, why should you react differently to your animal getting an entirely new cerebrum in the earlier thought experiment?

Someone might respond that it matters that the different anatomical structures involved in the production of 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 desire arises, different parts of a developing brain give rise to the pain and pleasure "sectors". Imagine one is in the brainstem as Shewmon has shown to be possible and the other is in the cerebrum (1997: 57-60). Would it be correct to say there were two thinking beings in one body? And if one is destroyed, is it correct to claim there is then only one thinking being? And would we say that we have a new thinking being 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"? I very much doubt it. I don't see any reason to identify ourselves with parts of the sentience-producing brain a la McMahan, nor with a larger being only if it

continuously possesses the same functioning brain-like structure a la (the pre-dualist) Unger, 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. But if you agree with me about this, then why maintain that we would have no prudential reason to care about one's animal if it received a new brain in the thought experiment?[24]

Perhaps you will argue that you would survive with *any* parts of your brain producing conscious states but not with a new brain. I would then ask for a compositional principle that makes the cerebrum, midbrain, cerebellum, brainstem etc. all parts of the same thing?[25] As far as I can see, what they have in common is that they are caught up in the same biological life. But so would be the assimilated brain in the transplant thought experiment. If you lack my compositional doubts, imagine that now and after a debilitating stroke that your pain is realized by the upper spine while pleasure has a cerebral basis.

In conclusion, if this alternative animalist approach provides a plausible conception of prudential concern, then we can avoid the metaphysical and practical problems of approaches that construe psychology as essential to us without abandoning the claim that identity matters. However, I must admit that it appears to be a rather thin limb to climb out on and so expect few readers to follow me. But my question then is where are the sturdier branches?[26]

Bibliography

Baker, Lynne Rudder. 2000. *Persons and Bodies: A Constitution Approach*. (Cambridge University Press).

Beauchamp, Tom and Faden, Ruth, with Nancy King. 1986. *A History and Theory of Informed Consent.* (Oxford University Press.)

Beauchamp, Tom and Childress, James. 2009. Principles of Biomedical Ethics. 6th ed. (Oxford:

Oxford University Press.

Bernat, James. (1998) “A Defense of Whole-Brain Concept of Death.” Hastings Center Report 28:2 (14-23)

Boorse, Chris. 1977. “Health as a Theoretical Concept.” *Philosophy of Science.* 44: 542-73.

Brock, Dan and Buchanan, Allen. 1990. *Deciding for Others: The Ethics of Surrogate Decision Making*. (Cambridge University Press.)

Damasio, Antonio. *Descartes’ Error: Emotion, Reason and the Human Brain* New York: Penguin Books, 2005.

DeGrazia, David. 2005. *Human Identity and Bioethics.* (Cambridge University Press)

Dresser, Rebecca. Dworkin on Dementia: Elegant Theory, Questionable Policy. *Hastings Center Report* 25:6, 32-38.

Ferlik, Andrew (1991) “Margo’s Logo” *Journal of the American Medical Association.* 265. 201.

Hawley, Katherine. 2005. “Fission, Fusion and Intrinsic Facts”, *Philosophy and Phenomenological Research*, 71:3, 602-621.

Hershenov, David. 2005. “Do Dead Bodies Pose a Problem for the Biological Account of Identity?” *Mind*, 114:453, 31-59.

Hershenov, David. 2008. “A Hylomorphic Account of Personal Identity Thought Experiments” *American Catholic Philosophical Quarterly*. 82:3. 2008.

Hudson, Hud. 2001. *A Materialist Metaphysics of the Human Person.* Cornell University Press.

King, Nancy. 2007. *Making Sense of Advance Directives.* (Washington DC: Georgetown University Press).

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

McMahan, Jeff. 2006. "An Alternative to Brain Death," *Journal of Law, Medicine, and Ethics* 44-48.

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

Noonan, Harold. 2003. *Personal Identity*. 2nd edition. (Routledge Press)

Noonan, Harold. "Persons, Animals and Human Beings" in J. Campbell amd M.O' Rourke, eds. *Time and Identity*. MIT Press. Forthcoming

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

Olson, Eric. 2002. 'Thinking Animals and the Reference of 'I." *Philosophical Topics*. 30, 1, 189-208.

Olson, Eric. 2004. "Animalism and the Corpse Problem." *Australasian Journal of Philosophy*. 82: 265-74.

Olson, Eric. 2007. *What Are We?* (Oxford University Press.)

Perry, John. 1979. "The Problem of the Essential Indexical." *Nous* 13, 3-21.

Persson, Ingmar. 1999. "Our Identity and Separability of Persons and Organisms." *Dialogue.* 38. 521-533

Regan. Tom. 1982. "What Sort of Beings Have Rights." *All That Dwell Therein: Animal Rights and Environmental Ethics.* (Berkeley, University of California Press).

Rosenkranz, Gary and Hoffman, Joshua. 1997. *Substance: Its Nature and Existence* (London: Routledge)

Shewmon, D. Alan. 1997. "Recovery from 'Brain Death': A Neurologist's Apologia" *Linacre Quarterly*. February, 1997, 30-96.

Shoemaker, Sydney. 1999a. "Eric Olson: The Human Animal." *Nous*. 33:3 496-504

Shoemaker, Sydney. 1999b "Self, Body and Coincidence." *Aristotelian Society Supplement.*.

Shoemaker, Sydney. 2004. Functionalism and Personal Identity – A Reply. *Nous.* 38:3 525-533

Sider, Ted 2001. *Four-Dimensionalism: An Ontology of Time and Persistence.* (Oxford University Press)

Tye, Michael. 2006. *Consciousness and Persons*. (MIT Press)

Unger, Peter. 1990. *Identity, Consciousness and Value*. (Oxford University Press)

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

Veatch, Robert. 1998. "The Conscience Clause: How Much Individual Choice in Defining Death Can Our Society Tolerate" in the *Definition of Death: Contemporary Controversies*, ed. Stuart Younger, Robert Arnold and Renie Shapiro (Baltimore: John Hopkins University Press) 137-60

Wear, Stephen. 1998. *Informed Consent: Patient Autonomy and Clinician Beneficence Within Health Care.* (Washington DC: Georgetown University Press)

White, R.J and Albin M.S., 1964. "Preservation of viability in the isolate monkey brain utilizing a mechanical extracorporeal circulation." *Nature*. 202: 1082-1083

White, R.J and Albin M.S. 1971. "Cephalic exchange transplantation in the Monkey." *Surgery* 70:1. 135-139.

White, R.J and Albin M.S. 1963. "Isolation of the Monkey Brain: In Vitro Preparation and Maintenance." *Science* 141. 1060-1061.

Wiggins, David. 2002. *Sameness and Substance Renewed*. (Cambridge University Press).

[1] Animalism is the view that we are essentially living animals. We persist as long as life processes continue, or at least the capacity for them is instantiated. According to the animalist, the exercise of our capacity for thought may make life valuable, but it has nothing to do with our origins, persistence or endings. The opponents of animalism usually defend some version of what has been called the psychological approach to personal identity. They stress that some kind of mental life is essential to our persistence.

[2] See author's article …citation withheld for purpose of blind review

[3] Pierre Cabanis, an 18th century scholar, notoriously made this claim about the brain secreting thought like the liver secretes bile. For some reason, the 19th century materialist Karl Vogt believed the point needed to be made even more emphatically so he declared: "*The brain secretes thought as the stomach secretes gastric juice, the liver bile, and the kidneys urine*." Countless neuroscientists make claims like the mind is the brain, or the mind is the functioning brain, or the mind arises from the functioning brain. It is a small step from this to identifying ourselves as whatever part of the organism strictly thinks our thought.

[4] Perhaps McMahan should take a page from the animalist treatment of the body and insist that the brain (or relevant part) ceases to exist when it loses its ability to sustain cognitive processes, just as the animal goes out of existence when it loses its ability to instantiate life processes. Animalists such as Olson and Hershenov don't believe that animals survive death as corpses. See Olson (1997, 2004) Hershenov (2005). This strategy would enable McMahan to avoid the spatially coincident person and brain part. But this identification strategy only works if McMahan can deny that persons can survive the replacement of their organic brain with a functionally equivalent inorganic brain. Other defenders of the psychological approach to personal identity, such as Baker and Unger, believe becoming organic is indeed metaphysically possible. They base this belief on the reason why it seems that ordinarily organic persons survive the replacement of their matter with other organic parts: such changes occur without any disruption of cognitive function, i.e., their thought is left unaffected. So the question for McMahan is why doesn't the same reasoning behind total organic part replacement allow the person to become wholly inorganic? And if the person can become wholly inorganic without disturbing thought processes, then the person can't be identified with the original organic brain which would no longer exist. Thus the unwelcome problem of spatially coincident thinkers returns.

[5] It may be that Baker believes the lack of a first person perspective and self reference is built into the nature of the human animal. So she doesn't need a conventional answer a la Noonan. But other advocates of spatially coincident entities do.

[6] Locke famously wrote: "We must consider what person stands for: which, I think, is a thinking intelligent being, that has reason and reflection, and consider *it self as it self*, the same thinking thing in different times and places."

[7] I think pronoun revisionism actually fails to avoid the four problems it is supposed to avoid because it is merely a linguistic or conventional solution. Let's say that the first person pronoun refers to not the individual who thinks or utters it but the one with psychological persistence conditions. (I don't remember being taught that as a kid!) Why

couldn't we then make up a word like 'Shmy' which would serve as the first person pronoun for the animal? So it too would be able to refer directly and without the possibility of misidentification to itself as the thinker of its thoughts. That is, it too would have an essential indexical of sorts (Perry 1979). But if that was possible, then the duplication and trivialization problems would return because then the animal would be person-like in its ability to think of itself as the subject of its thoughts. But even worse, the epistemic problem might return which would actually threaten the autonomy of the person. The problem is that when one is learning English words like 'I' (or a new one like 'Shmy') one has to grasp in thought (a Fodor-like language of thought?) the character and content before one can understand the word as expressing that. So there must be a representation of the meaning that is not in English when one learns such English words. But if the animal and person can then both think of themselves as the subject of their thoughts, the epistemic problem will return. They will know, thanks to pronoun revisionism, that the referent of the English word 'I' is the person and the referent of 'shmy' is the animal. But they can still wonder whether they are the animal or the person in their language of thought. If they can think about themselves in the way that is required to represent to themselves the meaning of English pronouns, then they can wonder which being they are.

[8] See King (2007) for the various forms advance directives can take.

[9] Recall that in the absence of pronoun revision there will be an epistemic problem. The human animal and the human person won't know whether they are animals or persons. They will either both think they are the animal, and one will be wrong, or they will both think they are the person, and the other will be wrong. So either the thinking animal or the thinking person will always be deciding upon an advanced directive while thinking it was the other.

[10] I actually think that momentary stages, despite their causal ties to other stages, are too short-lived to think. And if worms and other non-momentary entities think only in virtue of their momentary stages thinking (otherwise there would be the problem of temporary intrinsics), then since the stages don't think, neither do the other perduring entities composed of them.

[11] This is not just the stuff of science fiction and metaphysical thought experiments. White and Albin (1963, 1964, 1971) conducted brain transplant experiments with monkeys because they believed a similar procedure would be the best way to enable the paralyzed human being to move again.

[12] This replacement of our organic bodily parts with inorganic parts is the dream of the Transhumanist Movement. See the articles in Journal of Evolution and Technology Becoming More Than Human: Technology and the Post-Human Condition Special Issue (Volume 19 Issue 1).

[13] In a much referred to article, Dr. Andrew Ferlik famously described Margo, an Alzheimer's patient, as "undeniably on of the happiest people I have even known"(1991:201). Rebecca Dresser points out that most individuals "with dementia do not exhibit the distress and misery we competent people tend to associate with the condition….thus the 'tragedy' and 'horror' of dementia is partially attributable to the ways in which others respond to people with this condition" (1995).

[14] Readers may be wondering how we could know the interests of a human animal if the animal couldn't ever autonomously reflect upon its interests due to pronoun revision or the mistaken belief that it was the associated person? One possibility is that we could just speculate that given human nature, the human animal would want such and such. Perhaps the following alternative would be more helpful. The person (and the animal) could think "If I (the person) were the animal, I would like to be treated in such and such a way." If any creature can know the animals interests, one would think it was the overlapping person. It is not as if their psychologies could dramatically diverge and say one like dance and classical music and the other not.

[15] McMahan claims the mindless animal is to be viewed as morally akin to a corpse. He insists that while it must be treated with respect because a person once animated it, it doesn't have rights and can't be harmed. (2006, 48)

[16] I am assuming that they share the same interests though the animal can't express them as its own. Perhaps this can be determined by a substituted judgment of the type mentioned in footnote 14.

[17] For example, New York and New Jersey would allow an orthodox Jew to reject the brain death criterion of death. They would not be declared dead until the circulatory/respiratory criterion had been met (Veatch: 1998, 137-60).

[18] Baker says to imagine A and the psychologically continuous offshoots, B and C that replace her. "Suppose A had been on the verge of apologizing to her former fiancée for breaking off their engagement. Should B and C who now both have A's memories, both apologize to the former fiancé for the broken engagement? …Or suppose that A had promised D to buy her lunch. Are B and C both obligated to fulfill that promise? How could one be obligated to fulfill A's promise but not the other? Does A's promise entitle D now to two lunches?...Suppose that A was a politician who wanted to become the first woman Democratic presidential candidate. B and C, each of whom

remembers A's vow, are both infuriated by the unexpected (and unfair?) competition. Suppose that B becomes the first woman Democratic presidential candidate. B says elatedly…"I've totally fulfilled the intention that I remember before the operation." C says "dejectedly "since I am not the first Democratic woman presidential candidate, the intention I remember before the operation is totally unfulfillable." How can a single intention both be totally fulfilled and totally unfillable?" (2000:129).

[19] It might be thought that the animalist runs afoul of the only x and y principle when confronted by twinning. Dean Zimmerman believes that all three-dimensional materialists will fall prey to this. I am not so sure. A single cell organism (like an amoeba) would go out of existence when it divides regardless of whether one or both descendents survive. An organism, I take it, dies when that event that it is life ceases (or the capacities for life are no longer instantiated). And that seems to occur with cell division. I don't know that much about the cellular details but when the nucleus is being copied during the mitosis, it stops sending out a flow of chemically encoded directions to the rest of the cell and the event that is its life halts and that seems to be a good time to say that is the demise of the cell and the end of its life. It seems that the integration and cooperation characteristic of life ends there regardless of whether one or two cells emerge. Olson actually claims that is indeed the *demise* of the cell (1997: 114). Moreover, on Olson, van Inwagen, DeGrazia, Smith and Brogaard's animalist approach, no animal between 2 and 16 days post-fertilization (i.e., the period when twinning is still possible) goes out of existence when fissioning. The basis for this is their claim that *there is no such existing animal*, merely a bunch of single cell organisms that don't compose anything that is a mutli-cellular animal, its parts cooperating in the integrated way that a multi-cell animal does, maintaining homeostasis, metabolizing etc. as a unit. So the most common case of twinning doesn't provide an example of something going out of existence that would have survived if there had been only one successor. (Perhaps a sci fi case can be constructed of an adult animal fission which would be more difficult for me to deal with, but I would have to see the biological details.)

[20] It might help to imagine being told that a person psychologically continuous with you will suffer greatly in the future. This is bad but surely not as bad, or not in the same way, as being told that it is *you* who will suffer in the future.

[21] I have chosen to make this argument by appealing to Boorse's (1977) account of health and proper function. But I believe the argument could be carried out with rival accounts of proper function.

[22] One possibility, not pursued in the main body of this paper, is identifying the animal and the person, but claiming it has psychological persistence conditions. Such a being thus could be reduced to cerebrum-size and transplanted though it would cease to be alive during the process and hence is only contingently an animal. The claim that the person is the animal rather than spatially coincident or embedded within it avoids the extra thinkers and preserves the intuitions many have in transplant and irreversible coma cases. This view has a few modern adherents but Hershenov (2008) argues that it is actually the best way to construe the Catholic hylomorphic tradition. Catholic claims that we are made in God's image in virtue of our mentality rather than animality, and are bodiless when morally reformed in Purgatory, suggest we are only contingently living beings. And Aquinas's own position that every deceased human being is resurrected for eternal rewards in a body akin to that of a 32-33 year old which doesn't eat, drink, sleep and thus metabolize is at odds with the claim that we are essentially living organisms that persist by our parts being caught up in the same life. The substantial change that Aquinas envisions with the acquisition of rationality (delayed hominization) and then its loss (departed hominization) is akin to the substantial change that Olson has shown would occur in the transplant scenarios where animals are ascribed psychological persistence conditions. Olson suggests the idea is bad biology for it would lead to organisms popping in and out of existence - though his own alternative account has cerebrum-size persons popping in and out of existence in the transplant scenarios (Olson, 1997: 120-21). The allegedly bad biology infecting hylomorphism is that when an animal X with psychological persistence conditions has its cerebrum removed, found on the table where the operation took place is not the original animal X in a mindless state but a new cerebrumless animal Y that just came into existence. And, given the assumption that there is no spatial coincidence, when X, who has become cerebrum-size, is later placed in the mindless body of a third individual Z, that animal Z, pops out of existence rather than acquires the capacity for thought.

[23] Although DeGrazia is an animalist and doesn't believe sentience is essential to us, he suggests that the best interpretation of the transplant thought experiments is that sentience rather than self-consciousness is required for our persistence (2005:24).

[24] And if you admit that you have some prudential concern for the your animal with a new brain, then you can't also claim to have prudential concern for the being who earlier received your brain in a transplant.

[25] Incidentally, the animalist avoids the problem of the thinking animal having a thinking brain embedded within it by opting for a sparse ontology. Animalists claim that only organisms and simples exist (Olson and van Inwagen) – or only organisms, simples and some properly conjoined masses (Rosenkrantz and Hoffman 1997). They deny that brains exist, instead they claim that there are only atoms arranged brainwise (Olson 2007). Advocates of constitution and four-dimensionalism can't help themselves as easily to this sparse ontology solution. Four-dimensional theorists claims that a major appeal of their theory is that its endorsement of unrestricted composition avoids puzzles of vague existence that plague any theory that limits the composite objects to certain kinds (Sider, 2001:120-139). And constitution is prized for its ability to handle the paradoxes of of increase and decrease by claiming objects become constituted by earlier embedding or embedded objects (Olson, 2007). However, there is still one scenario of conjoined twins where the animalist who insists that there are no brains but only atoms arranged brainwise has a too many thinkers problem due to two animals sharing the same atoms arranged brainwise but no other parts. However, the animalist's having to tolerate the extra thinker turns out to be less objectionable than the constitution theorist's treatment of the problem. See my paper (withheld for purpose of blind reviewing) for an account of why the animalist has an advantage over the constitution theorist in dealing with overlapping thinkers. For a more sympathetic reading of the constitution theorist's treatment of the issue, see my (withheld for purpose of blind reviewing.)

[26] Acknowledgments withheld for purpose of blind reviewing.