

# 1. Human Beings, Human Animals, and Mentalistic Survival

*Denis Robinson*

## INTRODUCTORY REMARKS

The debate over personal identity is a complex one, involving many contrasting views, and ingenious and controversial arguments. But standing back from the complexity, we can see a couple of major and strongly contrasting groupings into which many recently defended views may be seen as falling. At one extreme, we find diverse forms of Psychological Reductionism, and at the other, views we may group under the label "Animalism".<sup>1</sup> Though Psychological Reductionism is (deservedly, I think) more popular,<sup>2</sup> Animalism also has its following, and has fairly recently been given an unusually blunt, succinct, and passionate defense by Eric Olson in Olson (1997).

The doctrinal opposition between these main groupings tends to line up rather neatly with a general methodological opposition which has ramifications in metaphysics generally: one where the key issue is, roughly, the relative priority in metaphysics of folk intuitions and a priori judgments, as against the findings of natural sciences.

But the very distance between these two main doctrinal groupings makes it natural to look for some kind of intermediate position, and for some corresponding methodological middle way for defending it. My aim in this paper is to inspect this intermediate terrain, and in so doing to draw doctrinal and methodological conclusions of my own. My sympathies are with Psychological Reductionism, rather than with

<sup>1</sup> An account of the various positions mentioned in these opening paragraphs will follow shortly.

<sup>2</sup> Johnston calls "dominant" a genre of Psychological Reductionism which he names "Wide Psychological Reductionism", citing Sydney Shoemaker, Anthony Quinton, and David Lewis as amongst its defenders (Johnston 1987: 61). Olson (1997: 170 n. 13) lists over thirty philosophers, including nineteen whom he refers to (p. 20) as "big names", as holding some kind of Psychological view.

Animalism, but it is not my aim here to defend the former, nor to attack the latter. Rather, I aim to bring out the importance in this area of very general ontological assumptions or presuppositions, and the way in which the outcome of applying a particular methodology may be radically affected by them. I shall build my discussion around one well-known attempt—Johnston (1987)<sup>3</sup>—to defend a middle way between the extremes of Animalism and Psychological Reductionism. I shall argue that the attempt is unsuccessful, that the middle ground it attempts to occupy is unstable, and that the methodological proposals invoked in its defense cannot do what is asked of them.

Not only is Johnston's view in Johnston (1987) (henceforth, "the Human Beings view"<sup>4</sup>) one which might be thought to combine elements characteristic of our main opposing camps, it also offers what might be seen as an attempt at a methodological middle way in support, since (as we shall see below) it invokes a priori intuitions about personhood on the one hand, while expressing caution about such intuitions, and prioritizing naturalistic ontological categories, on the other. I shall argue, however, that any substantive alternative to Psychological Reductionism which that methodology can be made to support will be even closer to Animalism than the Human Beings view, which is itself, despite the appearance of compromise, already quite close to an Animalist view.

To move on, we must begin by saying more about the kinds of positions referred to above. I shall say quite a bit about Johnston's position, the kind of ontological standpoint which underpins it, and the errors, as he sees them, which he wishes to avoid, before finally examining his methodological proposals, and the positive arguments for his view, to which they lead.

*Psychological Reductionists* agree that the constitutive criteria for personal identity over time are predominantly psychological. Various kinds of psychological connection may hold between a person's earlier

<sup>3</sup> It is important to realize that Johnston (1987) alone does not give an accurate impression of Johnston's overall position, as he in effect points out in Johnston (1987) n. 8, p. 64. Fully to appreciate his subtle and complex view one must read also Johnston (1989a), Johnston (1989b), and Johnston (1992b). Of these Johnston (1989b) is perhaps both the most important and the most difficult. A critique of Johnston's whole view would be a much larger undertaking than space here permits. My own take on the personal identity debate may be found in Robinson (2004).

<sup>4</sup> I capitalize the phrase "Human Beings" throughout, to signal that I am using it in Johnston's precisely articulated sense, not merely in an everyday sense.

and later mental states: combinations of such connections constitute the essential conditions for personal identity over time. Psychological Reductionists differ over the relevant connections and combinations, and their required causal underpinnings. (Johnston speaks of “Wide Psychological Reductionism”, where “Wide” connotes a relatively liberal attitude to the permissible kinds of causal underpinnings.)

*Animalism* is the view that the familiar persons of our acquaintance are strictly and literally identical with members of a particular animal species, *Homo sapiens*. Members of this species are typically persons for most (but perhaps not all) of their lifespans. There is no special problem of *personal* identity: identity-conditions for a person of this familiar kind are just identity-conditions for the relevant kind of animal, much like the identity-conditions for other animals. An animal is born, it lives, it dies. For these events and processes, there are no essential mental prerequisites, certainly no requirement of psychological connectedness or continuity. Thus, Animalism is less a view about the nature of persons in general, more a view about the nature of those with which we are mostly acquainted. If there are angels, thinking automata, Martians, or dolphins which are people, then the identity-criteria for those persons will just be those appropriate to such beings: there are no identity-conditions for persons per se.

As noted, the doctrinal opposition between Animalism and Psychological Reductionism aligns naturally, though not inevitably, with a methodological opposition. Psychological Reductionist views are typically justified largely by appeal to intuitions about the loss or preservation of personal identity in various kinds of real and imaginary cases: by the “method of cases”, as Johnston calls it. But this methodology has been called into question in a variety of ways.

Quine popularized the view that the analytic/synthetic distinction is at least hazy, and at best a matter of degree—some truths in which a term figures merely being more centrally constitutive of the term’s meaning, or the concept it expresses, some less. This casts a shadow on any correlatively understood enterprise of conceptual analysis, of which the “method of cases” is typically a core component. The shadow falls particularly on our ability reliably to recognize a priori the genuine (“metaphysical”) possibility or impossibility of some circumstance—as opposed, for instance, to its degree of familiarity or bizarreness, or its relationship to well entrenched empirical assumptions. Complementing both this aspect of Quine’s work, and his epistemological naturalism, we

have Kripke's and Putnam's work on rigid designation and natural kinds, suggesting that the essence of a kind of thing might need to be discovered by empirical inquiry, with conceptual analysis being, not merely epistemically suspect, but quite beside the point.<sup>5</sup>

These trends lead easily to the thought that we should reject the method of cases in favor of seeking where possible to identify the category of familiar persons with some *prima facie* appropriate natural kind. Biological kinds, and in particular animal species, are standardly taken as paradigms of such kinds. With a little help from evolutionary theory, Animalism emerges from these considerations as a natural rival to Psychological Reductionism.

In Johnston (1987) Johnston mounts a qualified critique of the method of cases, specifically in its application to the topic of personal identity, and proposes an alternative to Psychological Reductionism which, like Animalism, gives crucial precedence to the notion of a biological kind.<sup>6</sup> But he uses the notion of a biological kind in a more subtle way than does Animalism, in arriving at an account of the persistence-conditions for persons. Appealing to an alternative methodology, he argues that familiar persons are to be identified with "Human Beings": entities having persistence-conditions similar to, but not quite the same as, those of human animals. Indeed, though the kinds "Human Being" and "human animal" differ in their persistence-conditions, the crucial differences are rarely manifest in ordinary life, concerning as they do extreme and mostly imaginary circumstances.

Johnston's argument for this view appeals to the special role of certain psychological states, namely experiential memories, in our knowledge of personal identity; in this it is reminiscent of Psychological Reductionism, particularly in those neo-Lockean versions which make experiential memory crucial. Nevertheless, Johnston (1987) rejects the idea of straightforwardly mentalistic identity- or persistence-conditions. On the Human Beings view, psychological states and relations have a

<sup>5</sup> Though popular, the idea that Quine, Putnam, and Kripke provide us with reasons, good or bad, to replace conceptual analysis in the service of metaphysics with appeals to the findings and categories of the natural sciences is not inevitable. Frank Jackson's recent *Jackson (1998)* is one work which defends the view that these supposedly disparate methodologies should actually be seen as complementary.

<sup>6</sup> I shall not much discuss Johnston's attack on Psychological Reductionism as based on the method of cases. It revolves in part around a well-known conundrum for the method of cases due to Bernard Williams. For some criticisms of it, see Oderberg (1989).

merely evidential rather than a constitutive relation to personal identity. Any possibility that a person's survival might be sustained by psychological connections which transcend certain biologically based limitations, as in the fantasies of teletransportation, or of machine-assisted transfer of psychological states (or the informational states which realize them) from one brain to another, is ruled out. In all this, the view resembles Animalism. The key difference is that a Human Being, unlike a human animal, could persist through transplantation of its living brain into a different body, or in the form of the proverbial "brain in a vat".

“HUMAN BEINGS”, “HUMAN ANIMALS”,  
AND THE CONSTITUTION RELATION

Johnston's view is that the (normal) relation between human beings and human animals is one of *constitution*, using this word in a technical sense commonly employed, with minor variations, by contemporary ontologists. The constitution relation holds between material things which occupy the same place at the same time and are composed of the same matter but which are nonetheless not strictly identical.<sup>7</sup> An important special case is the relation between a material substance and the matter which composes it. Notoriously, the constitution relation is a time-relative or changeable relation, since the matter composing material substances may change.

It's important to notice that the constitution relation may hold between material substances—*bona fide*, everyday *objects*—and not just between a substance and its composing matter. For instance, a mountain might be composed of a single gigantic rock; subsequently thermal stress may cause the rock to fracture into small pieces, yet the mountain persists. The mountain is at one time, not at others, constituted by the rock. Having different histories, they are never identical. Similarly, a cloak composed of a button and a piece of cloth may lose its button, which may be destroyed. Subsequently the cloak is constituted (solely) of the piece of cloth. Cloak and cloth are never identical, since at

<sup>7</sup> At least not in typical cases. There is a debate about whether identity can ever be a case of constitution. Cf. n. 9.

an earlier time the button is part of the cloak, though it is never part of the piece of cloth.<sup>8</sup>

Thus, the constitution relation is importantly different from identity, despite being readily confused with it.<sup>9</sup> An important instance of this general distinction, is that between cases in which one thing “is” another, in the sense of constituting it—as the boulder, at certain times, “is” the mountain, or the cloth “is” the cloak—from cases in which an entity falls, temporarily or contingently, under a “phase-sortal” predicate. The latter are cases of genuine identity. A person begins life as an infant, passes out of childhood into adolescence, and, with luck, grows out of adolescence into adulthood. For a limited time, the person “is” the adolescent. But this is an “is” of identity, rather than of constitution. The person *ceases to be* an adolescent, but is nonetheless *identical with* that adolescent. The adolescent, on ceasing to be an adolescent, does not on that account cease to be.<sup>10</sup>

<sup>8</sup> These examples are from Robinson (1982: 319). Though Olson has elsewhere contributed to the literature on the constitution relation, in Olson (1997: 101–2) he is remarkably curt and dismissive about views appealing to it. But if it is hard to deny there may be clear examples of ordinary things standing in this relation, appeal to it ought to be a *prima facie* option in the personal identity debate.

<sup>9</sup> Another point of which these examples may remind us is that, even after we set aside instances of the “is” of predication, it is neither necessary nor sufficient for the concept of identity to be invoked by a given statement that terms appear linked by “is” or other forms of the verb *to be*. In fact, I believe that even such everyday locutions as “is identical with” need not signal identity in the philosophers’ sense: for they, too, may be used by the folk to express the constitution relation. It will be evident that I do not in general take speakers’ introspective judgments as to their meanings or concepts as an infallible guide, since ordinary speakers will not automatically distinguish particular utterances which attribute the constitution relation from assertions of identity in the strict, philosophers’ sense. Failure to note these points makes it spuriously easy to impute to common sense a strong commitment to Animalism.

As a four-dimensionalist I believe that the constitution relation can be analyzed in terms of identity of temporal parts: *x* constitutes *y* at *t* iff the appropriate temporal part of *x* is identical with the contemporaneous temporal part of *y*. In special cases this means that identity is a limiting case of the constitution relation. Johnston however rejects such an approach (see Johnston 1992a). It follows that, on his view, no Human Being is ever identical with a human animal, nor with any temporal part of a human animal.

<sup>10</sup> For present purposes, we may characterize a “phase-sortal predicate” as follows. A “substance predicate”, first of all, is one which applies to all and only members of a kind of entity, where members of a kind in the relevant sense share, *inter alia*, identity- and persistence-conditions. Applying a substance predicate to an entity thus carries implications for how it and members of its kind are to be individuated and tracked through time, and for which kinds of events they can and cannot, logically, survive. A phase-sortal predicate (such as “adolescent”, “blonde”, or “widow”), finally, is logically equivalent to the conjunction of a substance predicate with attribution of some temporary or accidental characteristic, and thus, by the rules for conjunction, carries similar implications. See Wiggins (1980).

The important contrast here is with the case of the boulder which, I say, does indeed cease to exist when it fragments. Thus we should *not* see ourselves as referring to the mountain itself, under the description “boulder”, when we refer to the boulder (though in many contexts it may be a matter of indifference which we refer to, so that we leave it indeterminate). Even more clearly, when we speak of the piece of cloth, in the example of the cloak, we are *not* merely speaking of the cloak itself under a phase-sortal description comparable with “adolescent”.

Here I follow Johnston in adopting the vocabulary of what I call a *substance-ontology*. Wiggins (1980) is a well-known exposition and defense of principles typical of such ontologies (though with, inevitably, a number of idiosyncrasies and optional extras). According to these ontologies, the world contains a *limited variety* of *kinds* of *substances*. Each of these kinds has distinctive criteria associated with it, criteria which, if articulated, would specify what kinds of events could count as commencing or terminating the existence of members of the kind, and how members of the kind are in principle to be identified, distinguished, counted, and traced over time. Important amongst the substance kinds are the natural kinds, including the biological kinds. Substance-theorists who prioritize natural kinds typically view the associated persistence-conditions as naturally, as opposed to conventionally, determined.

Consistently with these views, substance-ontologists typically deny the reality, or disparage the status, of arbitrary mereological aggregates of substances or other entities. They also typically deny the reality of temporal parts of substances, and insist on sharply distinguishing issues of change and identity for substances from parallel issues involving either events, or portions or quantities of matter.

Substance ontologies may differ widely in the range of kinds they recognize, and how they see those kinds as related. Roughly speaking, we can distinguish the *form* of a substance-ontology, from its *content*—the particular view it takes as to what the constraints on substance-concepts, and on substances, are. Natural kinds will typically be included, but that leaves open what other kinds might also be admitted. Suitably well-defined or determinate conditions for identity and persistence will always be required, but there is much room for debate as to what “suitably well-defined or determinate” means here. But however these issues are resolved, questions about substance-identity will for substance-theorists always ultimately come down to questions about either nominally or

metaphysically essential properties of kind-members: those properties possessed essentially by things belonging to the kind in question.

Thus we may say that on the substance-ontology approach, persons will comprise all or some of the members of all or some of the kinds of substances which typically have what it takes to be a person. For present purposes, and compatibly with all Johnston says, we can sum up what it takes to be a person, as the capacity for a sufficiently rich, sufficiently integrated kind of mental life. Consistently with all this, one *might* take it that the kinds of substance there are include *none* whose persistence-conditions coincide with anything which could plausibly be considered as *psychological* unity-conditions—conditions for the unity and persistence of a mind considered as such. An upshot of such an attitude would be that “person” might not qualify as a substance-concept, nor even as a phase-sortal concept, being more akin to a mere general concept like “fast mover”, or “red thing”. Given such assumptions, Animalism is a natural (and “naturalistic”) conclusion.<sup>11</sup>

Johnston adopts, not just the vocabulary, but the stance, of substance-ontology, as sketched above, including, it seems, the rejection of mentalistic persistence-conditions for bona fide substances.<sup>12</sup> Certainly his argument at a crucial point seems to rest on *some* restrictive set of background ontological assumptions—assumptions which effectively load the dice against any psychological account of personal identity. Johnston recommends a standpoint according to which psychological continuity and the like should be viewed merely as providing under normal circumstances relatively good evidence of personal identity,

<sup>11</sup> In addition to Olson, Animalists include Paul Snowdon, Peter van Inwagen, and David Wiggins. See Snowdon (1990), van Inwagen (1990), Wiggins (1980), and Wiggins (2001). Note that Wiggins’s Animalism is different in its details from, for example, Olson’s. Wiggins views the concept “person” itself as a substance-concept and indeed a natural kind concept, where the kind in question is precisely “human animal”. Nevertheless, he has been prone to suggest, obscurely, that the psychological nature of persons might form part of the essence of this kind, and so have some bearing on the individuation of persons, thus diminishing slightly the distance between his form of Animalism, and a Psychological view of personal identity, but creating a degree of distance between identity-criteria for this kind, and for other kinds, of animal. But as compared to his view in Wiggins (1980), in Wiggins (2001) this distance from other Animalists seems to have been lessened.

<sup>12</sup> It should be emphasized that, as I understand it, this is no implication of substance-ontology per se, though it accords well with the pre-eminent status typically accorded by substance-theorists to natural kinds.



much as do fingerprints or physiognomy. It would be as serious an error to take normally reliable psychological criteria as *constituting* personal identity, as it would be to take sameness-of-fingerprints as constituting personal identity.

Suppose we ask: could a machine, by reading the information from the brain of a human person and depositing it in another brain (real or synthetic), thereby bring about a continuation of the person's mental life and so provide a means of continuing the original person in existence in another body? Johnston suggests an attitude to this idea, much like that which an Animalist might take. From such a standpoint, the suggestion would be as bizarre and loopy as the idea that a machine duplicating fingerprints, or a plastic surgeon reproducing physiognomy, could accomplish the same effect. The most charitable view we could take of such ideas about fingerprints or physiognomy would be that they depended on radically mistaken overgeneralizations. Reflection on what is actually *done by* a fingerprint-duplicating machine, or a plastic surgeon, should be enough to make the mistake plain, since *they disturb the very regularities which make this evidence normally reliable*. Similarly, Johnston suggests, with the information-transfer machine.

Johnston's views thus run along similar lines to those of the Animalist, however they diverge at a crucial point. His views are, (i) that the relationship between a human animal, and a Human Being (in his technical sense), is not the relationship of identity, but the relationship of constitution, and (ii) that a Human Being could *in principle* outlive a human animal which formerly constituted it, though this possibility is one which does not get actualized in real life as we know it.<sup>13</sup> Indeed we may pretty much say that according to the Johnston of Johnston (1987), every actual individual Human Being *is* a human animal—*provided* that we understand this "is" as the "is" of constitution.

#### PROJECTIVE ERRORS AND THE REJECTION OF THE METHOD OF CASES

Johnston does not think that the error of those who misguidedly understand as *constitutive* criteria what are merely normally reliable *evidential*

<sup>13</sup> Setting aside, I would presume, various gruesome and (literally) short-lived borderline cases involving beheadings.

criteria for personal persistence arises simply from egregious overgeneralization. The relevant overgeneralizations gain spurious plausibility, Johnston thinks, from our tendency to commit what he calls “projective errors”. We are susceptible to a broadly Cartesian view of the self as a “bare locus” of consciousness: an enduring entity, distinct from anything physical, distinct too from any particular mental events, but underlying, sustaining, and unifying such events, simply by being a constant that-to-which-psychological-presentations-are-presented. No amount of physical continuity, nor even of psychological continuity, is absolutely necessary to the continued existence of such an entity. It is this view of selves as “soul pellets”<sup>14</sup> which Johnston describes as a form of “projective error”:<sup>15</sup> we are unaware of the physical underpinnings and complex processes which give rise to episodes of consciousness, and fall into the trap of projecting our lack of awareness into a lack in reality, taking it that there are no such underpinnings and complexities.

Johnston suggests that even those philosophers whose views about personal identity revolve around the explicit rejection of any such account of the self, nevertheless take seriously intuitions which amount only to silly overgeneralization, in part because they retain an unconscious residue of the soul pellets view.<sup>16</sup> For these and other reasons, Johnston proposes to reject the “method of cases”, which endeavors to extract criteria for personal identity from, principally, our intuitions about real and imaginary examples of personhood and its vicissitudes.

#### REIDENTIFICATION, MENTALISTIC SURVIVAL, AND THE BRAIN AS THE NATURAL ORGAN OF MENTATION

Johnston does not totally renounce all appeal to intuitions about imaginary cases. But he advocates a strict methodological regime according to which only those intuitions which cannot be traced to misleading

<sup>14</sup> Cf. Johnston (1989a: 372).

<sup>15</sup> Cf. Johnston (1992: 593).

<sup>16</sup> It’s worth noting here that Locke’s inauguration of the modern philosophical discussion of personal identity, used, in the specific context of an argument supposed to show the irrelevance of sameness of soul (or of organism) to personal identity, the example of the *soul* of a prince coming to inhabit the body of a cobbler. In other words, Locke might be accused of appealing to intuitions grounded in precisely the view he was attacking.

origins like those he identifies can lend support to an account of personal identity. Johnston's alternative proposal instead rests centrally on the observation that we normally identify and reidentify human persons relatively unproblematically, and proceeds by asking what kinds of entities *could* be unproblematically identified and reidentified in normal circumstances by the practices we habitually use.

Given this methodology, we might have expected Johnston to opt for Animalism, but in fact he opts for the Human Beings view. His argument includes two crucial steps. *First*, he asserts that "*if anything deserves the name of a conceptual truth about the relation between persons and minds, it is the claim that a person cannot be outlived by (what once was) his own mind.*"<sup>17</sup> (I'll call this claim MS, for "mentalistic survival": it says *survival of the mind suffices for survival of the person.*) *Secondly*, Johnston urges "*a properly naturalistic view of mental functioning*" as the characteristic functioning of a particular organ, the brain, claiming in this connection that "*talk of a mind is overly reified talk of an aspect of some minded thing.*"<sup>18</sup>

These are the two crucial steps from which Johnston concludes that we should view ourselves as Human Beings, entities normally constituted by a human animal, but such that in extraordinary (and so far imaginary) cases such as brain transplants, or brains-in-vats scenarios, they might be constituted merely by a living human brain, or by a creature resulting from transplant of such a brain. On this view, the death not of a human *animal*, but of a human *brain*, constitutes the death of a human person.<sup>19</sup>

How, we may wonder, do these two steps mesh with Johnston's proposed alternative methodology? Even if MS and the conclusion he derives from it strike us intuitively as plausible, by Johnston's lights that in itself counts for little. Is his appeal to this principle really an improvement on a direct appeal to intuition? This is probably the point

<sup>17</sup> Johnston (1987: 77). Olson cites this remark (Olson 1997: 12) in a context which suggests he takes it to show that Johnston supports a Psychological account of personal identity—a debatable reading which seems to me to take the remark very much out of context. Johnston's account is "psychological" in what is at best a stretched sense, and he explicitly opposes typical psychological theories. Cf. n. 29, below.

<sup>18</sup> Johnston (1987: 78).

<sup>19</sup> In discussion, it is often suggested that Johnston's view here amounts to the view that a person is identical with their brain. This claim, which Johnston rebuts, certainly does not follow from the view that a person is *in certain extraordinary circumstances* to be traced by tracing their brain. The fact that my car could survive destruction of its doors does not mean that they are not really part of it!

where Johnston's argument seems most to resemble traditional appeals to conceptual analysis. Does he merely capitulate to the method of cases at this point?

#### MENTALISTIC SURVIVAL, EXPERIENTIAL MEMORY, AND NON-MENTALISTIC INDIVIDUATION

Johnston does attempt to tie his appeal to MS to his proposed alternative methodology. The link is meant to be the role of experiential memory in self-identification. To see how, we must examine MS and the use Johnston makes of it. To this end, let's first compare MS directly with comparable claims about body organs.<sup>20</sup>

The claim "a person cannot be outlived by (what once was) his own *heart*" is false: we are familiar with cases in which the heart remains alive and is transplanted into the body of another person, after the death of its previous owner. Does the same possibility exist, in principle, for the *brain*? Johnston urges, plausibly to my mind, that we could not say the human animal as such would survive if the brain alone were kept alive and transplanted.<sup>21</sup> So a human animal can be outlived by what once was its own brain. Now if the person were (identical with) the human animal, it follows that the person could likewise be outlived by what once was its own brain. The denial of this possibility, conversely, is the upshot of Johnston's argument in Johnston (1987). He does not deny the surgical possibility, but claims rather that survival of the transplanted brain *would* be sufficient for survival of the original person.

There is a peculiar subtlety involved in differentiating Johnston's view from Olson's particular version of Animalism, since Olson argues—contrary to the claim I made above—that an animal cannot survive without a brain, nor vice versa, taking the view that where the living animal brain goes, there the animal goes. So Olson does hold that an animal could in principle survive being reduced to the condition of a brain removed from the skull which once housed it, and kept alive

<sup>20</sup> A full discussion of whether MS is true or a priori would require a detailed discussion of what might or might not count as "reidentifying a mind". We may reasonably dispense here with such discussion because Johnston's next move substitutes, rightly or (as I think) wrongly, the issue of reidentifying a brain.

<sup>21</sup> But note that Olson disagrees on precisely this point. See Olson (1997: ch. 6).

through artificial means.<sup>22</sup> It may appear therefore that whereas Johnston takes persistence for Human Beings, on his account, to *differ* from persistence for human animals, according to Olson's account it precisely *coincides* with the latter!

This does not threaten the generic difference between the Human Beings view and Animalism, since Animalists might have other views than Olson's as to necessary and sufficient conditions for animal survival. Furthermore, though Johnston, as we shall see, associates human mental life with the human brain, Olson's view is that it is not the whole brain, but the cerebrum, which is "that organ which is most directly responsible for your higher mental capacities such as reasoning and memory".<sup>23</sup> If Olson is right about this, then Johnston should say "*cerebrum*" wherever he in fact says "*brain*". In addition, Olson does not consider it necessary that the human animal, to survive, should retain a functioning cerebrum—the remainder of the brain, if it is alive and performing its normal functions, will do. In other words, for Olson, the cerebrum is an *inessential* part of the brain qua organ essential to survival. Thus the appearance of coincidence between Olson's and Johnston's views disappears once we take these subtleties into account: the part of the human animal essential for survival of Johnston's Human Being is in fact disjoint from the part essential, according to Olson, for survival of the animal itself.

Finally, and most importantly, this coincidence would at best be only an extensional equivalence. Even were Johnston's reasoning and Olson's to arrive at exactly the same conclusion as to what it minimally takes for a human person to survive—as it might be, continued survival and functioning of the brain *as a whole*—their views would remain intensionally different, since they differ as to which functions of the brain would earn it this crucial status.

This is particularly relevant in the light of the fact that Johnston views the relation of human animal to Human Being as one of constitution, and in addition takes the view that constitution always excludes identity. It seems then that even if human animal and Human Being always exactly coincided, due to the brain's functional versatility, Johnston would be required by his own lights always to see them as distinct, whereas Olson would see the allegedly distinct substance-category of Human Beings as entirely spurious.

<sup>22</sup> See especially Olson (1997: 131–42).

<sup>23</sup> Olson (1997: 9).

Returning to Johnston's main line of argument: he derives the claim that a person *can* survive through survival of their brain alone, from MS: the claim that it is a priori that survival of *mind* is sufficient for survival of person. And the second main step of his argument proceeds, essentially, by claiming that the truth of MS is preserved (though not, presumably, its a priori status!) if the word "brain" is substituted for "mind". He says in support of this: "A human mind is just a mode of functioning of a natural unit (e.g. a human organism or a human brain) whose conditions of persistence are storable in nonmental terms. This is the sense in which talk of a mind is overly reified talk of an aspect of some minded thing."<sup>24</sup>

Note how an implicit commitment to a restrictive ontology, most plausibly some kind of substance-ontology, emerges in the reference to *natural* units whose conditions of persistence are storable in *nonmental* terms. This is a pivotal point in Johnston's argument, and one at which his general ontological presuppositions come into view. Given the aim of discovering what kind of thing it is which we are most readily seen as unproblematically identifying and reidentifying in our everyday person-identificatory practices, it transpires that the candidate kinds are to be "natural units" such as organisms or body organs, or composite entities whose persistence-conditions may be formulated, *in principle non-mentally*, in terms of such natural units. Mentalistic persistence-conditions as such are ruled out.

Here a major general moral emerges. The results yielded by Johnston's methodology will inevitably be relative to the options the world is taken to serve up to us in the way of candidate kinds for our identificatory practices to be seen as tracking. Since different ontologies will take different stands on this, there is a danger that the choice of a particular ontology already tendentially limits the outcome of the inquiry before any appeal to identificatory practices can enter the picture. I believe this actually happens in the case of the answer Johnston supplies. Indeed I believe that a central weakness of Johnston's preferred methodology is its inevitable need for some pre-established restriction and ranking of the kinds of entities which are to be seen as candidate targets of our identificatory practices. The more general moral, however, extends beyond the context of Johnston's own methodology.

<sup>24</sup> Johnston (1987: 80).

Consider for instance the defense of conceptual analysis mounted in Frank Jackson's recent Jackson (1998). Johnston's and Jackson's methodological proposals share a common structure. Faced with the metaphysical question "What are Fs?" each in effect proposes a test to be applied to the members of some favored basic ontological pool. Those members of the pool which best pass the test, thereby qualify to be counted as the Fs. Jackson's test may be thought of as generated by Ramsifying an implicit term-defining "theory" of Fs, taken as revealed piecemeal by conceptual analysis; Johnston's (in the specific case where the Fs are the *persons*), as generated by asking what kind of thing it is which we are most readily seen as unproblematically identifying and reidentifying in our everyday person-identificatory practices. Thus Jackson and Johnston fish with different nets. But what entities you can fish out of an ontological pool depends not merely on what net you use, but on what the pool contains in the first place. If mentalistically individuated entities are not to be found in it, no net will fish them out. In other words, any metaphysical methodology which has this broad structure, identifying the Fs by applying a net or filter to some predesignated ontological pool, will only be as good as the reasons available for preferring that pool to begin with. A central weakness of Johnston (1987) is the lack of any overt argument in favor of the preference, albeit a fashionable one, for a pool which contains biological and other natural kinds, but which excludes mentalistically individuated substances *per se*.

Johnston's choice here (to allow only non-mentalistically individuated entities as *bona fide* substances) is not altogether arbitrary or unmotivated. He claims that the method of cases fails in this context, in part because "the" concept of a person, taken as what is in common to all the different conceptions of personhood which have historically been held or which may emerge from reflection on imaginary cases and the like, is impossibly vacuous. Nothing but a soul pellet could fit it, yet a capacity unproblematically to reidentify soul pellets would be impossible to account for; nor would it make sense to attach the kind of importance to the vicissitudes of soul pellets which we in fact attach to issues of personal identity. Thus it is natural for him to assume that his alternative method must to some extent rely on *nature*, rather than on *our concepts*, to demarcate the persons: to assume, in other words, that the kind "Human Person" must be constructible relatively straightforwardly

out of natural kinds. And in this context, as noted, Johnston identifies the relevant natural kinds as the biological kinds.<sup>25</sup>

Nevertheless, it would be possible to meet the demand that the persons be *in some broad sense* “naturally” individuated, and yet provide a greater range of options in seeking a kind to match with our ordinary person-identificatory practices. Discounting hopelessly gerrymandered or arbitrarily or extrinsically individuated “kinds” need not oblige us to reject *mentalist* or *psychological* persistence-conditions as Johnston does.

Johnston refers to “the naturalistic framework taken for granted by scientifically validated common sense”: but I doubt that respect for scientifically validated common sense demands that we reject the possibility of mentalistic persistence-conditions. It is admittedly a matter of considerable philosophical debate how to integrate the mental into a naturalistic world view: but scientifically validated common sense is happy to admit all sorts of entities which are not, but which supervene on, those which figure explicitly in actual science. There is for instance no science of bricks *as such*, but given that bricks supervene on molecules, and are relatively naturally individuated, scientifically validated common sense has nothing to say against them. Similarly, I can commonsensically count the ripples on a pond and see how far they travel, with a clear scientific conscience, even if I acknowledge that ripples cannot exist independently of the quantities of water which successively, by their causally interrelated movements, constitute them. It’s just not true that every bona fide materially constituted entity recognized by the union of common sense and science, and which persists and can be reidentified over time, counts as an independent substance which meets the stringent criteria laid down for substance-concepts by writers such as Wiggins, or which belongs to the biological natural kinds with which they are so taken.

In Johnston (1987) Johnston says nothing which specifically addresses this issue. Since Locke,<sup>26</sup> though, we should know better. Locke’s most profound and original contribution, I believe, in his discussion of personal identity, is not his proposal of a psychological criterion for

<sup>25</sup> Johnston does acknowledge the need for further argument in support of his prioritizing of non-mentalistically individuated natural units, as preferred targets for our reidentificatory practices, in his n. 19 to Johnston (1987).

<sup>26</sup> All references are to Locke (1961, Book II, ch. 27).



personal persistence, but something which is logically prior to that proposal—his recognition of the fact that, given an ontology of fundamental natural kinds (even including “supernatural” kinds such as spirits), we may also recognize the existence of *additional* well-individuated substance-kinds, distinct from more fundamental and autonomous kinds, but at all times supervening on them, because constituted out of them. This is the lesson he draws from consideration of the case of living things, and applies, in a brilliant and original analogy, to the case of persons. It is a lesson of which materialists should stay specially cognizant, since the alternative historically was to believe in an entire category of fundamental immaterial substances, “souls” or “spirits”, having their own primitive and unanalyzable persistence-conditions, and somehow guaranteed by their intrinsic nomic essence to sustain mental life and to individuate mental lives.

Locke in many ways fits the model of a substance-ontologist, attempting to argue for the cases of living things, and of persons, alike, that persistence-conditions for things of the relevant kind should be deducible from a careful statement of what it takes merely to qualify as a thing of that kind, and in effect arguing that in each case it is the persistence of an appropriate kind of *process* which is essentially constitutive of persistence for the relevant entities. Once again we see that the general form of a substance-ontology by no means automatically dictates the kind of prioritizing of biological kinds at the expense of mentalistic individuation, which Johnston must assume if his argument for the Human Beings view is to stand up.

Needless to say, there is no way fully to adjudicate the issues here touched on without a thorough discussion of the constitution relation, identity over time, and a host of other controversial topics in ontology.<sup>27</sup> Similarly, it would require a lengthy digression properly to defend my take on Locke. For my purposes it is sufficient to point out that Johnston is relying on assumptions which are by no means inevitable, which he does not defend, and which crucially determine the outcome of his argument.

<sup>27</sup> For instance, I am steering clear of the issue of four-dimensionalism which, as it happens, I endorse. Given unrestricted temporal partition and unrestricted mereological fusion, there will be no special problem about mentalistically individuated entities. But Johnston, as noted, in Johnston (1987) employs the language of substance-ontology. It is therefore dialectically more useful to point out that substance-ontology *per se* does not rule out mentalistic individuation, than to invoke the kind of apparatus—temporal parts and mereological fusions—which substance-theorists characteristically reject.

Let's return to that argument. It is not just outright Psychological Reductionism which Johnston renounces in renouncing mentalistic criteria for individuation or persistence, and the correspondingly richer set of kinds which might otherwise be regarded as eligible candidates for the title "Human Persons". Consider the intermediate or hybrid position I call the *Human Animal Stage* view. On this view, even if a person cannot be outlived by their own *mind*, their mind *can* be outlived by their body and brain. It holds that a person continues to exist so long as their mental functioning is substantively unimpaired, but that a person who suffers total permanent and radical amnesia and personality loss thereby ceases to exist. Subsequent development of new memories and personality will, on the Human Animal Stage view, constitute the coming into being of a new person, not the continued or renewed existence of the previous person.

Animalism differs from the Human Animal Stage view in denying that a person ceases to exist at such a moment of mental disintegration and amnesia. A Human Animal Stage theorist will hold that a single human animal may in such circumstances successively constitute two different people. But Johnston seems precluded from saying that radical psychological discontinuity would necessarily count as terminating a person's existence, or inaugurating the existence of another person, and hence (though in fact he does not consider such a view in Johnston (1987)) seems committed to rejecting the Human Animal Stage view.

A contrary impression may be given by the fact that Johnston tends to say merely that on his view a person *may* survive, *albeit in a mutilated form*, so long as their brain survives, and at one point in his introductory remarks he suggests a qualification, saying that "a human being could be reduced to the condition of a mere brain *so long as* that brain continues the human being's mental life".<sup>28</sup> In a similar vein, he says " 'human being' names a partly psychological kind", given that "the tracing of a human being gives primary importance to mental functioning among the various life functions exhibited by human beings".<sup>29</sup>

But Johnston repeatedly insists on giving the brain the basic individuating role when tracing Human Beings, as when he says "it is crucial

<sup>28</sup> Johnston (1987: 64). My italics.

<sup>29</sup> Johnston (1987: 79). Johnston's words at this point are reminiscent of Wiggins's version of Animalism.

to the tracing of a human being that there be something that is the continued functioning of that human being's brain."<sup>30</sup> In the light of this remark, it is clear that he cannot accept the Human Animal Stage view. The case raised above, in which the Human Animal Stage theorist would see a single human animal as successively constituting distinct people due to loss of mental continuity and connectedness, is one in which the brain does continue to function, hence by Johnston's lights must be seen as a case in which a single Human Being persists. Thus, Johnston should be seen as denying the necessity, as well as the sufficiency, of any sort of trans-temporal mental connections, for the survival of a Human Being. If mentalistically individuated entities as such are to be ruled out, we cannot make it any sort of condition for the survival of a Human Being *or* of its brain that the brain subserve any particular kind or pattern of mental functioning. The logical destination of Johnston's discussion is the unqualified claim that a person simply does survive just so long as their brain survives.

If this is the logical destination of Johnston's discussion there is something rather odd about his route to it. Let me explain.

Given his declared methodological assumptions, and his evident ontological assumptions, Animalism is and ought to be Johnston's natural starting point. The biological kind that at least our third-person, person-identificatory practices are most readily seen as tracking is the kind: Human Animals. But Johnston notes that we also habitually and unproblematically identify ourselves by the use of experiential memory. The deliverances of experiential memory, he says, are deliverances, true or false, about personal identity as such. As a fact of phenomenology, experiential memory always represents remembered experiences as experiences had by oneself: experiential memory does not deliver some "halfway house" judgment short of a belief about personal identity. Furthermore, its seeming aptness for delivering knowledge about personal identity would seem to survive even a Cartesian suspension of belief in one's body and in anything external to one's own mind.<sup>31</sup>

But if on the one hand we are already committed to finding a variant on the Human Animals view, in the sense of taking a *non*-mentally individuated biological kind of entity, or some composite constructed out of such entities, as the target of our ordinary modes of reidentification of

<sup>30</sup> Johnston (1987: 79).

<sup>31</sup> See Johnston (1987: 77).

persons, whereas on the other hand experiential memory seems suited by its phenomenology to detecting merely *mental* connections, we seem in danger of having to conclude that it is after all merely minds, *not* persons, which experiential memory is best seen as unproblematically identifying and reidentifying. It *is* specially suited to detecting *mental* connections: it is *not* specially suited to detecting *physical* or *biological* connections.

Hence Johnston's appeal to MS. He needs this principle to rule out such an option. Any case of reidentifying a mind must, necessarily and a fortiori, be a case of reidentifying a person. And finally—to complete the resolution of this impasse—the putative substance-kind “mind” is revoked, and the organic kind “brain” is substituted for it.

In short, minds enter the picture because of the relative *unsuitability* of experiential memory for the unproblematic identification of an organic entity such as a human *animal*. But once talk of a mind has done its job, the excessive reification is cancelled by reverting again to talk of an organic entity, but a different one—a human *brain*. The puzzle is that experiential memory has already been declared to be relatively unsuited—phenomenologically, at least—to reidentifying any such kind of thing unproblematically! I smell a rat.

In fact, it seems to me, Johnston's appeal to the role and peculiarities of experiential memory simply has the effect (unintended I am sure) of an exercise in misdirection. For all Johnston says, I believe, we might reasonably conclude that the things which can be ordinarily and unproblematically reidentified over time, in just the way in which we reliably and unproblematically reidentify ourselves and each other over time, are human animals. Johnston's preferred view gives a special status to the human brain, which provides the indispensable essence of a Human Being in his sense. I think it is important to see why his discussion of the role and character of experiential memory does not and cannot lend support to that conclusion. This is what I next attempt to show.

*What is “experiential memory” best suited, phenomenologically, to reidentifying unproblematically?*

Let's begin with the “first-person” phenomenology of experiential memory, and the claim that it represents an experience not simply as having happened, but as having been experienced by oneself. This

appeal to phenomenology seems to me at best a complete red herring. There simply is no logical room for an alternative phenomenology here: it follows just from the fact that *experience* is what experiential memory is memory of. One cannot recall an experience without recalling it as if it were one's own experience, because one cannot *have* an experience without having it as if it were one's own experience. A form of memory which informed one barely that some experience had occurred, without recreating that aspect of the phenomenology of experience, and *thereby* representing it as an experience had by oneself, would not be what we call "experiential memory". The phenomenology comes with the job description. What ought to determine the suitability of experiential memory as a normally reliable form of reidentification of a mind or of a self is something like the reliability of the information it delivers, interpreted as information about a mind or a self: not the form that information is packaged in. At least, not if the packaging is inevitable.

We can go beyond this criticism: further complexities arise concerning the phenomenology of experiential memory, obscure in their implications, relevant to Johnston's claims, but none of them supportive of his conclusions. At the very least, they highlight the lack of clarity of the precise claim which Johnston is making here.

For one thing, experiential memory very frequently provides us with information about the past states and behaviors of our *bodies*: the sleazy catch-phrase "let's go build some memories" suffices to remind us of this. Indeed, one might say that experiential memory frequently has a phenomenology which seems peculiarly well suited to *the reidentification of a body as such*. (Compare Descartes' Sixth Meditation remarks on the peculiar suitability of the faculty of imagination for providing us with information about the interactions of an *embodied* being with its environment.) Phenomenologically speaking, experiential memory often stops no more short of a judgment about the identity of a body than it does about the identity of a person (in contrast, be it noted, to the fact that it typically delivers nothing in particular in the way of judgments about the identity of *brains*).

For another thing, there is what I shall christen "the displaced viewpoint" effect. It may readily be confirmed by cross-examination that for many people, memory and imagination are alike in that remembering or imagining themselves performing some action often involves imagery of the action as seen from a displaced point of view. In remembering some particular occasion of going swimming, for instance, one may find

that what comes into one's mind is an image of oneself, seen from above, swimming in the water, rather than an image of the surface of the water as seen from an inch or two away, that is, from the point of view of the swimmer.

Now when one's memory takes such a form, it remains true that there is a particular first-person standpoint associated with the image: the remembered scene is viewed from a particular viewpoint. It could hardly be otherwise. The trouble with the displaced viewpoint effect, for Johnston's purposes, is not that in such cases experiential memory does not have an associated first-person standpoint, but that this standpoint is *not that of the person whose experience is being remembered!*

So it all depends on what we mean by "phenomenology". I may on one occasion remember seeing someone else swimming, actually having viewed them from above; or I might remember an occasion on which it was I who was swimming, but experience the displaced viewpoint effect, forming in my mind a memory image which duplicates that of the previous case. Typically, I will know perfectly well whether it is my own swimming, or my viewing of another's, which I am remembering. But I will not know this on the basis of the memory images per se. We have all learnt, after all, the Wittgensteinian lesson that imagistic accounts of thought are inadequate, since the content of such thoughts will depend crucially on the interpretation of the images, something which they themselves, qua images, cannot supply.

Perhaps, then, the proper way to understand "phenomenology" here is to take it as including, crucially, not just the character of memory images, but along with that, the *interpretation* of those memory images. Even so, given that memory images cannot supply their own interpretation, it's quite imaginable, whether or not it actually happens, that a person's experiential memory should be, in the following way, a little hazy: on a given occasion one might be confident enough one is remembering an actual event, yet, for all that, be in doubt whether the swimming figure in the memory image represents oneself, as in the displaced viewpoint effect, or someone else one actually watched swimming. Viewed in this light it's hard to see just how the phenomenology of experiential memory *does* fit it particularly well for the role of reidentification of selves or minds.

Perhaps this talk of memory images—or more generally, the idea that what is crucial here is recall of past experiences—is out of place. Despite his talk of phenomenology, Johnston himself says that the deliverances

of experiential memory are *propositions*<sup>32</sup> involving claims about personal identity. Perhaps his talk of “phenomenology” here is no more than dressed up talk of beliefs or judgments, and the claim is simply that usually when we remember some event, we remember that we were this participant rather than that? But this interpretation makes Johnston’s talk of “experiential memory” and “phenomenology” inappropriate—I *might* find myself believing some proposition about my past actions in the absence of *any* particular phenomenology. Worse, it’s hard to see how this claim on its own can carry the significance Johnston is seeking here.

At any rate, to repeat, it surely is not the *phenomenology* of experiential memory which makes it specially suitable for the reidentification of this, that, or the other, thing: it is the *reliability* with which its deliverances correlate with the facts concerning this, that, or the other, thing, which should be relevant here. Furthermore, as noted, Johnston himself says that what the phenomenology of experiential memory makes it specially well-suited for establishing is the existence of *psychological connections per se*—though he promptly transmogrifies this into the claim that it is only a mind *per se* which it is phenomenologically specially well-suited to reidentifying. And even this transmogrified claim cannot survive his subsequent, naturalistically motivated substitution of “brain” for “mind”. (This is arguable, it seems to me, even if we set aside any merely generic objection to apparent substitutions within intentional contexts, which is what claims to know identity, whether of mind or of brain, surely are.)

*The contingency of experiential memory as a means of self or brain reidentification*

In the light of these considerations, and of Johnston’s own insistence that we should not too hastily promote normally reliable evidential criteria into metaphysically essential constitutive principles, let’s reflect further on the *contingency* of the fact that this form of memory normally gives us reliable information as to who we are, or indeed of any of a number of related facts. If someone is having a putative experiential memory, there are a number of questions which may be raised about it.

<sup>32</sup> Johnston (1987: 77). My italics.

- a) Did an experience such as that which is apparently being remembered ever *actually occur*, and *cause* the apparent memory?  
(A putative experiential memory cannot be genuine unless it sufficiently matches some experience, in the sense of providing sufficient true information about the character of some real experience.)
- b) Supposing (a) answered in the affirmative, did the experience which sufficiently matches the memory, and which had a role in causing it, have the *right kind of role* in causing it, for the putative experiential memory to count as genuine?<sup>33</sup>
- c) Did the same *person* have the original experience, as is now having the experiential memory?
- d) Did the same *human animal* have the original experience, as is now having the experiential memory?
- e) Was the original experience realized by states of the same *brain* as the brain whose states are now realizing the having of the experiential memory?

How these questions are related to one another will be a matter for dispute between those who hold differing accounts of personal identity. But I claim that it ought to be agreed *by all parties* that in any given case of someone having a putative experiential memory, it is only contingent on certain in-principle-alterable facts about the world as we currently know it that *any* of them may be answered affirmatively. This ought already to be obvious from the fact that people are capable of psychotic delusions in which they imagine themselves to be, for instance, Napoleon. It follows again from the possibility in principle, exploited in such works of fiction as the movie *Total Recall*, starring Arnold Schwarzenegger, of technology capable of implanting false memories. Experiential memory indeed gives those of us who are not psychotic a highly reliable source of knowledge of who we are. But this is contingent on facts which could in principle change.

<sup>33</sup> If A has an experience and describes it, and B hears, reads, or encounters that description and is thereby caused to have an apparent memory matching A's original experience, this is not a case of genuine experiential memory, whether or not B is, as in the case of reading one's own old diary entries, identical with A. Which other causal connections would or would not be sufficient for genuine experiential memory must remain to some extent moot here, but clearly we must count the normal neurophysiological processes which underlie the formation of experiential memories in the human brain as sufficient.



Suppose we set aside the issue of psychotic delusions and other apparent memories which are false or fictitious in the sense implied by a negative answer to question (a). We might imagine, for instance, that we come to know that the world has been made free, through hand in hand advances in technology and social norms, from psychotic delusions, the degradations of stage hypnotists, spurious cases of “recovered memory syndrome”, and the like. Consistently with that assumption, a variety of further scenarios are possible.

*First case.* Suppose that it becomes technologically possible to perform brain transplants without damage to the normal processes by which memory-traces are laid down and preserved; and suppose that in other relevant respects not so far touched on, the world goes on much as it does at present. Then someone’s having a particular experiential memory *would* normally provide reliable evidence that the brain whose states realized the original experiences was the same *brain* as the one whose states were currently realizing those experiential memories. At the same time, it would (*pace* Olson’s views about animal persistence) provide somewhat less reliable evidence for identity of the relevant human organism if any. This is roughly in accordance with Johnston’s view, though of course the memories would not have the evidential status just described, for the particular person *having* those memories, *unless that* person was in possession of the relevant facts—unless they actually *knew about* the existence and functioning of brains, the possibility of transplanting them, and so forth. But we can imagine other scenarios.

*Second case.* Imagine instead that the technology of brain transplants stays in its current undeveloped state, but that the technology of prosthetic brain implants takes off. These implants, we might imagine, substitute modules relying on silicon chip technology for groups of brain cells, relevant information being copied from the original brain cells, stored temporarily, and then entered into the chip cluster with which those cells are replaced. Eventually an entire brain can be replaced in this manner. Thus degenerative brain diseases like Alzheimer’s may be defeated. Under such circumstances, experiential memory would provide a *more* reliable indicator of identity of human organism than of identity of brain: though once again, only those in the know about the relevant facts would be able to *employ* it as such an indicator. (Olson, once again, would dissent: his view is that under these circumstances the human organism would no longer exist.)

*Third case.* Imagine, finally, that technology was created capable of recording information from one person's nervous system and implanting it in the form of experiential memories in another's. It would then be entirely in order to accept the veracity of an experiential memory in the sense of believing it to be faithful to some actual experience—always to answer question (a) in the affirmative—whilst at the same time doubting or wondering whether it was indeed one of *one's own* experiences. If use of such technology was rife, the easy and uncomplicated ways in which we use experiential memory in tracing people would no longer be adapted to tracing things of that kind, and this would equally be the case whether one held the Human Animals view, Johnston's own Human Beings view, or even—it is important to note—some reasonably sophisticated version of Psychological Reductionism. An affirmative answer to question (a), even an affirmative answer based on experiential memory, would provide no reliable basis for answering questions (c), (d), and (e)—questions about identity, respectively, of experiencing person, of experiencing animal, and of experiencing brain.

In fact, of course, none of these possibilities obtains. As things stand, the actual role of experiential memory in unproblematically reidentifying persons accords well enough with both the Human Animals and Human Beings views of what a person is, but on each of these views its reliability in this role must be seen as contingent and a posteriori rather than as necessary or a priori. There are scenarios whose impossibility is merely technical—that is to say, scenarios which are not even nomically impossible, let alone metaphysically impossible—in which the special suitability of experiential memory for reidentifying either, or both, of these kinds of entities would be undermined. Furthermore, as things stand, experiential memory is also equally good for reidentifying persons conceived as the Psychological Reductionist conceives them!

#### CONCLUDING REMARKS

The upshot is that, as I see it, there is no way to combine the claims

- (i) that *experiential memory is specially well-fitted to identifying minds;*
- (ii) that *as a matter of a priori necessity, minds can't outlive the persons whose minds they are;* and

- (iii) that *the natural unit whose functioning underlies the phenomena overly reified as "minds" is the brain*, to arrive at the conclusion
- (iv) that *experiential memory is specially well-fitted to identifying persons, conceived as entities whose persistence-conditions are derivative on those of brains, as opposed to persons conceived as human animals*.

Claims (i) and (ii), if true at all,<sup>34</sup> are true only of minds qua minds: substitute "brains" for "minds" and any special suitability of experiential memory for the job of reidentification is undermined, and with it, any specially compelling reason based on that suitability for abandoning the Human Animals view in favor of the Human Beings view.

Experiential memory is, as (i) asserts, indeed well-suited *as things actually are* for the unproblematic reidentification of minds. Even this, I believe, is quite contingent. Still, if any of these kinds of entities—minds, human brains, human animals, Human Beings—may be seen as distinguished by the special suitability of experiential memory for unproblematically reidentifying its members, it will be the category of minds *as such*. And if as good materialists or naturalists, we really want to avoid reifying minds, we should here introduce as our surrogate not the materialistic category of brains (nor cerebra!), but the category of *mentally individuated persons*, conceived as Psychological Reductionists conceive them: possibly discontinuous physical entities which realize various mental states at various times, but such that those mental states collectively exhibit both synchronic and diachronic unity of a kind and degree appropriate to personhood.

Now of course such entities cannot be acknowledged if mentalistically individuated entities are ruled ineligible at the outset. Nor can they be acknowledged if one assumes a background ontology which is in some other way too restrictive. In principle such entities could coincide with mere temporal parts of human animals; numbers of them could be alternately and discontinuously constituted—or continuously and multiply constituted—by human animals, as in Jekyll and Hyde and other multiple personality cases; they could undergo fission and fusion; they could be discontinuously constituted first by some human animal, later by some android; and so forth. I've never found a convincing

<sup>34</sup> In what sense (ii)—which is to say, MS—is true, if any, is an interesting question, but one which space does not permit me to examine further here.

argument purporting to establish that such entities cannot be acknowledged consistently with the main tenets of substance ontology, provided that mentalistic individuation, which Johnston baulks at, is seen as acceptable. Nevertheless, substance ontologists (in line with their tendency to extend exclusive special privileges to natural kinds) are typically reluctant to admit such entities into their ontologies (though four-dimensionalists would have no problem with them).

It has by no means been my purpose here to *defend* Animalism, even though I believe that Johnston's argument in favor of his Human Beings modification of it does not succeed. Rather, my aim has been to point out that Johnston's sole argument in favor of modifying the Human Animals view does not stand up, and in particular that Johnston's proposed methodological alternative to the method of cases provides no good grounds for preferring his "Human Beings" modification. First, his methodology does not actually rule out his principal target, Psychological Reductionism, without having to be bolstered by an unargued and, I believe, indefensible ban on mentalistic individuation, or some equivalent form of ontological conservatism. Without good arguments for such assumptions, which he does not supply, Johnston's argument is essentially question-begging. But secondly, *even if* some suitably conservative ontological perspective is adopted, Johnston's methodology does not actually favor the Human Beings view over the Human Animals view. The kind "human animals" is certainly a more natural kind than the kind "Human Beings" which, as he admits elsewhere,<sup>35</sup> could include a creature consisting of the living body of a tiger onto which a living human brain had been successfully transplanted.

It is noteworthy, finally, that Johnston should purport to derive from his natural-kinds-and-normally-reliable-evidence methodology a conclusion which it does not favor, though it is relatively strongly favored by the alternative methodology of appeal to normal intuitions, which Johnston rejects. Derek Parfit, for instance, says "suppose that my brain is transplanted into someone else's (brainless) body, and that the resulting person has my character and apparent memories of my life. Most of us would agree, after thought, that the resulting person is me."<sup>36</sup> I think it is easy to go along with Johnston's argument because of these intuitions, which I am perfectly happy to accept. The point is that Johnston himself, officially, has no right to them.

<sup>35</sup> Johnston (1989b: 456).

<sup>36</sup> Parfit (1971: 4).

In a nutshell, Johnston's alternative methodology asks the following questions. What sorts of characteristics do we take as providing good "fingerprints" for persons? And what kind of thing is it which such characteristics would indeed provide good fingerprints for? Literal fingerprints work well, as we all know. Experiential memory does too. But experiential memory provides a reliable fingerprint for a human brain—hence for any kind of entity of which a human brain is an essential part—only contingently. And under the contingencies which in fact obtain, it provides exactly as good a fingerprint for a human organism. In one respect, indeed, it provides a better one. One needs much less special knowledge—in particular, one needs much less knowledge of human anatomy—to be in a position to know of the general reliability of experiential memory as a fingerprint for human organisms, or to be in a position to know that one is dealing with an exceptional or unreliable case—than one needs in order to know of the general reliability of experiential memory as a fingerprint for human brains, or for entities whose nominal essence includes them.

The normal intuitions about brain transplants reflect, I believe, not a view about which kinds of fingerprints provide the most reliable means of identifying which kinds of entities, nor merely a residue of bygone philosophies or religions, but our sense that what we really *care* about in caring about personal survival is psychological unity, so that we prefer *where possible* to identify personal unity with psychological unity. I believe, but shall not argue here, that this is a preference which neither the world, nor our conceptual commitments, once shorn of false metaphysics, prevent us from indulging.

#### REFERENCES

- Jackson, Frank (1998) *From Metaphysics to Ethics: a Defence of Conceptual Analysis* (Oxford: Clarendon Press).
- Johnston, Mark (1987) "Human Beings", *The Journal of Philosophy* 84: 59–83.
- (1989a) "Fission and the Facts", *Philosophical Perspectives* 3: 369–97.
- (1989b) "Relativism and the Self", in *Relativism: Interpretation and Confrontation*, edited with an Introduction by Michael Krausz (Notre Dame, Indiana: University of Notre Dame Press).
- (1992a) "Constitution Is Not Identity", *Mind* 101: 89–105.
- (1992b) "Reasons and Reductionism", *The Philosophical Review* 101: 589–618.

- Locke, John (1961) *An Essay Concerning Human Understanding*, ed. J. Yolton (London: Dent).
- Noonan, Harold (1993) "Constitution Is Identity", *Mind* 102: 133–5.
- Oderberg, David (1989) "Johnston on Human Beings", *The Journal of Philosophy* 86: 137–41.
- Olson, Eric T. (1997) *The Human Animal: Personal Identity Without Psychology* (Oxford: Oxford University Press).
- Parfit, Derek (1971) "Personal Identity", *The Philosophical Review* 80: 3–27.
- Robinson, Denis (1982) "Re-Identifying Matter", *The Philosophical Review* 91: 317–41.
- Robinson, Denis (1985) "Can Amoebae Divide Without Multiplying?", *Australasian Journal of Philosophy* 63: 299–319.
- (2004) "Failing to Agree or Failing to Disagree?: Personal Identity Quasi-Relativism", *Monist* 87: 512–36.
- Snowdon, Paul (1990) "Persons, Animals, and Ourselves", in C. Gill (ed.) *The Human Person and the Human Mind* (Oxford: Clarendon).
- van Inwagen, Peter (1990) *Material Beings* (Ithaca: Cornell University Press).
- Wiggins, David (1980) *Sameness and Substance* (Oxford: Blackwell).
- (2001) *Sameness and Substance Renewed* (Cambridge: Cambridge University Press).