When Do Persons Die?

Indeterminacy, Death, and Referential Eligibility

Ben Curtis^{1,2}

Published online: 12 May 2018
© The Author(s) 2018

1 Introduction

The topic of this paper is the general thesis that the death of the human organism is what *constitutes* the death of a person. All admit that when the death of a human organism occurs, in some form or another, this normally does result in the death of a person. But, some maintain, organismic death is not the same thing as personal death. Why? Because, they maintain, despite the fact that persons are associated with a human organism ('their organism'), they are not identical with their organism, and so they can, at least in principle and perhaps in actuality, die before or after their organism dies.

The thesis described above I shall call 'the two-deaths thesis' in virtue of the fact that its proponents believe that there are two kinds of death, viz. organismic death and personal death. The opposing thesis, according to which the death of a human organism *does* constitute the death of a person, I call the 'one death thesis'. After giving some background remarks in Sect. 2 of this paper, in Sect. 3 I argue that, if we take linguistic usage seriously, neither the two-deaths thesis nor the one-death thesis is true. I argue, rather, that so far as linguistic usage goes it is indeterminate which of them is true. In Sect. 4 I go further and argue that if one accepts a widely held metaphysical view, viz. that the world itself partially determines reference and the extensions of our concepts, then it is plausible that the one-death thesis is true despite the facts about linguistic usage.

✉ Ben Curtis
benjamin.curtis@ntu.ac.uk

¹ Arts and Humanities, Nottingham Trent University, Nottingham, UK

² University of Johannesburg, Johannesburg, South Africa

2 Background Remarks

In this section, before I come to my main line of argument, I offer some background remarks on the two-deaths and one-death thesis. I focus on one particularly well-known defence of the two-deaths thesis, viz. that given by Jeff McMahan in his seminal paper ‘The Metaphysics of Brain Death’. Nothing in the main argument that I give will depend upon the details of McMahan’s particular version of the two-deaths thesis, and I use it solely as a representative of its ilk. I have chosen McMahan’s view as my representative because his paper is probably the most widely read defence of the two-deaths thesis, and except for specialists in the area, it is probably the only one with which most are familiar.

I start by making a few comments on the brief arguments against the one-death thesis that appear in the early sections of McMahan’s paper. I start with these because they are presented in a manner that makes the one-death thesis appear to be much weaker than it is, which is an impression I wish to dispel, and also because those arguments serve as a useful base upon which to lay the ground for my main argument.

If the identity view is true (i.e. the view that persons are identical with their organisms), then it seems that the one-death thesis must also be true, for how could one and the same thing die at different times from itself? So anyone who believes the two-death thesis is committed to denying the identity view. McMahan is well aware of this, and begins his defence of the two-deaths thesis by noting that there are various arguments against the identity view, and he briefly presents a few of them. His purpose in the early sections of his paper, I think, is not to offer anything like knock-down arguments against the identity view, but to offer some initial motivation for the two-deaths thesis, for his real concern in the paper is to develop a specific version of the two-deaths thesis given that the identity view is false (although the case he makes for the two-deaths thesis in the remainder of the paper can be also be viewed as being independent argument against the identity view). Nonetheless, the arguments he gives in the early sections of his paper are one-sided, because whilst it is true that there are various arguments against the identity view, as anyone familiar with the literature on personal identity knows, there are also various arguments against the *non*-identity view. Indeed, despite the fact that each view has able defenders, there is no general consensus about which view is more plausible.¹ So, in an attempt to redress the balance here, I devote a few paragraphs to giving some brief responses to the arguments that McMahan gives against the identity view, and I present just one argument that has at least as much strength as those that McMahan presents, but that tells against the non-identity view instead.

The first of McMahan’s brief arguments goes like this: If we are identical with our organisms, then we come into existence when our organisms do, but this means that we once were single-cell organisms (i.e. zygotes). But ‘many of us find it hard to believe’ that this is true, so we ought to reject the claim that we are identical with our organisms.² This, I think, is a weak argument. To see why consider the sense

¹ McMahan himself is an able defender of the non-identity view. See chapter 1 of Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002).

² Jeff McMahan, “The Metaphysics of Brain Death,” *Bioethics* 9 (2) (1995): 91–126.

that the phrase ‘many of us must find it hard to believe’ must have in order for the argument to be valid. If it just means that many of us find ourselves in a conflicting doxastic state about the claim that we were once zygotes, and so unable to positively believe it for this reason, then the argument is invalid. I suffer a conflicting doxastic state about the claim that my student’s illness caused him to hand in his assignment late (he seems to be trustworthy but his illnesses seem to be suspiciously correlated with hand-in dates), but I would be rash if I was moved by this to reject his claim outright. So, the sense in which many of us must ‘find it hard to believe’ the claim is the sense in which we must find it to be wholly implausible. But is it true that many of us find the claim to be wholly implausible? There may be a fair few philosophers who find it to be so, but I wager many of them find it to be so due to their other philosophical commitments, and not because of some strong belief that arises when considering the claim impartially. It is an empirical question whether many of ‘the folk’ would find it to be wholly implausible (and if the argument is to have any weight, I would think that a large proportion would have to), but if my beliefs in this regard are anything to go by I doubt they would, for I not only find the claim to be not implausible when considered impartially, but I actually find it somewhat plausible. There seems to me to be nothing *obviously* wrong with the claim that we are biological organisms and so were once zygotes. There appears to be no contradiction involved in the claim, and it seems to reflect some aspects of our ordinary thought about ourselves. However, I should also say that I also find the negation of this claim to be somewhat plausible. Again, there seems to be nothing *obviously* wrong with the claim that we are psychological beings who come into existence in utero when the foetus develops sufficiently advanced cognitive capacities. Again, the claim seems not to be contradictory, and to reflect aspects of our ordinary thought about ourselves. That is, I find myself in a conflicting doxastic state with regard to the claim that we were once zygotes. So, I am unsure whether to believe that persons come into existence at the same time as their organisms, or to believe the negation of this. Both claims seem to me to have an initial kind of plausibility, but I am also aware that they can’t both be true. Moreover, I think that I am not unusual in this respect. And this is why I think that the argument McMahan gives is weak. It relies upon a claim about what we believe that seems to me to very doubtful.

There may of course be reasons for thinking the claim that we were once zygotes is false that go beyond how the claim strikes us when considering it impartially, and McMahan does mention one. He mentions that zygotes are unthinking things, and suggests it is hard to believe that we were ever unthinking things. But as soon as one reflects that when our organisms are in a deep sleep they are likewise unthinking (or, at least, it is possible that this is so, and I’m sure many believe it to be true), the mere fact that they are unthinking things can be seen to be a bad reason for denying that we were once zygotes (for I am certain that none of us think that we go out of existence when our organisms are sleeping). One might well respond to this line of thought by invoking the notion that when a human organism is sleeping it retains the *capacity* for thought, which is something that the zygote lacks. But this is a notoriously difficult notion to pin down, and no-one, so far as I know, has done so satisfactorily. The problem is to explain what such a capacity amounts to in a way

that does not entail that zygotes likewise possess such a capacity. A capacity, it seems, is a potentiality, and so to say that a sleeping organism has the capacity for thought is to say that it has the potential to have thought, meaning that were such and such changes to occur to it, it would at some time later come to entertain thoughts. But if this is what it means to say that a sleeping organism has the capacity for thought, so does any organism that is a zygote, for were such and such changes to occur to it, it too would at some time later come to entertain thoughts (i.e. were it to develop into a baby). Of course, organisms can be awoken much more quickly than they can develop from zygotes into babies, but why should the length of time be an issue here? And should anyone suggest that the length of time is in fact an issue, we can then point out to them that the length of time required to rouse an unthinking organism from a coma may be much longer than the nine months it takes one to develop from a zygote into a baby. So, as I suspect few would 'find it hard to believe' that persons continue to exist when their organisms are in a comatose state, I do not see how the notion of *capacity* can be explained in a way that justifies the claim that sleeping organisms possess the capacity for thought whilst zygotes do not. And so I find it hard to see why the mere fact that they are unthinking gives us reason to doubt the claim that we were once zygotes.

It also worth noting that McMahan mentions, in passing, the argument that we cannot once have been zygotes because in identical twin cases zygotes undergo twinning and develop into two distinct persons. He says:

If this is right, then identical twins have organisms that began to exist slightly later in the process of gestation than those of the rest of us. This also suggests an objection to the idea that we are organisms. For, if we are organisms, then one of us ceases to exist whenever monozygotic twinning occurs. If it is bad when one of us ceases to exist, thereby losing the whole of a life that would have been worth living, then twinning is bad. But no one believes this.³

I think this argument is weak too, as it is open to an easy refutation if one accepts certain other widely held metaphysical views. I have dealt with this issue elsewhere, however, so say nothing further about it here.⁴

McMahan's next argument against the identity view goes like this: If we are identical with our organisms, then we go out of existence when they do, but as our organisms continue to exist as corpses when they die, so do we, and so we 'survive death as corpses'. However, we all believe that persons go out of existence when their organisms die, so we ought to reject the view that we are identical with our organisms. (NB McMahan adds to the first clause of this sentence the qualification 'unless we believe that we are immortal souls'. This qualification is to be understood as holding in what follows too.)⁵ This too is a weak argument, I think, predominantly because it also relies on a doubtful claim about what we believe. Is it true that we all believe that we go out of existence when we die? Well, I for one do

³ Jeff McMahan (1995), op. cit.

⁴ See Benjamin L. Curtis, "A zygote could be a human: A defence of conceptionism against fission arguments," *Bioethics* 26 (3) (2012): 136–142.

⁵ Jeff McMahan (1995), op. cit.

not believe this. It seems to me quite plausible to suggest that we continue to exist *in some form* after we die. It cannot be claimed that we continue to exist as perambulating self-sustaining entities, it is true, but it can be claimed that we continue to exist as dead motionless entities, and I don't see any patent absurdity in this. Once more, I am not saying that I think this is clearly true, but it does not seem to me to be clearly false either. In fact I find myself, once more, in a conflicting doxastic state about it. As before, both the claim and its negation seem to have some plausibility, and each seems to capture at least some aspects of our ordinary thought about ourselves, but I am also aware that they cannot both be true. Again, perhaps there are good independent reasons for doubting the claim that we can exist as dead organisms, but in this case McMahan does not mention any.

Although I am not normally one to pick hairs, I do think that it is also worth noting here that McMahan's use of the term 'survive' is leading, to say the least. 'To survive', at least on one meaning of the term, means simply 'to continue living'. Consider, for example, how odd it would be for someone to say in reply to an enquiry about his garden: 'Well, the plants are all dead, but they have survived.' And I think at least some of the force of McMahan's argument here derives from our hearing a hint of contradiction upon learning that it is a consequence of the identity view that we 'survive death as corpses' (i.e. it suggests that we both continue to live and that we do not continue to live). Once one spells out this consequence of the identity view in unambiguous terms, i.e. as 'we continue to exist in the form of a dead motionless organism' it loses that hint of contradiction altogether.

The final brief argument against the identity view that McMahan mentions is a mainstay from the personal identity literature. It goes like this: Assume for *reductio* that persons are identical with their organisms. Now suppose we have a person P1 who is identical with his organism O1 at t1. Now consider that the brain from O1 is transplanted into a distinct organism O2 (I assume with McMahan here that organisms definitely do not go where their brains go, and so that this is a coherent description of the possibility). If everything is done skilfully, then at t2, after the operation, there will be a person P2 who is identical with the organism O2 but who is psychologically continuous with P1 at t1. P2 at t2 will have all the memories and personality traits that were possessed by P1 at t1, and will even claim to be P1. Because of these facts, most of us believe that P2 is identical with P1. But if this is so, then we have that $P1 = P2$, that $P1 = O1$, and that $P2 = O2$, from which we can derive that $O1 = O2$, which is plainly false. So the original claim that persons are identical with their organisms must likewise be false.⁶ I will return to this kind of argument later on, and so I content myself at this point by noting that arguments of this sort are far from clear-cut. They rely upon inducing the belief in us that persons go where their psychologies go. But it is far from clear that these beliefs, so induced, should be afforded much evidential weight. This is especially so given that examples in which psychological continuity between distinct organisms is preserved but no brain is transplanted induce the same belief in us, but that in such cases it is apparently possible, as Bernard Williams has shown, to induce in us the opposing belief that persons go not where their psychologies go, but rather where their

⁶ Ibid., p. 100.

organisms go, by merely re-describing the example. (I will have more to say about Bernard Williams's point later.)⁷

I turn now to briefly describing just one argument against the non-identity view that has at least as much weight as those that McMahan marshals against the identity view. It goes like this: If persons are not identical with their organisms, then it seems that persons and organisms are distinct but materially coincident entities, i.e. that they share their material parts at the times at which they both exist. But then persons and organisms share all of their intrinsic material properties at those times, and so assuming that a person's thoughts arise from (or supervene upon, or are grounded in) those intrinsic material properties that it possesses, if persons are thinking things, organisms must be too. In particular, organisms must have all and exactly the same thoughts that persons do. So, whenever a person has a particular thought, so does the organism. So, there are two distinct but indistinguishable thinking things located wherever a person is located: viz. the person herself and her organism. But we do not believe that there are two thinkers located wherever a person is located (how would we know which we are?). So it must be false that persons are not identical with their organisms.

I do not pretend that this is a knockdown argument against the non-identity view. There is much that could be, and has been, said in response to it. Indeed, McMahan himself devotes a few lines to responding to this argument, albeit, in my view unsuccessfully. He argues that psychological predicates do not apply to organisms. But, so far as I can see, he offers no reason for thinking that this is so other than to avoid the conclusion of the argument. But this will not do. How can psychological predicates that apply to persons fail to apply to organisms if persons and organisms possess the same intrinsic properties? (NB Harold Noonan has offered a more sophisticated response to this question.)⁸

But, although much could be said in response to this argument, it is at least as plausible as the ones presented by McMahan against the identity view, and in fact I find it to be somewhat more plausible than the first two that he presents. Unlike the claim that we were once zygotes, and the claim that we continue to exist as dead organisms when we die, I think the claim that there are two thinkers present wherever there is a person has no claim to initial plausibility at all, and is in fact positively implausible. It captures nothing of our ordinary thinking about ourselves and I strongly suspect that almost all would reject it outright. Still, there is no need for me to push this point, as it is independent of my main argument. So I am happy to rest content with the modest claim that this argument gives us at least as much reason to doubt the non-identity view as McMahan's give us to doubt the identity view.

I finish my background remarks by noting some further commitments that any proponent of the two-deaths thesis has. According to the two-deaths thesis we possess two concepts of death, one appropriate for organisms and one appropriate

⁷ Bernard Williams, "The Self and the Future," *Philosophical Review* 79 (2) (1970): 161–180, pp. 167–169.

⁸ See McMahan (1995), op. cit., p. 116 and Harold Noonan, "The Thinking Animal Problem and Personal Pronoun Revisionism," *Analysis* 70 (1): 93–98.

for persons. The concept appropriate for organismic death, on the two-death thesis, is presumably one according to which organisms die when they lose the capacity to function as an integrated whole in some sense (the details will not matter here). ‘Presumably’ because this is the ordinary concept of organismic death, and I see no reason why two-deaths theorists would want to deny it. And what concept of death is appropriate for persons on the two-deaths thesis? Presumably it is one according to which persons die when they think their last thought, or, perhaps, when they permanently lose the capacity for thought in some sense (again, the details will not matter here). ‘Presumably’ simply because there seems to be no other suitable conception in the vicinity. (Ask: If personal death differs from organismic death, and persons do not die when they lose the capacity for consciousness, then when do they die?)

3 Linguistic Usage

In this section I argue that our linguistic usage is such that it leaves the reference of the term ‘person’ indeterminate, and that therefore it is indeterminate whether the two-deaths or the one-death thesis is true. I have, in fact, already presented some of my reasons for making this claim. As already noted, some such as McMahan find it positively implausible that persons continue to exist after death as their dead organisms. I believe that McMahan is in the minority here and I believe that most are more like I am (or, at least, most of those who are not encumbered by pre-existing philosophical beliefs). I find the claim that we continue to exist as dead organisms to be quite plausible, although I find its negation to be so also, so I find myself in a conflicting doxastic position about it. That is to say, I think that the concept of a person, and the linguistic usage associated with it, is associated both with psychological persistence conditions and with biological persistence conditions. I suggest that we take this as evidence that our use of the term ‘person’ is indeterminate between psychological continuants and biological continuants, and so that it is indeterminate whether persons endure death as their organisms or whether they cease to exist when their organisms die and they lose the capacity for thought. The bare hypothesis alone gives a strong explanation for why I find both the claim that we can endure death as organisms, and its negation, to be plausible. However, the hypothesis can be further supported by independent argument.

First, the claim that the term ‘person’ is associated with psychological persistence conditions. Following Locke, the term ‘person’ is often thought to *mean* something like ‘a conscious, intelligent being’, and so given the widespread acceptance of this view, it is plausible that persons are in one sense, by definition, and so essentially, conscious intelligent beings.⁹ This is bolstered by the fact that if we were faced with a dead body, and asked directly whether it is a person or not, we would be likely to say that is not (i.e. we would be likely to withhold the application of the term ‘person’ to a dead body). Also, we can be induced to believe that persons can come to be associated with two distinct organisms at two distinct times by the kind of

⁹ John Locke, *An Essay Concerning Human Understanding* (London: Everyman): 180.

mind-swapping thought experiment mentioned above. But why should such a thought experiment induce such a belief in us? Because we consider how we would apply the term ‘person’ in such a case, and find that we would apply it to a purely psychological continuant. So it cannot be determinate that the term ‘person’ is *not* associated with psychological persistence conditions.

Second, the claim that the term ‘person’ is associated with biological persistence conditions. Despite the fact that if asked directly, we would not say that a dead body is a person, there are clear uses of the term ‘person’ that suggest that persons can exist as dead bodies, the most obvious here, of course, being its use in the common phrase ‘dead person’. But consider also the use of personal proper names. Consider a case in which someone turns up to a funeral unaware that it is her friend Smith’s funeral, and note how natural (albeit perhaps insensitive) the following response to her question of where Smith is sounds:

Response: He’s in the coffin at the front. Didn’t you hear, Smith died.

Other such phrases are common at funerals (e.g. “Today we bury our friend Smith...””) and I trust the reader can come up with many suitable examples. Also, as mentioned, we can be induced to believe that persons go where their organisms go, rather than where their psychologies go, by certain other thought experiments, to be discussed more fully shortly. But why should such thought experiments induce such a belief in us? Because we consider how we would apply the term ‘person’ in such cases, and find that we would apply it to a thing with biological persistence conditions (i.e. an organism). So it cannot be determinate that the term ‘person’ is *not* associated with biological persistence conditions either.

Of course, the committed defender of either the two-deaths thesis or the one-death thesis can respond to these arguments in various ways. For example, one-death theorists can respond to the point that we would say that a dead body is not a person, if asked directly, by drawing a parallel with certain other terms, such as ‘King’. Were we to be presented with a King after his abdication, we would likewise be likely to say that he is not a King, but this does not entail that Kings cannot endure abdication. The term ‘King’ is a phase sortal that only appropriately applies to something when it has certain properties, but those properties are nonetheless not essential and can be lost by the thing in question without that thing ceasing to exist. Similarly, one-death theorists can claim, the term ‘person’ is a phase sortal that only appropriately applies to something when it has certain properties (in this case, the most likely candidates are the Lockean psychological properties mentioned earlier), but that those properties are not essential and can be lost by the thing in question without it going out of existence too. And two-deaths theorists can respond to the point that we sometimes apply the term ‘person’ and personal names to dead bodies by claiming that these are mere *façons de parler*. They can claim that ‘dead person’ just means ‘the body of a person who no longer exists’ and that ‘Smith is in the coffin at the front’ just means that Smith’s *body* is the coffin at the front, and so on. But although I admit that there are arguments that can be given by the committed two-deaths and one-death theorists, I think that my diagnosis is much simpler. If the term ‘person’ is indeterminate between psychological and biological continuants, this explains why we find it natural to employ them to

refer to both, but in different circumstances. Certain circumstances serve to make one precisification of the term more salient than the other, although upon reflection we may not be so sure about our usage, and so neither precisification has a claim to be *the* correct precisification.

I return now to the last of the brief arguments that McMahan presents against the identity view, for I think a certain parallel can be made between it and another example from the literature on identity that makes my case stronger. The last of McMahan's brief arguments, recall, was the argument in which a description is given in which brains are swapped between two organisms, preserving the psychologies associated with each. This type of example induces us to believe that persons can be swapped from one organism into another, and that they therefore go where their psychologies go. As mentioned, similar thought experiments can be constructed in which no physical swap takes place, but where one organism is made to take on the psychology of another, and vice versa. Bernard Williams describes just such a case where we are asked to consider that we are associated with a body A, but that our memories and other psychological traits will be swapped into another body B. Simultaneously, the memories and psychological traits associated with body B will be swapped into the body associated with us, i.e. body A. Finally, we are told that after the procedure body A is to be tortured. If someone were to describe this case in this way, and we believed it were going to occur, we would likely feel relief that the torture would not happen to us, Williams suggests. But we would only feel this if we believed that we go where our psychologies go (i.e. if we believed we would be swapped across into body B). So, this thought experiment induces us to believe that persons go where their psychologies go, so can be swapped from one body, or organism, to another, and thus induces us to believe that the non-identity view is true, and so that the two-deaths thesis is true. However, Williams goes on to re-describe the case in such a way that the opposing belief, that persons go where their biological organisms go, is induced. I quote Williams at length, because I cannot better his description:

Someone in whose power I am tells me that I am going to be tortured tomorrow. I am frightened, and look forward to tomorrow in great apprehension. He adds that when the time comes, I shall not remember being told that this was going to happen to me, since shortly before the torture something else will be done to me which will make me forget the announcement. This certainly will not cheer me up, since I know perfectly well that I can forget things, and that there is such a thing as indeed being tortured unexpectedly because I had forgotten or been made to forget a prediction of the torture: that will still be a torture which, so long as I do know about the prediction, I look forward to in fear. He then adds that my forgetting the announcement will be only part of a larger process: when the moment of torture comes, I shall not remember any of the things I am now in a position to remember. This does not cheer me up, either, since I can readily conceive of being involved in an accident, for instance, as a result of which I wake up in a completely amnesiac state and also in great pain; that could certainly happen to me, I should not like it to happen to me, nor to know that it was going to

happen to me. He now further adds that at the moment of torture I shall not only not remember the things I am now in a position to remember, but will have a different set of impressions of my past, quite different from the memories I now have. I do not think that this would cheer me up, either. For I can at least conceive the possibility, if not the concrete reality, of going completely mad, and thinking perhaps that I am George IV or somebody; and being told that something like that was going to happen to me would have no tendency to reduce the terror of being told authoritatively that I was going to be tortured, but would merely compound the horror. Nor do I see why I should be put into any better frame of mind by the person in charge adding lastly that the impressions of my past with which I shall be equipped on the eve of torture will exactly fit the past of another person now living, and that indeed I shall acquire these impressions by (for instance) information now in his brain being copied into mine. Fear, surely, would still be the proper reaction: and not because one did not know what was going to happen, but because in one vital respect at least one did know what was going to happen—torture, which one can indeed expect to happen to oneself, and to be preceded by certain mental derangements as well.¹⁰

What are we to make of the fact that different descriptions of the same case induce different beliefs in us about the persistence conditions of persons? Once more I think that we should take this as evidence that our use of the term 'person' leaves it indeterminate whether we go where our psychologies go or where our organisms go. If this is so, it is unsurprising that by using certain forms of words, and descriptions of cases, that one precisification of the term can be made salient, and by using another description the other precisification can be made salient. Much the same occurs in Hobbes' well-known ship of Theseus example where it seems clear that the correct conclusion to draw is that our linguistic usage is indeterminate.

Consider a ship S1 at t1 that undergoes a gradual (plank-by-plank) replacement of all of its parts until at a later time t2 there exists a ship S2 that is spatiotemporally continuous with S1 but has none of S1's parts. Ordinarily, we would be perfectly happy to assert that S1 = S2. Now consider a ship S3 at t1 that has its parts gradually removed and, as they are taken away, they are put back together again elsewhere until at a later time t2 there exists a ship S4 that is made from all of S3's parts. Ordinarily, we would be perfectly happy to assert that S3 = S4. But now suppose that, in fact S1 = S3, and so that the two cases are in fact partial descriptions of exactly the same case. In other words, suppose that an original ship had its parts slowly taken away and replaced, but that the parts that were taken away were put back together to form a ship elsewhere. Now, because S1 = S3, we cannot assert both that S1 = S2 and that S3 = S4, because then it follows that S2 = S4, which is plainly false. So what do we say? Perhaps one has the opinion that spatiotemporal continuity is more important than part-sharing in matters of ship identity over time. In that case one thinks that in fact the second description is misleading and that S1/S3 is identical with S2 and not S4. One may allow that part-

¹⁰ Bernard Williams, "The Self and the Future," *Philosophical Review* 79 (2) (1970): 161–180, pp. 167–168.

sharing is sufficient for identity in cases where ships are taken apart and put back together elsewhere *without* replacements being made, and so think that the persistence conditions for ships are disjunctive affairs, or one may reject part-sharing out of hand and maintain that one simply went wrong when one was induced into believing that $S3 = S4$. Similar comments apply if one has the opposite opinion that part-sharing is more important than spatiotemporal continuity in matters of ship identity over time, in which case one will believe that the first description is the misleading one and that $S1/S3$ is identical with $S4$ and not $S2$. I have to admit that, in this case, I find myself once more in a conflicting doxastic state. Both the claim that $S1 = S2$ and the claim that $S3 = S4$ seem plausible, but I know that both cannot be true. And in this case it seems clear that this is due to the fact that the term ‘ship’, as least as far as our linguistic usage goes, is indeterminate. It is associated *both* with persistence conditions involving spatiotemporal continuity and with persistence conditions involving part-sharing, and indeterminately picks out each kind of continuant. But this case is parallel with the body-swap case, and so I think the same conclusion should be drawn in that case as is drawn in this case. So, I think that we have good reason for thinking that, so far as our linguistic usage goes, the term ‘person’ is indeterminate between psychological continuants and biological continuants, and so it is indeterminate whether the two-deaths thesis or the one-death thesis is true.

4 The Dual Constraint Doctrine

There is an extension of the above argument available to anyone who holds a particular metaphysical doctrine. The doctrine, which I shall call ‘the dual constraint doctrine’ (or ‘the DC doctrine’ for short), is the doctrine that our linguistic practices and the world itself deliver dual constraints on reference. According to it our linguistic practices only partially determine what our terms pick out, with much of the slack being taken up by the existence of certain properties, the instantiation of which makes things into eligible referents (or, as some have put it, ‘reference magnets’). The DC doctrine was first propounded by G. H. Merrill, and later developed by David Lewis, in response to an argument due to Hilary Putnam.¹¹ I do not intend to give the details of Putnam’s argument, nor rehearse the reasons why Merrill and Lewis thought that the DC doctrine was required to refute it. Indeed, I do not intend to discuss the reasons for believing the doctrine at all, for my aim here is simply to draw out its consequences. Suffice it to say that, in fact, unless one is a global descriptivist, the DC doctrine is not strictly required to refute Putnam’s argument, for a causal constraint on reference would do just as well in that regard. Nonetheless, the doctrine has seemed to many (including myself) to be

¹¹ See G. H. Merrill, “The Model-Theoretic Argument Against Realism,” *Philosophy of Science* 47 (1) (1980): 69–81; David Lewis, “New Work for a Theory of Universals” in *Australian Journal of Philosophy* 61 (1983): 343–377; David Lewis, “Putnam’s Paradox,” *Australian Journal of Philosophy* 62 (3) (1984): 221–236; and Hilary Putnam, “Models and Reality,” *Journal of Symbolic Logic* 45 (3) (1980): 464–482.

independently plausible, and it now has the status of a ‘standard view’ (at least, in the metaphysics literature), and is held by many.¹²

The properties that engender results of referential eligibility are the more-or-less *natural* properties, and the more natural a property is, the greater its power to engender such results. According to Lewis’s account, there is a class of elite properties (the perfectly natural properties) which is comprised of the fundamental properties identified by physics, such as mass, charge, and spin (assuming, that is, that current physics is right about these properties being fundamental). These properties are ones that bestow the highest degree of referential eligibility upon the things that possess them, but other less-than-perfectly-natural-but-still-pretty-natural properties also bestow a high degree of referential eligibility. How natural a property is, according to Lewis, depends on how simply it can be defined from the perfectly natural ones:

... physics discovers which things and classes are the most elite of all; but others are elite also, though to a lesser degree. The less elite are so because they are connected to the most elite by chains of definability. Long chains, by the time we reach the moderately elite classes of cats and pencils and puddles; but the chains required to reach the utterly ineligible would be far longer still.¹³

No doubt there are other possible conceptions of naturalness in the vicinity, but quite how we identify them will not matter here. All that will matter is that we can recognise that some properties are more natural than others, and that some are natural enough to bestow a high degree of referential eligibility on those things that instantiate them. The view that Lewis holds, that physics identifies the most natural properties of all, but that the special sciences (biology, chemistry, etc.) identify properties with a high degree of naturalness, is the most widely held view, and the most plausible, for we know that those working in the special sciences have a good grasp of how the properties they identify can be defined from the those identified by physics. So properties identified by the special sciences have an excellent claim to be properties that bestow a high degree of referential eligibility. By contrast, artefactual properties (e.g. *being a chair*) are not nearly as natural, and so have much lesser claim to be such properties.

I will shortly illustrate the above with an example. But before I do it should be noted that although referential eligibility pulls against the other constraint on reference, viz. linguistic usage, it does not necessarily trump it. Where we use our terms in such a way that we *clearly* intend to refer to a less than very natural kind of entity, then it is overwhelmingly plausible that we refer to it, whether or not there is some very natural kind of entity in the vicinity. For example, tree-trunks cut into suitable sizes are sometimes used as tables, and referred to as such. Take a particular

¹² The DC doctrine is been defended by, amongst others, Brian Weatherson, “What Good are Counterexamples?”, *Philosophical Studies* 115 (1) (2003): 1–31; Theodore Sider in “Ontological Realism”, in David Chalmers, David Manley, and Ryan Wasserman, eds., *Metametaphysics* (Oxford: Oxford University Press) (2009), pp. 384–423; and J. R. G. Williams’ influential “Reference Magnetism and the Reduction of Reference” (Unpublished).

¹³ David Lewis, op. cit., p. 228.

example of such a table and suppose the term 'T' is introduced as a name for it (i.e. as a name for a *table*). Were we to somehow alter T so that it clearly remained a tree-trunk whilst becoming entirely unsuitable for use as a table, this would be a case in which T is destroyed despite the fact that *being a tree-trunk* is a much more natural property than *being a table*. (Of course, the tree-trunk itself would endure, but the tree-trunk is not identical with T, for T is a table.) As another example consider that the term 'cyborg' is a term which is used quite explicitly to refer to beings with both organic and mechanical or electronic parts. Functional micro-electronic devices have already been permanently embedded within the bodies of arthropods (specifically, moths) creating hybrid creatures of this sort, and it is implausible to suppose that the term 'cyborg' when applied to them actually refers to their moth parts only rather than the sum of their moth and non-moth parts, whether or not their moth-parts instantiate a particularly natural property that the non-moth parts do not. Similarly to before, suppose the term 'C' is introduced as the name for such a cyborg. If C were to undergo a procedure to have its non-moth parts removed, this would result in the destruction of C, even if the moth that partially overlapped with the cyborg prior to the procedure were itself to survive. In these cases our linguistic usage is clear and (relatively) determinate, and so natural properties play little to no part in engendering reference-fixing results. It is in cases where linguistic usage is indeterminate that natural properties come into play. The following fictional example illustrates such a case clearly:

Consider a community of speakers, the As, who use the general term 'drog' to speak about certain black four-legged objects that they see skulking in the dark at night. Suppose the As only ever speak about these objects when they see them by exclaiming 'There goes a drog'. Suppose that the As have no particular conception of these objects, in the sense that they do not conceive of them as falling under any higher-order sortal kind term such as 'animal' or 'artefact' (i.e. they are ignorant about whether these objects are biological entities or mechanical entities, or for that matter, supernatural entities), but they are able to track these objects across space and time as we track most ordinary objects, i.e. (roughly) in terms of spatiotemporal continuity. Suppose further that the As only ever see black four-legged instances of these objects, and that each instance that they see remains well integrated across those periods of time that they track them, i.e. they remain integrated wholes that are clearly demarcated from their environment.

Now consider the following scenarios:

1. Suppose there also exist certain grey objects that are otherwise indistinguishable from those black things that the As call 'dregs', and ask: do these otherwise indistinguishable grey objects fall into the extension of the A's term 'drog' or not? (i.e. Are such grey objects dregs?)
2. Consider what would happen were a member of the As to catch one of the those four-legged things they call 'dregs', and were to remove one of its legs (leaving it otherwise intact and capable of independent movement as before). Would the three-legged thing that results be a drog, and if so would it be the same drog or a different drog from before? (i.e. Would this be a case in which the As had

destroyed a dog, or a case in which they had merely altered the appearance of a dog?)

What is at issue in question 1 is the synchronic identity conditions for dogs. The question, that is, is whether an object that instantiates a particular set of properties at a time counts as being a dog. By contrast it is the diachronic identity conditions for dogs are at issue in question 2. The question, that is, is whether an object that is a dog can survive various changes in its properties. One might think that both questions can be resolved by consulting the modal intuitions of the As. One might think, that is, that one could ask the As (i.) whether they would consider a grey object that is otherwise indistinguishable from a normal instance of a dog to be a dog, and (ii.) whether they would consider the removal of a leg from a dog to result in its destruction. One might think that if each of the As unhesitatingly answered ‘Yes’ to either question then this would settle the matter in the affirmative, and if each unhesitatingly answered ‘No’ to either question then this would settle the matter in the negative. And one might be right. But suppose that the As would have no clear modal intuitions either way. Suppose that whilst perhaps some As would find themselves able to pronounce clearly on the matter, most would remain puzzled, unsure about how to answer each question. Suppose, that is, that most would find both answers to each question equally plausible when considering them impartially. In such a situation, it would seem, the linguistic usage of the As leaves it indeterminate whether grey objects are dogs, and whether dogs can endure the removal of a leg.

However, now suppose that the objects picked out by the term ‘dog’ are materially coincident with dogs, i.e. that the objects that the As refer to using the term ‘dog’ share their spatial and temporal extents (during those periods of time across which the As track them) with the biological entities that we call ‘dogs’ and that fall under the higher-order sortal kind term ‘animal’. Because *being a dog* is a particularly natural property, the DC doctrine plausibly entails that, despite the As (at least collective) uncertainty about the answers to 1 and 2, the As do indeed refer to dogs using the term ‘dog’, and that therefore (because dogs can indeed be grey) the grey objects are dogs and (because dogs can endure the removal of a leg) dogs can endure the removal of a leg. So, dogs, *as the As conceive of them*, can be grey and can survive the removal of a leg, even though the As themselves are (collectively) unsure about whether this is possible. In particular, the proponent of the DC doctrine can maintain that there is a natural property (the property of *being a dog*) that is instantiated by black and grey dogs alike, and that is not lost by a dog when it has a leg removed. Given the indeterminacy in the As usage of the term ‘dog’, these facts come into play to resolve the indeterminacy such that the A’s term ‘dog’ ends up referring to dogs.

Given my conclusion in Sect. 5, the application of this point to the topic of death should now be obvious. If the DC doctrine is true the fact that the term ‘person’ is indeterminate does not settle the matter of what the term refers to, for this also depends on whether psychological continuants are more natural than biological continuants, or vice versa – i.e. on whether either instantiates a property that is natural enough to bestow referential eligibility upon it. And here I think that

matters are quite clear. Biological continuants – organisms – are clearly the more natural entities. They are studied by biologists who have a great general scientific understanding of their properties and how they relate to the properties of physics. Organisms are, in fact, paradigms of natural entities. This is in contrast with psychological continuants, which, as they can swap from one organism to another, and could perhaps be sustained within a wholly non-biological ‘body’ (as there is nothing in principle preventing a person’s psychology being transferred into a machine), are far from natural. So, I think that it is entirely plausible in this case, just as it was in the case of the As with their use of ‘drog’, that despite our (collective) uncertainty about whether persons can survive organismic death, that in fact they can. Organisms are far more eligible as candidates for reference than psychological continuants, and so our indeterminate use of the term ‘person’ notwithstanding, it is to organisms that we refer using that term. And, as a consequence of this, if the DC doctrine is true, it is plausible that we are organisms, have the persistence conditions of organisms, and die when our organisms die.

5 Conclusion

I have argued that our use of the term ‘person’ is indeterminate between psychological and biological continuants, and that given this the best case that can be made is that it is indeterminate whether the two-deaths thesis or the one-death thesis is true. But, furthermore, our indeterminate use of the term ‘person’ notwithstanding, if one adopts the DC doctrine, the argument can be extended and a good case made for the truth of the one-death thesis. If the DC doctrine is true, it is plausible that the indeterminacy of the term ‘person’ is resolved by the fact that *being an organism* is a particularly natural property, and that as such, we refer to organisms using it. As a consequence, persons are identical with their organisms, and so die when their organisms die.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.