

Four-Dimensional Animalism

David B. Hershenov

I. Introduction

The Four-Dimensionalist recognizes the existence of entities, temporal parts, that the Three-Dimensionalist does not.¹ Your animal will have a temporal part that exists for the first half of its life, another for the first quarter, and even one for the first moment of its life. And your animal will consist of an infinite number of other temporal parts: some composed of only minimally conscious temporal parts, others containing just robustly self-conscious temporal parts, and still others including both thinking and non-thinking temporal parts. Also, unlike many Three-Dimensionalists, the Four-Dimensionalist typically adopts unrestricted composition and thus holds that any collection of objects will compose another. That means there will even be an object consisting of the reader, some sand on the Jersey coast, and the Hanging Gardens of Ancient Babylonia.

The puzzle confronting the Four-Dimensionalist is which of these countless entities that have some thinking parts should be designated a *person*? Since Hud Hudson has thought longer and harder about this topic than anyone else with whom I am familiar, I will concentrate on his solution. He presents an exclusion principle that denies personhood to objects possessing any parts that don't contribute to thought. Thus the human animal is not the human person because it begins life with mindless embryonic temporal parts. I will contest this claim of non-identity, but without rejecting anything else in Hudson's conception of Four-Dimensionalism.² So I will accept, though just for the sake of argument, that we persist in virtue of temporal parts, that

¹ Informally, a temporal part of an entity will exist only at a time and will then overlap all of the entity's other parts that exist at that time. More formally: something is a temporal part of *x* during interval *T* if and only if (i) the object exists at but only at times in *T*, (ii) it's part of *x* at every time during *T*, and (iii) at every moment during *T* it overlaps everything that's part of *x* at that moment [6, p. 59].

² Hudson personally favors a version of Four-Dimensionalism that he has named *The Partist View* which replaces temporal parts with an analogue he calls spatio-temporal parts [7, p. 65]. I will ignore his favored version of Four-Dimensionalism for ease of presentation. (He actually does the same for most of his book.) Nothing will be lost, for everything said here about standard Four-Dimensionalism can be translated into the vocabulary of the Partist.

composition is unrestricted, that epistemicism is the appropriate treatment of vagueness, that human persons are complex entities composed of only material parts, and that “person” is maximal and thus there are no persons embedded within other persons.

The first part of the paper will be a response to Hudson’s claim that human persons can’t be identified with animals who, though mindless at their origins, have the potential for later thought.³ Hudson argues that if having *later* thinking states were sufficient for being a person, then, assuming unrestricted composition and temporal parts, there would be an infinite number of entities that are persons. I will instead argue that one can identify human persons and human animals on the basis of a gen identity relation that will not extend the title of *person* to other entities that earlier have mindless temporal parts preceding their thinking ones. The key is to appeal to a type of causal relationship unifying temporal parts that is typical of entities that belong to a natural kind.

This will still leave us with two good candidates for the title ‘human person’: the human animal that is initially mindless and then later self-conscious, and the entity favored by Hudson that is capable of self-conscious psychologically continuous reflections at every stage of its existence.⁴ With which of the two candidates for personhood are we to be identified? Hudson argues that just as it is arbitrary and unmotivated to consider an object with unthinking (embryonic) stages to be a person when there is a candidate available with only thinking temporal parts, so too is it unprincipled to view the person as including bodily components of the animal like hair and fingernail that don’t contribute to the production of thought. Since Hudson maintains that the parts relevant to production of thought and personhood are found “beneath the

³ See Hershenov [5] and Hudson [6] for two competing conceptions of the bioethical implications of perduring human beings.

⁴ I’m using stages to designate very brief, perhaps momentary, temporal parts.

skin” in the brain and central nervous system, no animal really has the potential to someday be thinking in the strictest sense.⁵ So if a human animal will never come to think in the strictest sense, then we shouldn’t consider that animal to be a person in virtue of any potentiality for thought. However, drawing upon ideas of Damasio [3], Noë [11] and Olson [12], I will argue that the thesis that only neurological components beneath the skin of the animal contribute to thought and thus compose the person cannot be sustained. I’ll first conjecture why Hudson may have been wrongly led to reduce the size of the person. Even if incorrect as a matter of intellectual biography, my diagnosis will reveal a problem in trying to exclude certain parts of the animal from the person on the grounds that they don’t contribute to thought.

Regardless of whether persons are found wholly beneath the skin or not, there will still be a candidate for personhood consisting of just thinking temporal parts overlapping the older perduring animal, the latter including mindless embryonic temporal parts. So it might seem that an exclusion principle will favor the entity with only thinking parts over the entity that possesses parts superfluous to thought. However, in the second part of the paper, I will argue that our intuitions about the persistence of persons are best explained by appeal to a biological (or animalist) account of personal identity. Our intuitions that we would survive certain hypothetical changes as indicated by what appears to be prudential concern for the resulting individual can’t be accounted for by any criterion of psychological connections and continuity, or even the continuation of the brain’s capacity for mere sentience. My contention is that only an appeal to a criterion that identifies us with a future person in virtue of sharing the same biological life can make sense of such responses. Since we take ownership of the thinking temporal parts in virtue of their being biologically related to each other, it would thus be arbitrary and unmotivated to

⁵ Perhaps thought could be ascribed to the animal derivatively since it has a part that non-derivatively thinks. The analogy would be to stating the car is noisy in virtue of it having a truly noise-making part, the horn [10, p. 93].

claim the person consists of just the thinking temporal parts of the animal since these are not unified by immanent *mental* causation. If this is correct, then even if it is wrong to deny that our thought is produced by just a temporal part of the animal's central nervous system, I can still claim that such temporal parts are not linked to each other across time by their realizing psychological continuity.⁶ Instead, they are linked by realizing mental states of the same living animal, i.e., they are caught up in the same life processes. Moreover, that same biological relationship linking the various thinking temporal parts also unifies the unthinking parts of the animal with the thinking parts. This provides us with grounds to argue that of the countless Four-Dimensional entities possessing some thinking temporal parts, we are to be identified with the living human animal. Furthermore, if any entities warrant the label of 'person', we do, for surely any theory that posits we are not persons is wrong. Then by helping ourselves to Hudson's maximality principle – there are no persons embedded within another person - we can consider any entity with *only* self-conscious temporal parts to be but a proper part of the human person.⁷

II. Why Four-Dimensional Human Animals Don't Appear to be Persons

Hudson believes Three-Dimensionalism to be false. He bases this not on an appeal to temporary intrinsics, compatibility with relativistic physics, or considerations about vagueness, but to Four-Dimensionalism's superior handling of well-known problems of material constitution such as fission, embedded parts, the statue and the clay [7, p. 58]. Moreover, once one accepts unrestricted composition, Four-Dimensionalism has an appeal that Three-Dimensionalism lacks [13, pp. 229-31]. A Three-Dimensionalist is not at liberty to accept unrestricted composition without some very counterintuitive results. If one's notion of unrestricted composition includes

⁶ Thus in the second half of this paper, the claim that not all of the animal's parts contribute to the person's thought can be added to the list of Hudson's assumptions that are accepted for the sake of argument

⁷ This denial of two kinds of *human* persons isn't meant to rule out there being other kinds of persons such as the robotic, angelic and divine.

the idea any plurality of things composes one *and only one* other entity, i.e., there are no spatially coincident entities compose of the same parts, then there is an advantage to advocating a Four-Dimensionalist approach. The reason is that the parts that compose you at any moment will soon be somewhat dispersed as you exhale, perspire etc. Eventually your composite atoms will be scattered across the region. If you want to avoid having to picture yourself as possessing human shape for but a fraction of your existence, then it is better to understand yourself as composed of temporal parts that are themselves fusions of temporal parts of atoms.

Hudson looks unfavorably on attempts to reject unrestricted composition. He offers a defense that includes both analyzing the intuitions behind those who reject it, and showing the costs of doing so. He conjectures that those who deny unrestricted composition have mistakenly allowed their interests to determine their ontology [7, p. 107]. While there is no name for the various objects consisting of a grain of sand in the Sahara Desert and a drop of water in the Indian Ocean because there is no human interest in them, Hudson insists that we aren't justified in restricting our account of what exists to what we are interested in. Not only do the interests of human beings change, but there could be non-humans with very different interests. Moreover, there is no principled way to allow some scattered objects and not others [7, p. 108]. Not only are the United States and its various institutions scattered objects, but its citizens are swarms of scattered microscopic particles. So once one admits Congress and Hawaii into one's ontology, it is very hard to leave out the scattered object consisting of the drop of water and the grain of sand.

Let's now examine Hudson's reasons to claim that if composition is unrestricted, then the person can't be identified with the animal. He insists that the animal is not a person for, at best, human animals and persons would merely share *some* of their temporal parts. The typical human animal consists of thinking stages through most of its life and non-thinking stages during its embryonic months. There is also a distinct perduring creature, perhaps a large temporal part of

the animal, whose temporal parts are all capable of thought. Which one is the person? Hudson finds it more compelling to identify the person with the entity consisting of *only* thinking stages rather than something like the animal which also has non-thinking temporal parts. However, such a principle, if left unqualified, would mean that your temporal part that exists for the duration that you are reading this sentence would be a person embedded within you. Since there are countless things that consist of only thinking temporal parts in a Four-Dimensional metaphysics, Hudson contends that the only non-arbitrary selection of stages deserving the label “person” is an aggregate of thinking stages which are *not* embedded within a larger, similar thinking being. Thus “person” is a *maximal* concept.

Hudson further specifies that it isn't any kind of thinking that belongs to the maximal person. For example, *merely* sentient stages aren't sufficient for personhood. What is needed are self-conscious thoughts appropriately related via psychological continuity and connectedness [7, pp. 122, 130-31, 144]. Even that is not enough since a nuclear explosion could have vaporized Hannah while she was reflecting upon her thoughts, and then in an incredible cosmic coincidence, a psychological duplicate of her could materialize in a far off galaxy [7, p. 132]. Since Hannah has died rather than suddenly switched locations, the existence of psychologically continuous stages isn't enough to make them belong to the same person. There must also be the right kind of causal connection, an *immanent cause* involving earlier thinking stages bringing about later thinking stages [7, pp. 134-36].

Hudson contends that appealing to the animal's *potential* to have later temporal parts manifesting thought will not succeed in rendering the animal identical to the person. He acknowledges that might work for a Three-Dimensionalist metaphysics that restricts composition and denies that there exist spatially coincident entities [7, pp. 125-26, 152-53]. On such an account of persistence, it is better to maintain that the mindless embryo is identical to the later

minimally sentient newborn, self-conscious adult, senile geriatric and terminally ill, irreversibly comatose patient. But the Four-Dimensionalist typically accepts unrestricted composition and so there will be countless objects that have thinking stages. There will even be an object that consists of an ancient Egyptian necklace and President Roosevelt. If mindless embryos are persons because they have later thinking parts, then there was an additional person present in ancient Egypt in virtue of the necklace that was an early temporal part of an object composed of it and the later thinking Roosevelt stages. But surely that object isn't a person endowed with the moral status that is typically thought to accompany personhood. When Roosevelt died, the world didn't lose more than one person. If everything that has thinking stages for a time is a person then Roosevelt's death would have involved an infinite number of deaths.

III. The Components of a Person

A. Natural Development

So the problem Hudson presents us with is that if we want to deny a person exists in ancient Egypt due to the necklace existing there and being part of an object that includes the later Roosevelt thinking stages, then we must also deny that the mindless human embryo is a person in virtue of its future thinking temporal parts. My response will involve arguing that there are grounds for claiming that some but not all potential thinkers are themselves persons even when they haven't yet manifested that potential. There is a way to distinguish potential thinkers via the relations unifying the stages of a natural kind in order to then claim that the animal is also a person while other beings that have thinking stages for just some of their existence are not persons. The idea is roughly that the mindless embryonic stages are the same kind of stages of the latter thinking person – i.e., they are all animal stages. There are mindless animal stages linked by life processes to thinking animal stages. They are all living stages of an animal. Their diachronic (as well as synchronic) unity is due to their parts being caught up in the same life

processes. They are stages of the same token of a natural kind, not parts of two things of distinct kinds cobbled together in virtue of the principle of unrestricted composition.⁸ The gerrymandered entity composed of the ancient necklace and President Roosevelt doesn't have later *necklace* stages that happen to think. This suggests an explanation of why it seems much less plausible to ascribe the capacity of thought to the necklace-Roosevelt entity when only the necklace is present than it is to so ascribe it to the mindless stages of Roosevelt. The capacity is not found in the developmental telos of the necklace. It is not the nature of the earlier stages to give rise to later thinking stages. Compare the necklace/Roosevelt composite with Roosevelt himself. One finds a telos programmed into all the stages of Roosevelt, even the mindless ones.

So the idea is that there is a principled distinction between things that have thinking parts at one time in their existence but not at another. The mindless animal stages that are part of an entity that later thinks are stages of one and the same animal. The later thinking stages are also animal stages united by life processes. But the mindless necklace stages are not part of an entity that later thinks composed then of necklace stages. I suspect only the human animal will have its *mindless* and *thinking* stages bound by the same unity (gen identity) relation. And the reason there is no *animal* composed of you up to this moment and another reader after this moment is that there is not the appropriate *immanent causation* characteristic of life processes: the earlier stages of a life causing the successive stages of the same life. Likewise for the composite of the scattered gametes and the reader that resulted from their fusion. There are *three* lives involved. The *same* life doesn't link them diachronically or synchronically.

So we can grant that mindless human animals are persons without having to bestow the title on every object which has mindless stages preceding its thinking ones. However, there *may*

⁸ Hudson admits that appealing to natural kinds is the best option for his rivals who want to identify persons and animals [8, p. 233]. But he thinks the notion of natural kind is "too obscure" to be effective.

also be a single relation, psychological continuity, unifying all of the thinking stages of the person. So Hudson could appeal to the existence of a non-gerrymandered, ‘natural’, psychological unity relation in to order apply the label ‘person’ to the perduring object consisting of only thinking stages. Thus my approach might seem to commit us to there being two kinds of persons – some that are mindless for a time, others that are always thinking. And that admission will run afoul of the maximality principle that persons are not to be found within larger persons. However, I shall put forth arguments in part IV that reveal the best candidate for the label ‘person’ to be the one that was once a very little mindless animal. So it will not be, as Hudson claims, “arbitrary and unmotivated” to identify the human person and human animal.

B. Contribution Determines Composition

I have so far ignored another Hudson-inspired argument regarding why we should not consider the mindless embryonic animal to be a person on the grounds that it has potential to develop into a thinker. Hudson, makes the surprising claim that perduring persons are not temporal parts of animals. Rather, persons are “certain proper temporal parts of the brain and central nervous system of living human organisms” [7, p. 147]. The basis for this claim is that the person is composed only of those parts that contribute to its cognition. The entire animal doesn’t produce thought, merely part of it does. So if the mindless animal is never going to develop to where it can *directly* produce thought, there is little reason to identify the person with the animal who will, at best, come to think only derivatively in virtue of some of its parts *really* doing the thinking.

Hudson insists that just as it is unprincipled to identify the person with a perduring animal that possesses non-thinking temporal parts, so too is it to identify the person with any of the temporal parts of the animal since many of the animal’s spatial parts have nothing to do with thought production. Hudson explains: “Rather, once again, the only non-arbitrary choice would

be an object each of whose parts plays *a contributory role* in supporting a psychological profile constitutive of personhood” [8, p. 224]. Even though Hudson can’t say exactly which parts are so involved, nevertheless, he claims that since he can rule out “such parts as one’s forearm [8, p. 219]...some parts of the hand...” [8, pp. 224-25], “finger nails and bone-marrow...” [8, 143-44] as making a contribution to thought, that is enough to sustain his thesis that persons are not temporal parts of animals. This leads Hudson to conclude that the person is not only to be found “within the lifespan” but also “beneath the skin... of the human animal.” [8, p. 220].

I am skeptical of the view that only certain parts of the animal contribute to the production of thought and want to offer a rather speculative diagnosis of what might be the source of Hudson’s error. Even if I am wrong about the source of his error, the diagnosis will still reveal that ‘a contributory role in supporting a psychological profile’ won’t restrict the person’s boundaries in the manner Hudson envisions. My suspicion is that Hudson is misled by the truth that the animal could continue to think if reduced in size to 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 Hudson thinks produce thought. The animal needs to be alive to think. Following van Inwagen [18, pp. 81-87], 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.⁹ 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. 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.

The basis for denying that we thinkers are merely parts of animals doesn't just lie in the fact that the living body contributes the life support necessary for any brain activity. The neuroscientist Antonio Damasio argues that the brain's constant monitoring of the body, its receiving and sending of the messages, is necessary for the working of the normal mind [3, pp. 223-44]. Even *partially* cutting off inputs to the brain in those suffering spinal chord injuries causes changes in the state of mind. Damasio's acceptance of "the idea that the mind derives from the entire organism as an ensemble" [3. p. 225] leads him to reject the assumptions underlying one of philosophy's most famous thought experiments - the brain in the vat. He claims the disembodied brain floating in a vat of nutrients, without perfect duplication of the inputs and stimuli outputs, might not even be able to think. For similar reasons we should reject Hudson's view which amounts to considering the person to be "a brain in a living vat." Damasio explains:

In brief, neural circuits represent the organism continuously, as it is perturbed by stimuli from the physical and sociocultural environments, and as it acts on those

⁹ My stress on the contribution of life processes to thought production should not be interpreted as denying that the non-living can think. If thought could occur after the cessation of life processes, mechanical substitutes would be needed. The brain in the vat envisioned by philosophers needs the vat to function. The thinker, if composed by what contributes to thought production, would then have such mechanisms as parts. If contribution determines composition, we should speak not of a thinking brain in a vat but of a thinking brain/vat composite.

environments. If the basic topic of those representations were not an organism anchored in the body, we might have some form of mind, but I doubt that it would be the mind we do have...the body contributes more than life support and modulatory effects to the brain. It contributes a *content* that is part and parcel of the workings of the normal mind [4, p. 226].

It may be that there are other sources of Hudson's claims about the parts of the nervous system being the components of the person rather than the entire organism. He is, after all, a product of the intellectual community that widely holds the view that the brain produces thought much as the stomach's digestive tract produces gastric juices. In Alva Noë's diagnosis, such views are based upon what he calls *The Foundational Argument*. The argument draws primarily upon direct stimulations of the brain and the experience of dreaming as evidence for the claim that the brain's neurology is sufficient for thought. This, in turn, provides reasons to think we persons consist just of the brain's parts. I will briefly sketch the reasons Noë provides to reject it. His preferred understanding is "that brain, body and world together maintain a living consciousness" [11, p. 42]. Consciousness arises from the dynamic activity of the organism with the world. Consciousness "is something we achieve..., more dancing than digestion" [11, p. xii].

Noë points out that if our thoughts were just produced by the brain, then one would expect stimulation of the brain to reproduce the thoughts that were realized there. But the plasticity of the brain suggests some reasons to be skeptical of the view that we are found within our skull. Quite revealing are studies of newborn ferrets that had their eyes wired to parts of the brain used for hearing. One would think that such creatures would hear the results. Instead of hearing with their eyes, they saw with parts of the brain previously used to hear. The lesson Noë draws from "the character of conscious experience vary(ing) even though the neural activity underpinning it does not change...(is) that what determines and controls the character of conscious experience is not the associated neural activity" [11, pp. 53-54]. Another revealing experiment enabled the blind to have a vision-like experience. Paul Bach-y-Rita placed 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. Noë insists 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" [11, pp. 58-59].

Noë does admit that some stimulus of the brain causes sensations [11, pp. 173-4]. For example, electrodes placed in the brain can give rise to sensations of light or illusions of motion. But he objects that is no reason to think *all* experiences could be so triggered. One obstacle to such a conclusion is that there is feedback as a result of the body's role in changing its relation to the environment, thus making conscious experience far more than the result of a pattern of brain inputs. Furthermore, such manipulations are affecting existing consciousness, not generating it. However, even if the future brings brain in the vat technologies that create hallucinations corresponding to *all* of our experiences, that still wouldn't mean the brain has produced consciousness. At best, it means that consciousness is produced by the combination of brain, the vat and the experimenter. Producing changes in consciousness isn't the same as claiming that consciousness is produced by the brain. Finally, the brain in the vat doesn't have veridical experience so it doesn't capture all experience unless one can maintain the skeptical thesis that the real world is just a virtual world, a Matrix-like creation.

The second part of the Foundational Argument is that dreams are held to show that consciousness is not a dynamic bodily production but something that just transpires inside us

[11, pp. 177-78]. But this may be assuming that every experience can also be dreamt. 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. Moreover, it seems that dreams depend upon earlier experiences of the entire waking animal that are the product of dynamic engagement [11, p. 180].

Given Noë and Damasio's arguments, I doubt that Hudson can rely upon the notion of "a contributory role" in supporting personhood to so shrink the size of the person. Another explanation of the failure has been offered by Eric Olson, who speaks of *direct involvement* with the production of thought instead of the near equivalent *contributory role*. Olson 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 [12, pp. 91-98]. 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" [12, p. 92]. 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.¹⁰ Instead, the fault lies in the notion of *directly involved* being unprincipled.

Although Hudson doesn't say much about the crucial notion of *contribution*, perhaps a refined notion can evade or mitigate the criticisms made so far. Perhaps I have been conflating the causal and the constitutive [2, p. 982]. A more nuanced conception of *contribution* might distinguish an instrumental causal condition in the past from the neurological basis constitutive of thought in the present. Borrowing from Clark, the constitutive reading of contribution could be limited to the "vehicles of mental states and processes," the physical realization of *information* that governs actions [2, p. 966]. Moreover, a distinction could be made between content and vehicle [2, p. 966]. The vehicle is just neurological.¹¹ Its content may be of extra-neurological states of the body or environment, as Damasio and Noë speculate, but the person is just composed of the vehicle. The body may play an instrumental causal role in the production of thought, but the neurology is sufficiently constitutive of thought. I'm willing to assume, for the sake of argument, that the life support needed for thought need not be simultaneous, i.e., the animal could die but the brain could very briefly continue to function with the resources that life processes earlier delivered. Thus life processes have an instrumental causal role in the production of thought, but are not constitutive of thought. They are more akin to the object dropped on the foot and the damaged tissue that "drives" the later neurological realization of the painful sensation. Let's also assume Olson is wrong and that the notion of "directly involved" is just vague rather than unprincipled, so Hudson's epistemicism will guarantee that the person has precise boundaries within the animal. With such assumptions, Hudson might claim we have good

¹⁰ So Hudson's epistemicism won't help him.

¹¹ Even Damasio claimed a brain in the vat with perfect duplication of inputs and outputs of the embodied brain may have the same experiences [3, p. 228]. But see note 9 for why all that might show that it is the brain/vat complex that is thinking.

reason to believe that the person's present temporal part consists of whatever parts of the neurology now are the vehicles of the representation, and that person's later temporal parts will involve the physical realization of psychologically continuous states. However, readers will see in the second part of the next section (IV) why that won't work even if this refined notion of *contribution* turns out to be able to demarcate the neurological thought producing elements.

IV. The Human Animal is the Only Person

A. The Collapse of Psychological Continuity into Biological Continuity

I will now offer a very different line of reasoning for identifying the Four-Dimensional human animal and the human person.¹² I will show that the psychological continuity and connectedness criterion favored by Hudson [7, p. 144] *collapses* into animal identity. What I mean by *collapses* is that there are cases which tend to elicit from us descriptions of one thinking entity being identified with another thinker that cannot be explained by a psychological criterion being satisfied. The intuitions we have there about identity can only be accounted for by both thinkers being the same animal. So what we want to say are stages of a persisting person in cases involving the dreaming and the awake, the rational and the demented, divided and reunited minds, can only be construed as such if an appeal is made to the biological persistence conditions of animals.

The first problem for the psychological account of identity involves a twist on Reid's famous critique of Locke's memory criterion [15]. Locke claimed that one's identity extended as far back in time as one's memories. Reid revealed a failure of transitivity by envisioning an old general who could remember his first military campaign as a young soldier, the young soldier could recall being flogged as a school boy for stealing from an orchard, but the general couldn't remember being flogged. Therefore, the general is *not* identical to the boy, yet he is identical to

¹² The arguments can be extended to three-dimensionalist denials of the identity of the human animal and person.

the young soldier, who is identical to the boy. This absurdity could be avoided by appealing to psychological continuity, i.e., overlapping chains of psychological connections [14, pp. 206-09]. Psychological continuity involves the general being able to remember a time (his first military campaign) at which he could remember being flogged. So an overlap of memories will suffice in lieu of a direct memory connection. But the transitivity problem returns with a modified version of Reid's scenario that Perry named the *Senile General* case [15, p. 19]. The senile general could remember being flogged (or remember a time at which he could remember a time that he was flogged). So he is identical to the boy. The young soldier could remember being flogged, so he too is identical to the boy. But the general couldn't remember his more recent experience as a young soldier, nor could he remember any other time at which he then could remember his first military campaign. This renders the general identical to the boy but not identical to the young soldier, who is also identical to the boy. So if they are to be identified, as it intuitively seems they should, an appeal to their being the same animal can do what an appeal to psychological continuity cannot.

I don't think it will work to claim that the general and the young soldier can be identified as long as they both recall, via direct memory or overlap of recollection, the child beaten in the orchard. Advocates of a psychological account of identity generally don't want to say that if Hannah appeared to cease to exist at noon and just by coincidence, a psychological duplicate of her popped into existence a moment later, Hannah would have reappeared elsewhere. As we noted earlier, the reason why Hannah hasn't survived is that there isn't the appropriate causal relationship between Hannah and the duplicate. There is thought to be some kind of immanent causation that must link the psychological states of Hannah at one time with psychological states

at a later time.¹³ Notice that there would be no immanent causation between the temporal parts in the altered Reid case where the general becomes senile late in life and as a result can't remember or have an overlap of memories (or any other psychological ties), to the young soldier at the time of his first military campaign, but both the general and the young soldier can recall (or have an overlapping chain of memories back to) the child beaten in the orchard for stealing an apple. So even though both the general and young soldier have psychological contents that are immanently caused by the young boy, there is an absence of the requisite causality between the temporal parts linking the perduring general and young soldier. The senile general is psychologically the same as he would be if he had just woke up from a coma that he had been in from the time before his first military campaign as a young soldier to the present time.¹⁴

Attempting to preserve the identity between of the general and the young soldier in virtue of the common memory link to the young thief and the transitivity of identity relation would be as illegitimate here as it would be to claim that the two products of fission are the same person because they are both linked to the pre-fission person. The post-fission mental states wouldn't be appropriately immanently causally linked to each other, thus the thinkers are not identical. Likewise, for the modified Reid case, in which the young soldier and the general do not have the appropriate causal links between their respective mental contents, neither having access to or causal overlap with the others.

¹³ This relation is more important to persistence than spatio-temporal continuity, so says Dean Zimmerman [19]. Hudson concurs and draws upon Zimmerman's work [7, pp. 34-37].

¹⁴ It would be a mistake to endorse, as Perry does, Grice's suggestion that the general is the young soldier because he could have remembered the event [15, pp.19-20]. That is too promiscuous a criterion. There will be no psychological change that isn't identity preserving because the resulting being could have remembered the earlier events if it wasn't for say the neurosurgeon's intervention in a Parfit spectrum-like case or the comatose could have recalled if not comatose and so on.

I imagine that someone might object to the analogy because there aren't two concurrent branches in the senile general case as in fission cases. However, fissioned hemispheres and transplantation could be staggered. Someone's upper brain could be divided, one hemisphere removed and frozen and the other transplanted. When the recipient of the first transplant dies, the remaining hemisphere is thawed and transplanted. The Four-Dimensionalist would surely not claim the two post-transplant branching series of stages were the stages of the same person. The Four-Dimensionalist reasoning here isn't fully explained by the fact that two conscious states emerging from fission occur at the same time, rather, it is due to their not being causally related to each other.

A second scenario where a psychological criterion of identity collapses into a biological one involves a temporary division of a mind. Consider Parfit's *My Physics Exam* scenario where there is just a short term loss of a unified consciousness due to cutting the corpus callosum so one person can direct both hemispheres to work on different parts of a test [14, pp. 246-48]. The hemispheres are reunited after the dual work is done. As Parfit himself notes, the most plausible response is that there was one person temporarily cut off from himself. To account for that intuition, something other than a single causal chain of psychological continuity must be relied upon. Four-Dimensionalists usually qualify the criterion of psychological continuity for cases involving fission and fusion so the result is that there are two distinct persons continuous with the same earlier stage. They do so by insisting that psychologically continuous x and y are stages of the same person if there is no stage z that is psychologically continuous with x or y but simultaneous and distinct from either y or x [1]. So during the exam there are two streams of thought that have stages that are simultaneous but distinct from each other, thus ensuring that there is not a single person despite their both being psychologically continuous with shared earlier stages. But this will deliver the counterintuitive result that there is not a person with the

briefly divided mind but that there were two persons present at that time since they involve simultaneous but distinct stages. If the intuitive response is to be preserved, then it appears that we must appeal to a rather ad hoc modification of the psychological criterion or claim that it must be because it is the same animal doing the thinking.

One can also undermine the psychological continuity criterion for identity by taking issue with Locke's account of Socrates awake and Socrates asleep [9, p. 343]. Locke conjectured that if sleeping Socrates was psychologically cut off from waking Socrates then they would not be the same person. Imagine that your waking and dream states are not psychologically connected. You cannot recall your dreams and these dreams don't follow from your waking life. I suspect that few readers would follow Locke and deny that they were states of the same person, interpreting the psychological disconnect as evidence of two people sharing a body. Since there isn't any psychological continuity between the waking and the sleeping, then what makes them the same person must be that they are the same living animal. It might help drive the point home if readers imagine that medical technology reveals that every night they have horrible nightmares though the following day always wake up without any memories of such dreams. I suspect that if readers could prevent these nightmares by doing something when awake, they would. And readers would do so for prudential reasons, not moral concerns about alleviating the suffering of another.¹⁵

¹⁵ I believe these ideas of one person whose thoughts are cut off from himself can be extended to challenge McMahan's interpretation of the dicephalus – consisting of one animal with two heads - as being two persons [10, pp. 35-39]. It might be argued that while Sleeping and Waking Socrates should not be considered two persons sharing a body, the lesson cannot be extended to McMahan's two headed case for Sleeping Socrates and Waking Socrates are not thinking concurrently and the concepts made use of by Sleeping Socrates are acquired from Waking Socrates. I argued that these differences don't undermine the lesson by analogy in [4].

A fourth scenario undermining psychological continuity theories relies upon our reactions now to the possibility of future pain after the onset of amnesia or even more debilitating impairments [10, 17]. Consider the prudential concern many envision having for the being with their brain after a stroke undermines the brain's capacities for rationality and self-consciousness, leaving a mere sentient child-like mind. If told earlier that the being with our damaged brain will suffer horrific pains unless we take on almost as much physical pain before losing our memories and capacity for self-consciousness, most of us would consent to the lesser pain to ensure the greater does not transpire. Such a show of apparently prudential concern for an animal in the future, despite the absence of psychological continuity and the reflective capacities associated with personhood, suggests an adherence to an animalist, i.e., biological account of our identity.

What I have been hoping to get readers to recognize with the *Sleeping Socrates*, *Physics Exam*, *Senile General* and future pain scenarios is that there is a divergence between the psychological criterion and our intuitions about our survival. Only a reliance upon animalist identity conditions can to accommodate our judgments of persistence. I suspect, however, that some readers will offer an alternative interpretation. Their response is that it is not psychological continuity that matters to our persistence, but the capacity for mere sentience - minimal thought and feeling. As long as the same brain sustains sentience, then the individual survives despite memory loss and even some mental fragmentation.

Hudson contends that an individual suffering "profound senility" would not be a person [8, p. 222]. There wouldn't be the requisite self-consciousness and psychological continuity. But our prudential concern in the typical philosophy thought experiments suggests we would survive

a loss of mental capacity.¹⁶ So while I think this should lead Hudson to abandon his belief that *we* are essentially self-conscious persons [8, p. 218], given unrestricted composition, it need not lead him to deny that there are beings that are essentially self-conscious with psychologically continuous stages. However, if anything deserves the title “person”, we do. So given Hudson’s commitment to a maximality principle, he should accept that we persons are not essentially self-conscious psychologically continuous thinkers, but merely self-conscious for just a period of our lives. We’re persons because of our capacity for self-consciousness, but that capacity need not be actualized during all of our stages. Of course, even if Hudson were to admit this, it still wouldn’t commit him to identifying the human person and the human animal. He could instead claim we persons are identical to a maximal being composed of all merely conscious stages rather than only self-conscious, psychologically continuous stages. McMahan and Unger offer Three-Dimensional versions of this thesis [10, 17], claiming that we survive as long as the same brain produces sentience (consciousness). So what I propose to do in the next section is provide thought experiments which suggest that our prudential concern reveals that we persons believe ourselves to be not even essentially sentient. These thought experiments reveal that the future sentient states we are concerned with can be deemed ours only if they are united by a biological criterion.

B. The Collapse of Brain-Based Psychological Identity into Biological Identity

My contention is that a Four-Dimensionalist can be brought to see that the human animal is the only person by drawing upon our concern for our stages that are devoid of the traits that characterize personhood. Our prudential concern towards our adult conscious animal in the

¹⁶ Hudson assumes that appeals to personal identity thought experiments will end in stalemate [8, p. 217]. I am more sympathetic to his later “acknowledge(ment) that my dismissal of the fanciful thought experiment defense may have been uncharitable and over-hasty...” [8, p. 233].

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 will hold 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 to psychological accounts of personal identity. The first involves your brain being destroyed and replaced by a new brain. The second thought experiment involves you swapping brains 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 depends upon 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 [16, pp. 57-59] 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. And I doubt readers would assert that there is a new thinking being, one of the same kind as they, 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 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. 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. I don't see any reason to identify ourselves with parts of the consciousness-producing central nervous system [7, 10],

nor with a larger being only if it *continuously* possesses the same functioning brain-like structure [17], 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.

Perhaps you will initially argue that you would survive with *any* parts of your existing brain contributing to the production of conscious states, but would perish if your brain ceased to exist. I believe opposing intuitions can be elicited. Imagine that now and after a debilitating stroke that your pain is received and realized (in some sense) by the upper spine while pleasure has a cerebral basis. I assume that pondering this prospect doesn't eliminate our now having prudential concern for the post-stroke creature in pain that lacks the capacity for self-conscious reflection. It seems that the best explanation of why these would be your pains and pleasures is that the parts involved with producing them are caught up in the same life, i.e., they belong to the same animal.

So it appears that the two most prominent psychological criteria of diachronic identity (*self*-conscious psychological continuity or brain-based persistence of *mere* consciousness) can't deliver the intuitive response - that there is but one and the same thinker in the stroke case. What can do so is the animalist account in which human persons and human animals are identical. Thus it makes sense to claim that the only person in the stories is the animal. As long as our animal can have pleasures and pains into the future, we have some prudential reason to obtain the former and avoid the latter.

Our attitudes of prudential concern provide additional reasons to reject Hudson's idea that the person consists of that which beneath the skin directly produces thought. Thus even if some sense can be made of the proper part of the animal being what directly produces or contributes to thought at this moment, say the neurological realization of some information, it doesn't seem to be the entity for which we have prudential concern. If different parts of our

organism would later constitutively contribute to painful sensations, we would be prudentially concerned with preventing these feelings. The neurological states that matter to us are those of the *same* organism, they are not identified by causal connections between earlier and later neurological states that would underlie a psychological continuity or mere (brain-based) sentience thesis.

If you share my attitudes to the individuals with maimed or reduced brains, then why maintain that we would have no prudential reason to care about one's animal if it received a new cerebrum in a thought experiment after the old was destroyed?¹⁷ And if you admit that you have some prudential concern for your animal with a new cerebrum, then you can't also claim to have prudential concern for the being who would receive your cerebrum in a second thought experiment that involves a brain swap between you and your clone. This is not to deny that you can care about the recipient of your functioning cerebrum even though that person will not be you. I don't even have to endorse the claim that your commitment to the human animal with your original brain ought to be less than your concern for yourself with a new upper brain. My point is just that you cannot have *prudential* concern for both since prudence is *self* concern. So I don't have to claim you are irrational to care about the other person/animal who receives your functioning cerebrum in the transplant swap scenario where you stay behind as an animal with a new upper brain. Such concern would be no more irrational than caring more about your spouse or your child than yourself. Nor do I have to follow my fellow animalist, Eric Olson, and treat sympathetically the Parfit-Shoemaker claim that what matters to us is not identity but psychological continuity [12, pp. 42-72].¹⁸

¹⁷ Likewise, would it not be good for your embryonic child to grow a normal brain? I have elsewhere tried to make the case that mindless organisms have interests in their well-being [5].

¹⁸ Part of what I mean by *identity mattering* is that we must be identical to the future subject of our psychology if there is not to be some drop in prudence-like concern for that thinker. I'm not committed by this thesis to our caring

So once readers see that thinkers are best individuated by life processes, it becomes arbitrary to claim only part of the animal is a person. One can still, on the basis of unrestricted composition, claim that the person consists of only scattered thinking stages of organisms before and after the stroke-induced coma and injury. But the stages of the animal don't have the right causal connections. Such a 'person' is an artificial, gerrymandered product of the principle of unrestricted composition, not an entity possessing either a natural biological or psychological unity between its stages. Calling such an entity a person would be as suspect as claiming the first half of my life and the second half of your life would compose a person. There is no immanent *mental* causation between the thoughts of the person who suffers the stroke-induced brain damage and temporary coma, and the later pains and pleasures. Likewise for the other scenarios discussed. If immanent causation is needed, then it would be in the form of life processes unifying sleeping and waking Socrates, the senile general and the young thief, the later stroke victim and the earlier rational self, the merely sentient newborn and the later reflective child, or the divided and then reunified mind studying for Parfit's physics exam. So we see that our prudential intuitions, our belief that we are persons if any entities are, and the maximality principle all serve to indicate that the human animal is the least arbitrary candidate for the persistence of the person in the above cases.¹⁹

prudentially about our later irreversibly comatose organism. But this thesis and my belief that the human person is identical to the human animal does commit me to caring about my animal's future psychology even if that thinking animal has a new brain and hence no psychological continuity to me now.

¹⁹ I would like to thank Adam Taylor for considerable help with this paper.

References

- [1] Brueckner, Anthony and Buford, Chris. 2008. The Psychological Approach to Personal Identity: Non-branching and the Individuation of Person Stages. *Dialogue*. 47: 377-386.
- [2] Clark, Andy. 2009. Spreading the Joy? Why the Machinery of Consciousness is (Probably) Still in the Head. *Mind*. 118:472. 962-963.
- [3] Damasio, Antonio. 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Penguin Books.
- [4] Hershenov, David. 2004. Countering the Appeal of the Psychological Approach to Personal Identity. *Philosophy* 79: 445-472.
- [5] Hershenov, David. 2011. "Embryos, Four-Dimensionalism and Moral Status. In *Persons, Moral Worth, and Embryos: A Critical Analysis of Pro-Choice Arguments from Philosophy, Law, and Science*. Ed. Stephen Napier. Springer.
- [6] Hudson, Hud. 1999. Temporal Parts and Moral Personhood. *Philosophical Studies* 93: 299-316.
- [7] Hudson, Hud. 2001. *A Materialist Metaphysics of the Human Person*. Ithaca: Cornell University Press.
- [8] Hudson, Hud. 2007. I am not an Animal! In *Persons: Human and Divine*, Eds. Dean Zimmerman and Peter van Inwagen, 216-234. Oxford: Oxford University Press.
- [9] Locke, John. 1975. *An Essay Concerning Human Understanding*. Ed. Peter Nidditch. Oxford: Oxford University Press.
- [10] McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.
- [11] Noë, Alva. 2009. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York: Hill and Wang.

- [12] Olson, Eric. 1997. *The Human Animal: Identity without Psychology*. Oxford: Oxford University Press.
- [13] Olson, Eric. 2008. *What are We? A Study in Personal Ontology*. Oxford: Oxford University Press.
- [14] Parfit, Derek. 1983. *Reasons and Persons*. Oxford: Oxford University Press.
- [15] Perry, John. 2008. 2nd ed. In *Personal Identity*, ed. John Perry. Berkeley: University of California Press.
- [16] Shewmon, Alan D. 1997. Recovery from 'brain death': a Neurologist's Apologia. *Linacre Quarterly* 64:1, 30-96.
- [17] Unger, Peter. 2000. "The Survival of the Sentient." *Philosophical Perspectives*. Ed. James Tomberlin. 14: 328-345.
- [18] Van Inwagen, Peter. 1990. *Material Beings*. Ithaca. Cornell University Press.
- [19] Zimmerman, Dean. 1997. Immanent Causation. *Philosophical Perspectives*. Ed. James Tomberlin, 11: 433-471.