

Brain Death and Irreplaceable Parts

Christopher Tollefsen

I. Introduction

Could a human being survive the complete death of his brain? I am going to argue that the answer is no. I'm going to assume a claim that is common to all the Aristotelian-Thomists engaged in the controversy: that you are I are human animals, organisms of a particular species, *homo sapiens*. As such, we are not numerically identical to any of our parts: no human person is identical to a hand, liver, skin cell, brain, or soul, but each is rather a composite, of form and matter, soul and body – however one wants to describe it at some level of resolution or other.

I'm only concerned with human persons in what follows, so whenever I talk about organisms, you can take it for granted that I'm talking only about human organisms.

I'm going to take it for granted that in some cases, an organism of some sort exists after the complete destruction of the brain that heretofore was housed in the skull associated with that organism. If it turns out that the arguments for thinking that this never happens – that what looks like a residual organism is really only a coordinated set of organic structures – are sound, then so much the better. My argument would not be vitiated by that, just rendered somewhat redundant. But my argument goes through even if there really is an organism of some kind, which is good, because that is the case that poses the toughest objection to the claim that you do not survive the complete death of your brain. After all, if there is an organism after the loss of the brain, it is natural to think it is the same one that was there before the brain perished.

Here is the argument in very brief form: the human organism cannot survive the loss of its irreplaceable parts. Hence when an irreplaceable part is gone, whatever remains is not numerically identical with the organism that existed previously. If you were the original organism, then you do not survive as the residuum. The brain is such an irreplaceable part, therefore, etc.

One will want to know, of course, what I mean by irreplaceable parts; which parts are, in fact, irreplaceable; and why it is that the loss of such parts has such consequences; my paper will be an attempt to answer all these questions.

II. Irreplaceable Parts

The idea that some parts of the human animal are irreplaceable provides, I'll suggest below, a key to understanding much of the controversy over brain death. But some precision is needed as to which sense of "irreplaceable" is being used. I think there are three possible senses.

First, it could simply mean that nothing else can substitute for it; once it is gone, it is gone. This might be true of an individual's sex cells. You probably can't replace sperm cells in a human male in such a way that they would be *his* sperm cells. The same seems true of a woman's oocytes. No sex cells that came from elsewhere, a donor, or a lab, could be integrated in the right way into the biological economy of the organism.

By contrast, there are plenty of parts that are replaceable in this sense in a human organism. A human animal can lose a hand, and a gifted surgeon can graft a donor's hand to its arm. Its body will, if all goes well, integrate the living or potentially living tissues of the donor hand into its own biological life, and the hand will become his or hers, unlike, as Aristotelian-Thomists tend to agree, a prosthetic limb made out of plastic and metal, which will never be so integrated.

And although a human animal's sex cells can't be replaced, I do think that a penis transplant to a man who has, say, lost his penis in an accident does restore a penis to that man. So in principle it seems to me that some sex organs are in this way replaceable, though not, I think, the gonads, since they are responsible for the generation of the sex cells.

So in this sense, some parts of the human animal are replaceable, others are not. But this does not satisfactorily capture the sense of "irreplaceable" that lurks in the debate over brain death, because the loss of irreplaceable parts in this first sense is compatible with continued existence of the animal whose part it was. So here is a second sense of "irreplaceable": a part is irreplaceable in this sense if the animal cannot live without it. It seems that if there is something irreplaceable in this sense, then when the irreplaceable part ceases to exist, so does the substance of which it was a part.

But this can't be the intended sense either, since a heart is irreplaceable in this sense – the animal can't live without it. But it is not true that the end of the animal's heart automatically means the end of the animal, because the heart *is* replaceable in the *first* sense: with a heart transplant, the animal can continue to live.

What is needed, I think, is something that is irreplaceable in both senses: in this third sense, something is irreplaceable for an animal if and only if it is something the animal cannot live without but that is also such that there is no substitute for it possible. When it is gone, it is gone and so is the animal to which it once belonged. If the brain is an irreplaceable part of human beings in that sense, then whole brain death would therefore mean the end of the human animal, because the animal can't live without it, and there is no substitute.

Here is another way to put the point. A part is irreplaceable to an animal if and only if its absence means that any remaining animal or other organism will fail to be numerically identical to the animal which once had the part. In the brain death case, this would mean that when the brain is entirely destroyed, any remaining animal or organism fails to be numerically identical with the organism that once housed the brain.

It is natural to ask: what is it about a part that could make it irreplaceable in this very strong sense? We can easily see what it is about a part that could make it such that an animal can't live without it – many parts are like that in virtue of the vital functions they fulfill. But the strongly irreplaceable parts will presumably have to fulfill vital functions *and* be unavailable for substitution. What could make a part be like that?

The answer given through much of the dialectic over brain death is that a part is irreplaceable in the strong sense when it plays the coordinating, integrating, and directing role that is necessary for the integrated organic function of the human animal as a whole. Something that does that is clearly irreplaceable in the sense that the animal could not live without it – to be alive is to have the right kind and degree of coordinated integrated function, and something needs to accomplish that. But it also seems plausible that such a part would be irreplaceable in the other sense as well, for, as Moschella argues in her contribution to the brain death symposium, self-directedness is essential to an organism's existence; the coordinating and integrating need to come from within the organism itself. But suppose there is an organ or part that plays this integrating role; then a substitute would seem to be in some way from without, external.

The integrating organ thus stands on different footing from other organs that an animal can't live without but that can be *integrated* into the organism – presumably by the *integrating* organism's performance of its function.

One problem with running the argument like this, though, is that it can seem, in the face of Shewmon's examples, to beg the question. Is this a living human being? If it is, it is an integrated whole, a single living organism. But it can't be, because it does not possess the brain, which is the organ that does that integrating, and maintains the whole as a whole. But that can't simply be a conceptual truth about the brain; and its empirical adequacy is jeopardized by the data. The question we have is whether this apparent organism really is one, and we shouldn't be able to appeal to the presence or absence of the brain to decide the case.

As a side note: I take it something like this thought motivates the move made by Grisez and Lee. For them too it is true that an irreplaceable part is integrating, but their argument works by identifying that which is missing in the residual organism, namely, the capacity for sentience. This is a non-question-begging approach, though it has in common with other approaches the claim that the brain is indeed an integrating organism; it just does not have that claim as a premise.

In this paper, I plan to approach the phenomenon of irreplaceability by way of a somewhat different route. I'm going to argue that there are two irreplaceable parts, the brain and the soul. That claim also is common to most of Aristotelian-Thomists; they agree that a human organism can't live without its brain or its soul, and that there is no substitute for either. On the traditional approach, there is a parallel explanation available

for why these are irreplaceable, viz., that each in some sense does necessary coordinating, integrating, and directing work for the organism as a whole.

But since I am not taking that path, in order to make the argument work, I'm going to have to make some claims that, at least until fairly recently, would have put me well outside the Thomistic mainstream. As we'll see, these claims are now much less universally denigrated; but they still create difficulties, especially if one claims to be an animalist, and I won't be able to deal with all those difficulties.

What is the claim that is true of both the brain and the soul that identifies each organ as irreplaceable in the properly strong sense? It is this: assuming that you are a human being, both the brain and the soul are the only proper parts of the human being which you could exist *as*, under certain, admittedly extreme, circumstances. So my argument needs to defend the following claims: that this is true of the brain and of the soul; that it is true *only* of the brain and the soul; and this identifies the brain and soul as irreplaceable in the suitably strong sense. If these claims are true, then it will follow that no organism that is biologically the successor to a human being but that lacks that human being's brain is that human being; and the same is true of the soul.

III. You Could Exist as Your Brain and You Could Exist as Your Soul

So let's start with the claim that you could exist as your brain. Could you be reduced to life simply as a brain? I think it is the most extreme and unlikely possibility, but I think it could happen. Your brain is removed from the rest of the organism that usually houses it, and is kept alive – in a vat, of course, since it is a philosophy example. No doubt the form of consciousness available to such an immensely mutilated entity would be itself quite mutilated – unimaginable to us, probably. Maybe no form of consciousness could be maintained except in potency, but few I think, would deny that if the brain were kept alive, and then transplanted into a new organism, then the waking organism would be you. And I think it would make sense as well to think that you had survived through the process.

First, for example, your head might have been severed, then the brain excavated, then transported, and then surgically inserted; you would have become much smaller, physically, but then you would have been restored to something more like your original size, height, weight, etc. But you would be present throughout the narrative.

I think the same is true of your soul. Could you survive as only your soul? Currently, there is an argument among Thomists about whether St Thomas believed that you could. Those who think that you could survive as a soul are called “survivalists”; those who deny this “corruptionists.” It seems to me that St Thomas was in the latter camp, despite some texts where at least he suggests otherwise, for he asserted, as is well known, that if my body is not raised to eternal life, but only my soul is, then neither am I raised to eternal life, for “I am not my soul.” The exegetical claim has recently been advanced by Turner Nevitt on the basis of a number of passages in which Aquinas denies that Christ

was a human being during the 3 day period before his resurrection; it has also been argued extensively by Patrick Toner.

It would be too much to wade further into the exegetical waters here, but while I think the corruptionists have the better of the interpretive argument, I've become convinced in recent years that the survivalist position is true, even if it is not Thomas's. It is true that I am not my soul. Neither am I my brain. To exist only as a soul would be to exist in a radically deprived state, one the condition of consciousness of which can barely be imagined. Existing only as a soul would thus be existing in a radically mutilated state, as would existing as a brain.

But the prospects for understanding what it means to say that my soul continues to exist after my death while I do not – and enjoys a form of consciousness aided by the Divine; and is punished; and engages in certain acts; and is the suitable object of petitionary prayer; and yet is not *me* – the prospects for understanding what these claims amount to are dim. The soul is the principle of all my acts as an embodied creature, and I will never exist as I most fully ought to without the restoration of my body, but it seems to me that I can exist as a soul.

In the next section I will defend the claim that you could survive *only* as either your brain or your soul. Before I do so, let me make three qualifying remarks.

First, it is necessary to say something about the expression “could exist *as*.” My claim is not quite that you could be numerically identical to your brain, or numerically identical to your soul. Those claims would be problematic, for anything you *could* be numerically identical to is that which you *essentially* are. If you could be numerically identical to your soul, then you would essentially be a soul. That is a false claim if the starting point of the paper is true: you and I are essentially matter-soul composites.

So *existing as* is not the same as numerical identity. One possibility is that it could be some form of, or something like, a constitution relationship. David Oderberg defends such a claim about both souls and brains (Oderberg speaks of heads); Eleonore Stump also defends this view. In such cases, a whole is constituted by a single part, but the whole is not identical to that part, as can be seen by considering the differing modal properties of each.

A second aside: the fact that I argue that you can exist as your soul without your brain, and that both brain and soul are marked as irreplaceable because you can exist as each might suggest that I think you can exist as your brain without your soul. But of course that couldn't be the case; you are still ensouled when existing as a brain. I don't think that is a problem for my argument, but I await correction.

A final aside: while I say at the beginning that I assume the truth of animalism throughout the paper, I confess to being tempted to think that what is true of your brain and soul is also true of the organism that you are. In *some* sense I think it is true to say that you are

that organism, but I think as well that you – the ‘I’ that goes with that organism – is not reducible to the organism that it – in some sense – is. Our personal existence transcends our organic existence, but not by being the existence of some *thing*, such as a soul, that is different from the living body; indeed, such transcendence claims seem necessary if my claims about what you can survive as are true. But I leave this deeper issue for another time.

II. True Only of Brain and Soul

Are there any other contenders for the role of parts that you or I could exist as? I think there is only one other. This contender is the brain *hemisphere*. One might think that if I could be reduced to a brain, then I could be reduced to a brain hemisphere. Just keep whittling. After all, any reason to think you could be reduced to a brain applies equally strongly to the hemisphere: those who undergo an anatomical hemispherectomy lose half their brain and still manage to experience the range of human consciousness and function. In principle, it seems they could survive the removal of that hemisphere only and its maintenance in the proverbial vat.

But this is simply a case in which the brain is reduced in size to that of (roughly) a hemisphere. The hemisphere *is* the brain, not half a brain. Similarly, if some bit of tissue is removed from your liver for donation, it is not the case that you now have half a liver; you have a whole liver that is smaller.

Now this might pose a problem for my overall argument if it could be the case that both hemispheres could survive as persons. We can agree, I expect, that it is not the case that the two separated hemispheres together are one person. So there are two other options: either one of the hemispheres would be me and the other not, or neither would be.

The first option would be strange but not a problem for my argument; it would require a new person to come into existence that is constituted by a small hemisphere-sized brain. It would not pose a challenge to my claims about what parts *you* could exist as; you could still exist as a brain. The second possibility is incompatible with my argument as we’ll see in the next section. So I need pre-emptively to deny that it is possible, and I’ll do that by denying that either option is really possible.

The only evidence we have to support the claim that you could exist as a hemisphere sized and shaped brain is that it seems you can exist *with* only half a brain after a hemispherectomy. But this is not *exactly* right: in this procedure, a hemisphere is removed from the brain, but it is not the case that the brain is reduced to merely a hemisphere. Rather, deep brain structures remain after the surgery, and so does the brain stem, so that the remaining brain is considerably different from the brain material that is removed. I think that the success of anatomical hemispherectomy gives us reason to think, at the very edges of what is possible, that we could exist as a brain with only one hemisphere; but it gives us no reason to think that we could exist as a hemisphere proper.

So I conclude that the brain and the soul, and only the brain and the soul, are the parts that you or I could exist as.

IV. These Organs are Irreplaceable

In this section I argue that parts you could exist as are irreplaceable in the strong sense; that is, this feature of any part identifies it as irreplaceable. In an appendix, I say something about why we should expect this to be the case. For now, the proof that they are.

Recall the definition of an irreplaceable part in the strong sense: A part is irreplaceable to an animal if and only if its absence means that any remaining animal or other organism will fail to be numerically identical to the animal which once had the part. Put in the terms used earlier, a part is irreplaceable to an animal iff it is something the animal cannot live without but that is also such that there is no substitute for it possible. When it is gone, it is gone and so is the animal to which it once belonged.

Suppose you are an organism (O), but there is some part of that organism (P), such as an organ, that you could in some circumstance exist as, albeit in a radically debilitated state. Suppose that this part is removed and maintained, and so you continue to exist as P; and suppose further that removing P leaves, in this instance, a residual organism (R) – the previously existing organism minus the organ that you now exist as. Grant that this residual organism is alive, one, self-integrating, etc.

Obviously, if you exist as the removed part P, then you do not exist as the residual organism R, so R is also numerically different from O, the original organism, because O *was* you prior to the removal of P. If R was identical to O, then it too would be you.

Now suppose that P is not removed but destroyed, and there remains behind a residual organism exactly like R in the previous scenario. Could R be you now, since there is no other contender? Could it be numerically identical to O?

It could not be. For whether R, the residual organism, could be numerically identical to that earlier organism O should not – I would say cannot – depend upon what happens to something *else*, namely P. David Wiggins calls this the *Only a and b* rule. Where a and b are candidates for the relationship of numerical identity, whether they are numerically identical cannot depend upon what happens, or how they are related, to some third entity c. It can't depend on whether there exists or does not exist a *better contender* for the identity relation.

Similarly, whether the organism that is at one time you (O), and the residual organism (R) that exists after the removal (or death) of P, are numerically identical must depend only on facts about those two entities, O and R, and not on whether or not that third thing P continues to exist or not. So if $O \neq R$ when P survives detached from R, then $O \neq R$

ever, including the case where P merely happens to be completely destroyed, rather than removed and maintained.

This means that P, in virtue of being a part that you could exist as, is irreplaceable in the strong sense. Obviously, its loss means that no residual animal organism is numerically identical to the animal or organism that once had P as a part. Hence P is something the animal cannot live without; and nothing can substitute for P.

V. Application to Brain and Soul

Now everyone agrees that when your soul departs, nothing that might remain, even if alive and one and self-integrating, is identical to the organism that your soul once informed. But it follows from the argument just made that when your brain departs, whether because it perishes, or because it has been preserved and maintained in a vat, nothing that might remain, even if alive, and one and self-integrating, is identical to the organism that once housed your brain. Both your soul and your brain are irreplaceable if you can exist as either, and this was what the argument set out to show: *brain-absent* residual organisms are not numerically identical to their brain-present predecessors, and so you will never exist as a brain dead organism.

Appendix I: A Definitional Question

I define irreplaceability in an odd way that might seem odd:

A part is irreplaceable to an animal if and only if it is something the animal cannot live without but that is also such that there is no substitute for it possible. When it is gone, it is gone and so is the animal to which it once belonged.

Here is a different way I could have defined it: a part is irreplaceable if and only if it is something *you* cannot live without and no substitute for it is possible. When it is gone, it is gone and so are you.

But this definition won't do, because by this definition, the brain is not irreplaceable: if you can exist as a soul, then *you* can exist without your brain; it is not irreplaceable *for you*. But that is not where the focus needs to be anyway: it needs to be, not on what *you* can or cannot live without, but on what the *animal* can or cannot continue to exist without. So the definition makes no reference to "you", although identifying which parts are irreplaceable does: those parts are irreplaceable that you could exist as.

Appendix II: What is the connection between what you can exist as and irreplaceability?

The parts you could exist as are in fact irreplaceable; we know this by application of the *Only a and b* principle. But why? I think we see in the two parts identified, the brain and the soul, a coming together of the features that made other parts irreplaceable in the two weaker senses we identified. On the one hand, we had parts that were irreplaceable

because they performed a vital function, and thus you could not live without them, like a heart. But substitution was possible; you might have an artificial heart.

On the other hand, a very small number of parts were irreplaceable in the sense that there is no substitute for them. Sex cells seem to be the primary case. Why should these cells be irreplaceable? I think it is because they are connected in a deep way to our identity: our biological identity determines which sex cells can plausibly be considered ours, because such cells are the link between ourselves as potential reproducers and subsequent human beings considered as our biological children. Without the right kind of connection, as mediated by *our* sex cells, subsequent children simply are not biologically ours. So those cells do not admit of substitution.

Now brains and souls seem to share both these features. To take the easy case first: your soul performs a vital function, since it is the animating principle of the organism; and it is deeply, constitutively, linked to your identity. Your soul cannot be the soul of another person, nor can any other soul be yours. So the soul is irreplaceable in the strongest sense.

And of the brain too these seem true. We did not want, prior to making the argument, to say that the brain was irreplaceable because it performs an integrating function; but the brain *is* irreplaceable, and a reason we can posit for that is that it does indeed play an integrating function for the organism that you are. I am not abandoning the idea of the brain as the integrating organ any more than Grisez and Lee. And that integrating function is certainly vital.

Moreover, the brain seems, in a way analogous to the soul, linked to your identity, both biologically and personally. The brain, unlike every other organ, provides the material substrate for the acts that are most deeply linked to your identity as a person – your thoughts, choices, sensations, memories, emotions, etc. So, like the soul, it has both the features that together give a part a claim to strong irreplaceability.