CHAPTER 4.2

Persons, Bodies, and Human Beings

Derek Parfit

Of the possible criteria of personal identity, there were three that I took seriously in the book that Judith Thomson discusses.[1] On the *Wide Psychological Criterion*, for some future person to be me, we must be psychologically continuous. On the *Physical Criterion*, which I shall here rename the *Brain Criterion*, we must have the same brain. On the *Narrow Psychological Criterion*, we must both be psychologically continuous and have the same brain.[2]

My discussion of these criteria was confused. At one point, I endorsed the Wide Psychological Criterion (*RP* 208). That was a mistake. As I later claimed, we should not try to decide between these criteria.[3] In the so-called "problem cases," where these criteria conflict, questions about personal identity would be indeterminate, and empty. If we knew the facts about both physical and psychological continuity, there would be nothing further to discover.

According to some writers, even if there is nothing to discover, we should at least consider how we can best refine our concept of a person, by adopting one of these criteria. But that, I believe, is not worth doing. First, as I shall argue again below, personal identity is not what matters. Nor should we try to find some criterion that would make identity coincide with what matters. Such coincidence could not be complete. Unlike what matters, for example, numerical identity cannot hold to different degrees. And, if we try to preserve the coincidence between identity and what matters, it will be harder fully to shake off the view that identity is what matters.

Though I believe that we need not choose between my three criteria, we cannot discuss persons without making some assumptions about personal identity. And, in some of my arguments, I assumed

(1) If there will be a single future person who will have enough of my brain to be psychologically continuous with me, that person will be me.
(2) If some future person will neither be psychologically continuous with me, nor have enough of my brain, that person will not be me.

These claims I shall call *my view*. This view seemed to me uncontroversial, since it is the shared or common element in my three criteria.

Judith Thomson rejects my view. She accepts a fourth criterion, which I dismissed, too lazily, in a single sentence.

1 Thomson's View

Thomson accepts:

The Bodily Criterion:
(1) We continue to exist if and only if our bodies continue to exist.
(2) Our bodies would continue to exist even if our brains were replaced.

On Thomson's view, if her brain were destroyed, and my brain were successfully transplanted into her empty skull, it would be she who would wake up, mistakenly believing that she was me.

Thomson sees little need to defend her view. She calls it "the ordinary view," and she asks: Can anyone think otherwise?

The answer is Yes. I know many people who have considered this question, and most think otherwise.[4]

Even if we all thought otherwise, Thomson's view might still be true. But, before considering Thomson's arguments, it is worth suggesting why hers is *not* the ordinary view.

To apply Thomson's view, we must decide what would count as the continued existence of the same body. We, and our bodies, would survive the loss of certain limbs and organs. But how far could this process go? Thomson says that she wants her body to continue to "support consciousness." Even if she could not act, she writes, "Mere consciousness wouldn't be nothing: I could still ruminate on the nature of the universe."[5] What would be required for someone's body still to support consciousness?

Suppose that I became paralyzed, and lost sensation in all of my body below the neck. Suppose also that, because of damage to my heart and lungs, my brain was connected to a heart-lung machine. If these things happened, both I and my body would continue to exist. I could still see, and hear, and speak, and do philosophy.

Suppose next that, in a second operation, my head was disconnected from the rest of my numb and paralyzed body.[6] Would I then cease to exist? On Thomson's view, the answer depends on whether, if my head continued to exist and to support consciousness, that would amount to my body's continuing to exist.

Assume, first, that it would not. Thomson's view then implies that I would have ceased to exist. If we accept this view, we must decide how we should regard my surviving head. Even if I had ceased to exist, my head would be that of a conscious being, with many of the properties of an ordinary person. This being would have a mental life like the one that I had before my head was disconnected from the rest of my body. It could admire the sunrise, talk to my friends, and dictate the rest of my unfinished book. Given these facts, it would be hard not to regard my head as a living

person's head. But, on this version of Thomson's view, this person cannot be me. My head must now be the head of a new person.

If we accept that conclusion, this case challenges the view that identity is what matters. Suppose that I knew in advance that, if my head were disconnected from the rest of my body, my brain would function rather better, and would no longer be subject to migraine attacks. On that assumption, as I shall argue later, I ought to accept this operation. It would be irrelevant whether, after my head was disconnected, it would become the head, not of me, but of a new person.

Thomson might give a different answer. She might claim that, if my head continued to exist, after the rest of my body was destroyed, this would amount to my body's continuing to exist, though in a diminished state. On this version of Thomson's view, even after this operation, I would still exist.

Our story might continue. Suppose that, after another operation, the blood going to my head came, not from a heart-lung machine, but from someone else's heart and lungs. And suppose that my head was then grafted onto the rest of that other person's body. That other person we can assume to be Thomson, whose head had earlier been destroyed in some accident.

Would these further operations make it true that I would cease to exist? Thomson's answer must be Yes. She believes that, if her body was given a new brain, it would still be the same body, and she would therefore still exist. It could not affect the identity of Thomson's body if its new brain retained its covering of bone and skin. Thomson's view thus implies that, at the end of our story, it would be she who would have my head.

If we accept this view, our story again suggests that identity is not what matters. Thomson may think that no objection. As she remarks, so far as the metaphysics is concerned, whether personal identity matters is not what matters. Our story also, however, challenges Thomson's view. We have now imagined four operations. My head would first be connected to a heart-lung machine, then disconnected from the rest of my body, then reconnected to Thomson's heart and lungs, then grafted onto Thomson's headless body. On Thomson's view, I would not survive all these operations; by the end of this story, my head would have become hers. It is hard to accept this view. In each of these operations, my head would be unharmed. We could even suppose that, throughout our story, there would be no interruption in the stream of thoughts and experiences that would be directly related to my brain. It is hard to believe that, of this continuous stream of experiences, the first would be had by me, and the last by Thomson.

It may be objected that I am appealing here to a Sorites argument, of a kind we know to be unsound. I may seem to be claiming only that, in this series of operations, there could not be a point at which I would cease to exist. And that may seem no better than the claim that, if I kept losing single hairs, there could not be a point at which I would become bald. But my claim is not of this kind. There could, I agree, be a sequence of events in which what starts as one person's body ends as the body of someone else. And, in claiming that such a sequence might occur, we need not suppose that there would have to be some point at which the first person ceased to exist. We could claim that, in the middle of this sequence, it would be indeterminate which person existed.

My claim is different. I believe that, when we consider my imagined story, most of us would be confident that I would survive. Given what we know about physiology, most of us believe that, if our heads will continue to receive the blood that our brains need, and our brains will continue to function in a normal way, we ourselves will continue to exist. What happens below our necks, though it may affect the content of the experiences that are directly related to our brains, cannot, we assume, make it true that these experiences will not be ours.

2 Cases and Intuitions

Thomson's view is, I have claimed, implausible. It may be objected that, in supporting that claim, I have appealed only to one example. Might there not be cases where it is *my* view that is implausible?

It will help to classify the cases where I and Thomson disagree. We are asking when it is true that some future person will be me. On Thomson's view, that person will be me if and only if we have the same body. On my view, that person will be me if we have the same brain, and are uniquely psychologically continuous. If we have different brains, and are not psychologically continuous, that person will not be me. In other cases, on my view, there is indeterminacy, or no answer to the question whether that person will be me.[7]

We can next distinguish two kinds of disagreement. Thomson and I *strongly* disagree when there is some claim that she thinks true, and I think false. We *weakly* disagree when there is some claim that she thinks true, and I think indeterminate, or neither true nor false.

If some future person will be psychologically continuous with me, let us say, for short, that this person will have *my psychology*. Thomson and I strongly disagree in any case where

(A) some future person would have my brain and psychology, but a different body,

or

(B) some future person would have my body, but a different brain and psychology.

There are six kinds of case. (A) would be true when some future person would have my brain and psychology, but either (1) someone else's body, or (2) a new body. (B) would be true when some future person would have my body, but either (3) a new brain and psychology, or (4) someone else's brain and a third person's psychology, or (5) someone else's brain and a new psychology, or (6) a new brain and someone else's psychology.[8]

These imagined cases would involve three kinds of operation. One we have already described: the transplanting of some part of one body into another. I shall assume this part to be, not just a brain, but a whole head. That puts more

pressure on Thomson's view, but in a way that is not unfair. I shall also assume that, if a different head were grafted onto the rest of some headless body, that would not make this a different body. Whether we must accept that assumption I shall reconsider later.

The second kind of operation is the creation of new heads, or bodies. In some cases, these would be exact copies of some existing person's head or body, as in my imagined case of Teletransportation (*RP*: 199–201). Such cases may always be technically impossible. But they are not deeply impossible, or against the laws of nature. And they are easy to understand. We know what it would be for some head, or body, to be exactly copied.

The third kind of operation, which is less straightforward, is the remodeling of some existing person's brain, in a way that destroys all psychological continuity. In some cases, this remodeling would make this brain functionally equivalent to the now destroyed brain of some previously existing person. The resulting person would then be psychologically continuous with that previously existing person. In other cases, this remodeling would not copy any other brain, and the resulting person would be psychologically continuous with no one. Since it is unclear what would be involved in such brain remodeling, these examples are harder to assess.

Let us now consider our six kinds of case:

> Case (1), or *Head Transplant*: Thomson's head is destroyed, and my head is grafted onto the rest of Thomson's body. The resulting person is psychologically continuous with me.

This is the case discussed above. On my view, I survive, but with a different body below the neck. On Thomson's view, she survives, but with my head. Most people's intuitions, I have claimed, here support my view.

> Case (2): My head is grafted onto the rest of a newly created body. The resulting person is psychologically continuous with me.

In one version of this case, the newly created torso and limbs are exact copies of mine. As before, on my view, I survive. On Thomson's view, I die, and the person with my head is someone new. Most people's intuitions, I believe, again support me.

> Case (3): My head is destroyed, and a newly created head, which is unlike mine, is grafted onto the rest of my body. The resulting person is psychologically continuous with no one.

On my view, I die, and a new person has what is left of my body. On Thomson's view, I survive, but with a new head. It may make some difference, here, that this head had no previous owner. But I believe that our intuitions still support me. If we knew that our heads will be destroyed, and that the rest of our bodies will then be connected to dissimilar heads, few of us would expect to wake up again.

The remaining cases can be discussed together:

		This person is, according to:				
		Thomson's Bodily Criterion	My view	The Brain Criterion	The Narrow Psychological Criterion	The Wide Psychological Criterion
If some future person will have:						
(1)	My brain and psychology, but Thomson's body	Thomson	Me	Me	Me	Me
(2)	My brain and psychology, but a new body	A new person	Me	Me	Me	Me
(3)	My body, but a new brain and psychology	Me	A new person	A new person	A new person	A new person
(4)	My body, but Thomson's brain and Williams's psychology	Me	Not Me	Thomson	A new person	Williams
(5)	My body, but Thomson's brain and a new psychology	Me	Not Me	Thomson	A new person	A new person
(6)	My body, but a new brain and Thomson's psychology	Me	Not Me	A new person	A new person	Thomson
(7)	A new body, but my brain and Thomson's psychology	A new person	–	Me	A new person	Thomson
(8)	My body and brain, but Thomson's psychology	Me	–	Me	A new person	Thomson
(9)	My body and brain, but a new psychology	Me	–	Me	A new person	A new person
(10)	A new body and brain, but my psychology	A new person	–	A new person	A new person	Me
(11)	My brain, but Thomson's body and psychology	Thomson	–	Me	A new person	Thomson
(12)	My brain, but a new body and psychology	A new person	–	Me	A new person	A new person
(13)	A new brain, but my body and psychology	Me	–	A new person	A new person	Me

Figure 1

Case (4): My head is destroyed, and Thomson's head is grafted onto the rest of my body. Thomson's brain has been remodeled, in a way that copies the brain of some other previously existing person, whom we can suppose to be Bernard Williams. The resulting person is psychologically continuous with Williams.

Case (5): As in (4), Thomson's head is grafted onto my headless body. Thomson's brain has again been remodeled, but in a way that does not copy anyone else's brain. The resulting person is psychologically continuous with no one.

Case (6): My head is destroyed, and a new head is grafted onto the rest of my body. Since this new head contains a replica of Thomson's brain, the resulting person is psychologically continuous with her.

In all three cases, on Thomson's view, I survive, but with her head. On my view, since my head has been destroyed, and the resulting person is not psychologically continuous with me, that person is not me. In all these cases, as in (3), I die. Our intuitions, I believe, still favor this view.

I conclude that, in all the cases where my view and Thomson's strongly disagree, her view is less plausible.

There are also cases where our views weakly disagree. It is worth listing both kinds of case. And we can include, in this list, what is implied by our other three criteria (see figure 1).

This list can be explained as follows. The views we are considering appeal to three elements: body, brain, and psychology.[9] In the cases in Group B, all three elements have different origins. In Case (4), these elements come from three existing people. In the other cases in this group, one element is newly created: the psychology in (5), the brain in (6), and the rest of the body in (7). In the remaining cases, two elements stay together, but the third has a different origin. In Group A, this third element is the body; in Group C, it is psychology; in Group D, it is the brain.

We can next compare the implications of our five views. In Cases (1) to (3), things are simple. In all three cases, Thomson's view gives one answer to our question, and the other views agree on a different answer.

In Cases (4) to (6), Thomson again gives an answer which the other views reject. Thomson claims that, in each case, the future person would be me. On every other view, that person would *not* be me. But, while the other views agree in rejecting Thomson's claim, they do not agree on who this person would be.

Case (4) is the one where there is most disagreement. In this case, which might be called the *Neapolitan Sandwich*, the resulting person would have Thomson's head, the rest of my body, and be psychologically continuous with Williams. Our views give five different answers. On Thomson's view, the resulting person would be me. On the Brain Criterion, he would be Thomson. On the Narrow Psychological Criterion, he would be a new person. On the Wide Criterion, he would be Williams.[10] On my view, this person would not be me, but it would be indeterminate whether he would be Thomson, Williams, or a new person.

Though I believe that, in Case (4), there is indeterminacy, Thomson and I still strongly disagree. On my view, while none of the other answers is determinately true, Thomson's answer is determinately false.

My view may seem inconsistent. Why do I treat Thomson's answer in a different way? Since I believe that there is indeterminacy, and I claim that we should not try to choose between the other views, why do I reject Thomson's view?

There is no inconsistency. Indeterminacy is always limited. Even if we should not choose between some views, we should reject others. Suppose it were claimed that, if some future person will have my heart, that person will be me. We should reject that view. If my heart were transplanted, we could insist that the resulting person would not be me. Thomson claims more plausibly that, if some future person will have my whole body below the neck, that person will be me. But I am also inclined to reject this view. I am inclined to insist that, if my head were destroyed, and Thomson's head were grafted onto the rest of my body, the person with Thomson's head tomorrow would not be me.

My rejection of Thomson's view might be misunderstood. When Thomson claims that it would be I who would have her head, and I deny that claim, I am not making a different prediction about what, in such a case, I could expect to happen. On my view, if I knew that someone would wake up tomorrow with Thomson's head and the rest of my body, that description tells me what would happen. I can truly claim that this person would not be me. But, in stating that fact, I am not giving further information about what would happen. I am merely giving an acceptable redescription.[11]

That I can truly give this redescription is, I have claimed, a *conceptual* fact. But that claim could also be misunderstood. I do not mean that, when Thomson rejects my redescription, she shows ignorance of our concept of a person. What I mean is, roughly, this. When we consider this imagined case, most of us would accept my redescription; and Thomson's arguments do not, I believe, show that we are mistaken.

I am, I have said, *inclined* to reject Thomson's view. But I might be persuaded that, here too, there is indeterminacy. On my view, the "problem cases" raise no problem, since there is nothing to discover, nor is it important how, given our concept of a person, these cases can be redescribed. So I might be persuaded that Thomson's is among the views between which we need not choose. Indeed, as I shall argue later, even if Thomson's view was determinately true, that should not affect our conclusions about what matters. But, since most of us assume that our identity is what matters, it is worth discussing whether, as I am inclined to believe, we can reject Thomson's view.

I have discussed the cases where Thomson's view seems to me determinately false. In the remaining cases, her view is not, I believe, determinately true.

In

> Case (10): My body is destroyed, and a replica created. The resulting person is psychologically continuous with me.

On my view, it would be indeterminate whether the resulting person would be me, or be a new person. But this is one of the cases where, though there is indeterminacy, one description would be most convenient. It would be best, I suggested, to call my Replica a new person (*RP*: 205). On Thomson's view, that is not merely the best description; it is straightforwardly true. I suspect that, in this case, many people would find Thomson's view more plausible than mine. But the intuitive difference is, I think, much less than in Cases (1) to (6).

In the similar

> Case (13): My head is replaced with an exact replica. The resulting person is psychologically continuous with me.

On Thomson's view, the resulting person would here be me. On my view, it would again be indeterminate whether this person would be me, or someone new.

Here, I believe, our intuitions slightly favor my view. Suppose we accept Thomson's claim about Case (10). We believe that, if my whole body is destroyed, I cease to exist. On this view, for me to continue to exist, it is not enough that there will be a replica of my body. If we accept that claim, we may doubt that Case (13) is, as Thomson claims, quite different. If my head will be destroyed, it may not seem enough that there will be a replica of my head. True, this head will be grafted onto the rest of my body. But that may not seem decisive. After all, if this head were Thomson's, most of us doubt that I would survive. And, if I would not survive either if my whole body were destroyed and copied, or if my head were replaced by a head that is not a copy, why should we be confident that I *would* survive if my head were replaced by a copy?

Consider next

> Case (9): My brain is remodeled, in a way that removes all of my psychological features. The resulting person is psychologically continuous with no one.

On Thomson's view, the resulting person would be me. On my view, it is indeterminate whether this person would be me, or someone new. Thomson's claim may again seem more plausible than mine. When he considers such a case, Williams (1970) argues forcefully that, however much my brain is interfered with, I would continue to exist. This, I conceded, might be the best description. But that concession may again seem inadequate.

Our assessment of this case turns upon the view we take of such brain remodeling. As Shoemaker points out, the required brain surgery might have to be extreme (Shoemaker and Swinburne 1984: 87). In considering a case like (9), we cannot appeal to what we believe about ordinary cases of psychological discontinuity, such as those involving amnesia, or severe brain damage. In those cases, many psychological features are retained. And, even when, at the conscious level, such features have been lost, their underlying neurophysiology is largely undisturbed. If we wanted wholly to remove all psychological connections, we would have to destroy very many of the links between the nerve cells of the brain. And, to create a new psychology, we would have to create new links, in different patterns. It is not intuitively obvious that, if my brain were *zapped* in such a way, and the resulting person was wholly unlike me, I would still exist. Those who make that claim need to appeal, as Williams does, to some argument.

Consider next the similar

> Case (12): As in (9), my brain is remodeled, in a way that removes all psychological features. My head is then grafted onto the rest of a new body.

On my view, as in (9), it is indeterminate whether I survive. On Thomson's view, (9) and (12) are quite different. She believes that, if my remodeled brain stays in my body, I survive; but, if my head is transplanted, my brain becomes the brain of someone new. As before, we may doubt that this makes all the difference. Our intuitions may be closer to my view.

Of the remaining cases, (8) and (11) are like (9) and (12); and, in (7), Thomson's view is close to mine. I conclude that, in three of these cases – (8), (9), and (10) – Thomson's view may better fit most people's intuitions. But these are cases where our views only weakly disagree; and, in some of the other cases of weak disagreement, our intuitions may favor me. In the more important cases, where I and Thomson strongly disagree, Thomson's view conflicts sharply with most people's intuitions.

Our intuitions may be mistaken. So we must now consider Thomson's arguments.

3 Thomson's Arguments

Thomson argues by elimination, or by rejecting other views. She does not directly discuss my view; but, in rejecting other views, she may give sufficient grounds for rejecting mine.

Consider first her remarks about the Brain Criterion. Thomson appeals to a version of

> Case (11): My brain is remodeled so that it becomes like Thomson's brain, and is then transplanted into Thomson's empty skull.

On the Brain Criterion, I survive, though with Thomson's body, and psychologically continuous with her. Thomson says that, in such a case, we would all agree that I do *not* survive. If that is so, sameness of the brain cannot be sufficient for personal identity. Thomson might also appeal to Case (13), in which my brain is replaced with an exact replica. On the Brain Criterion, in such a case, I do not survive. On Thomson's view, I *do* survive, though with a different brain. If that is so, sameness of the brain cannot be necessary for identity.

I agree that, in these cases, the Brain Criterion gives the wrong answers.[12] Where Thomson claims these answers to be determinately false, I claim only that they are not determinately true. Suppose that we accept Thomson's stronger claim. Should we then conclude, as Thomson seems to do, that the sameness of the brain is not even *relevant* to personal identity?[13]

That conclusion does not follow. Our examples involve three elements: sameness of the brain, sameness of the body, and psychological continuity. On what I shall call the *Majoritarian View*, none of these elements is either sufficient for personal identity, or necessary. But all three elements are relevant, since any *two* elements are sufficient. If some future person will have either my brain and body, or my brain and psychology, or my body and psychology, that person will be me. In all other cases, that person will not be me.

Compared with Thomson's, this Majoritarian View better fits most people's intuitions. These views differ only in Cases (1) to (6), and these are precisely the cases

where Thomson's view is least plausible. The Majoritarian View, in these cases, gives the same answers as the Narrow Psychological Criterion. And, compared with Thomson's, these answers are easier to believe. So, whenever these views differ, Thomson's view is harder to believe.

Despite its intuitive appeal, the Majoritarian view faces theoretical objections. Majority voting, though an excellent procedure, is seldom more than that. My present point is only this. On the Majoritarian View, the sameness of the brain is always relevant to personal identity. That claim cannot be challenged by appeals to cases like (11) or (13), since the Majoritarian View, in those cases, coincides with Thomson's view. And, in the six cases where that claim *is* challenged, because these views conflict, that claim is supported by most people's intuitions.

Consider next the Narrow Psychological Criterion. On this view, one of our three elements, the sameness of the body, never matters. The other two elements are each necessary for identity, and together sufficient. For some future person to be me, we must both be psychologically continuous, and have the same brain.[14]

Though Thomson rejects the Narrow Criterion, she says little to explain why. She may assume that, since we should reject the Brain Criterion, we can also reject all views that take the sameness of the brain to be relevant to identity. I have just explained why that does not follow.

Thomson might reply that, on the Narrow Criterion, the sameness of the brain is claimed to be, not merely relevant to identity, but necessary. And that stronger claim may seem to be refuted by a case like (13), in which my brain is replaced by a replica. This example does, I agree, cast some doubt on the Narrow Criterion. We may find it hard to believe that, as this criterion implies, if my brain were destroyed and a replica were grafted in its place, I would cease to exist. But this example cannot, I think, refute this criterion. Every view may have some counterintuitive implications. One example is Thomson's claim that, if my brain were destroyed and her brain were successfully grafted into its place, I would continue to exist. We may find that just as hard to believe.

Thomson suggests a different ground for rejecting the Narrow Criterion. She believes that, if we accept some version of Psychological Criterion, we should prefer the Wide Version. Both criteria agree that, for some future person to be me, we must be psychologically continuous. But, on the Wide Criterion, this continuity could have any cause. On the Narrow Criterion, it must have its normal cause: the continued existence of enough of the same brain. Thomson suggests that it cannot be important how such continuity is caused. If the cause is abnormal, why, she asks, is that "not good enough"?

I agree that, in an account of what matters – what has rational or moral significance – it is irrelevant how such continuity is caused. But it could still be true that, in stating a criterion of identity, we should require that such continuity should have its normal cause, the continued existence of the same brain.[15] Thomson says that she cannot understand why we might prefer the Narrow Criterion, which makes that requirement. But the rest of her paper provides one answer. She advances arguments against the Wide Criterion, and some of these arguments do not apply to the Narrow Criterion. If Thomson's arguments are good, they give us reason to prefer that view.

Thomson also appeals to

> Case (14): I take some drug which causes my brain to "imprint" on my liver. After the rest of my body is destroyed, my liver is transplanted into Thomson's body, where it causes the remodeling of Thomson's brain. The resulting person is psychologically continuous with me.

Such a case, Thomson suggests, challenges the Narrow Criterion. Suppose we believe that, if my brain were transplanted into Thomson's body, and carried with it my psychology, I would continue to exist. Thomson suggests that, if the transplanting of my liver had the same effects, we should take the same view. We should believe that, here too, I would continue to exist. We should therefore drop the requirement that, to constitute identity, psychological continuity must have its normal cause. Nor could we plausibly require that, even if the causal story is abnormal, it must involve the continued existence of *some* part of the body. Why, she asks, should *that* be required? We should move to the Wide Criterion, which allows such continuity to have any cause.

Thomson's example does not, I think, support these claims. As Thomson remarks elsewhere, on any plausible criterion of identity, there could be borderline cases, where our questions would have no clear answer.

On Thomson's own view, for me to continue to exist, my body must continue to exist. Consider the cases in my Physical Spectrum, in which different proportions of my body would be destroyed and replaced. In some of these cases, on Thomson's view, it would be indeterminate whether I continue to exist. But, as Thomson would agree, such cases do not cast doubt on her view. Thomson's imagined Case (14) is of this kind. If we accept the Narrow Criterion, we reject Thomson's claim that, for me to continue to exist, my body must continue to exist. It is enough, on our view, that there should continue to exist one part of my body: the part which supports consciousness, and carries with it psychological continuity. We believe that part to be my brain. Thomson asks us to imagine that my liver, amazingly, acquires some of the properties of a brain. If we accept the Narrow Criterion, we may be puzzled by this example. But, as before, such a case casts no doubt on our view. Life can be distinct from death, though there are some cases in between.

We can now turn to Thomson's main arguments, which are directed at the Wide Criterion. I shall ask whether, if these arguments were redirected, they would have force against my view.

Supporters of the Wide Criterion, Thomson says, sometimes appeal to the claim that we can imagine switching bodies, or surviving without a body. Thomson challenges these appeals to what we think we can imagine. She may be right. But these remarks do not apply to my view.

Thomson also appeals to our intuitions. She discusses a case like (8), in which, by remodeling Robinson's brain, surgeons make the resulting person psychologically continuous with Brown, whose whole body has been destroyed. According to the Wide Criterion, the person with Robinson's body would now be Brown. Thomson calls it "odd idea" that "tinkering" with Robinson's brain could have that result. To apply this objection to my view, we must change the example. We must suppose that, rather

than remodeling Robinson's brain, surgeons destroy this brain, and transplant Brown's brain into Robinson's empty skull. On my view, the resulting person would be Brown. That is not an odd idea. It is what most of us, on reflection, believe. If that idea is wrong, we need some other ground for thinking so.[16]

4 The Extrinsicness Objection

Thomson then appeals to what Williams calls the "reduplication problem." If we accept the Psychological Criterion, we cannot exclude the possibility that there might be psychological continuity between a single person at one time and two different people at some other time. But one person cannot be the same as each of two different people. So our criterion must be stated in a way that avoids that implication. My original statement was:

> X today is one and the same person as Y at some past time if and only if X is psychologically continuous with Y . . . and there does not exist a different person who is also psychologically continuous with Y. (*RP*: 207)

Thomson objects that, in this formulation, our criterion is circular, since it assumes that we can tell whether Y is continuous with some person *different* from X. A criterion of identity should explain what makes people the same or different; so it should not assume an answer to this very question.

One reply would be to claim that, in describing personal identity *over* time, we are entitled to assume that we can distinguish people who exist at the *same* time. But my reply was to revise this criterion, so that its final clause becomes "and this continuity does not take a *branching* form." Such branching can be described without assuming answers to questions about personal identity.

Thomson's main objection is that, however we revise this clause, the Psychological Criterion makes identity depend, implausibly, on accidental and extrinsic facts. On this criterion, whether some future person will be me does not depend only on the intrinsic nature of the relation between us. It also depends on whether that relation holds *elsewhere*. Judgments of identity, Thomson claims, should not depend on such extrinsic facts. Let us call this the *Extrinsicness Objection*.

This objection has been advanced by several writers; and several others have suggested answers. Since these answers seem to me sufficient, my comments will be brief.

First, I believe that we should reject this objection's premise. Truths about identity can depend on extrinsic facts. And, when we see why that is so, it should not seem puzzling.

To illustrate this objection, let us consider two of my imagined cases. In the *One-Sided Case*, the right half of my brain would be destroyed, and the left half would be transplanted into the empty skull of another similar body. In the *Two-Sided Case*, or *My Division*, both halves would be transplanted. In both cases, someone would wake up with the left half of my brain. Let us talk of the *left-brained person*, where that phrase is not a name, but only a description. On the Psychological Criterion, in the

One-Sided Case, the left-brained person would be me. In the Two-Sided Case, the left-brained person would not be me. He would be a new person, who can be named *Lefty*.

Assume that what actually happens is the Two-Sided Case. And suppose that you are Lefty. On the Psychological Criterion, if what had happened was the One-Sided Case, you, Lefty, would never have existed.[17] Several writers think that absurd.

We should state more precisely what these people think absurd. There is a similar-sounding claim which we would all accept. If things had gone differently elsewhere, that might have caused things to go differently here, in a way that affected you, or even prevented your existence. But that is not the claim that we are now considering. We can suppose that, if what had happened was the One-Sided Case, in which the right half of my brain was destroyed, this would have had *no* effects on what happens to, or in, the left half of my brain. In both cases, this half-brain would be transplanted into the very same body, and the resulting left-brained person would have the very same, or at least exactly similar, thoughts and experiences. What is thought to be absurd is the view that, in the One-Sided Case, that person would be me, while, in the Two-Sided Case, he would *not* be me, but Lefty. If the history of the left-brained person is, in both cases, just the same, how could events elsewhere make a difference to who this person is?

Properly understood, this conclusion is not, I think, absurd. It will help to consider a simpler pair of cases: those involving Hobbes's famous *Ship of Theseus*. In *Case One*, this ship is dismantled, plank by plank, and is later reconstructed by antiquarians. In *Case Two*, when each plank is removed, it is replaced, so that a working ship continues to exist, though it becomes entirely made of new planks. On what Nozick calls the *Closest Continuer* view, in Case One, the antiquarian's ship *is* the original ship, which has been reconstructed. In Case Two, since the continuously working ship is a closer continuer, it is claimed to be the original ship, and the antiquarian's ship, though made of the original planks, is here claimed to be a different ship.

Wiggins objects that such a view

> licenses the following as a possibility: we could walk up to the antiquarian's relic, seen as a candidate to be Theseus' ship, and say that, but for the existence of its rival [the continuously working ship] it would have veritably coincided as a ship with Theseus' original ship. But the idea that in *that* case it would have been Theseus' very ship seems to be absurd.

This view is absurd, Wiggins suggests, because it implies that "anything might have been a numerically different entity from the entity it actually is" (1980: 95).[18]

Such a view need not have that implication. On this view, we need not claim that this *very ship*, though in fact different from the original ship, might have been identical to that ship. Our claim could be re-expressed as follows. Though this set of planks actually constitutes a different ship, if there had not been the continuously working rival ship, this *very set of planks* could have been correctly claimed to be constituting, here and now, the reconstructed original ship.

Similar remarks apply to the One-Sided and Two-Sided Cases. In both cases, there would exist the very same left half-brain, in the same body, and there would occur what would be, in one sense, the very same thoughts and experiences. On the

Psychological Criterion, our claims are these. If my right half-brain is destroyed, this left half-brain and body, and these thoughts and experiences, can be correctly called mine. If my right half-brain is not destroyed, this left half-brain and body, and these thoughts and experiences, cannot be called mine, but must be described as those of a new person. What happens elsewhere does not, mysteriously, affect what happens here. In the most straightforward sense, the very same things happen here. What events elsewhere affect is only the *label* that we can apply to what happens here. That is no more puzzling than the fact that, if my father has another child, though in a way that has no effects on me, that could make it true that I am not my father's only child. Whether I am my father's only child is partly an extrinsic fact: it does not depend only on my father's relation to me. Why should not the same be true of our identity over time?

Some would reply that, while such a claim can be true of a genuine *relation*, like being the only child of, it cannot be true of numerical identity. That is not a genuine relation, since it does not hold between *two* items. Or at least, if identity is regarded as a relation, it must be wholly internal. That *I am me* must be a truth that is wholly about me.

If these remarks fail to convince, there is a different reply. Thomson takes the Extrinsicness Objection to count in favor of her own view. But that is not so.

We can first consider how, if we accept the Narrow Criterion, we might try to avoid this objection. On this view, for some future person to be me, that person must have enough of my brain. Wiggins once suggested that, by "enough," we could mean *more than half*. That would avoid the reduplication problem, since there could not be two future people who each met that requirement. But, with that revision, the Narrow Criterion would be unacceptable. People could survive, as Thomson notes, even if they lost the use of half their brains.[19]

Similar remarks apply to Thomson's Bodily Criterion. For Thomson's view to avoid the reduplication problem, it must similarly claim that, for some future person to be me, that person must have more than half my body. But that requirement is also unacceptable. If half my brain and body were destroyed, but the remaining half continued to function and to support consciousness, what was left would still be the same body, and we cannot plausibly deny that I would still exist. Thomson herself claims that a body might continue to exist if it retained its more important organs, even if *less* than half this body continued to exist, because more than half had been replaced. We are now supposing that a full half continues to exist, and that there is no replacement. Nor does a body need to retain all of its more important organs. Just as we need only half a brain, we need only one lung, or kidney, and only half a liver. In the case of some vital organs, such as the heart or stomach, someone with only half a body would need reconstructive surgery, and some artificial aids.[20] But Thomson should agree that, if half this person's body continued to exist, and function, this person would still exist.[21]

I conclude that, even on Thomson's view, personal identity could depend on extrinsic facts. If my body was divided, and only one half survived, the resulting person would be me.[22] If both halves survived, the two resulting people could not each be me. So, when we ask whether each is me, the answer depends upon what happens to the other.

Since Thomson's own view faces the Extrinsicness Objection, this objection cannot support her view. Thomson might reply that, when applied to the Bodily Criterion, this objection has less force. She might repeat Williams's remark that, when bodily continuity takes a branching form, that fact is not, in the same objectionable sense, extrinsic. But the same could be said in defense of the Narrow Criterion, or the Brain Criterion, or my view. If the dividing of the body is not an extrinsic fact, nor is the dividing of the brain.

Every plausible criterion, it seems, faces the Extrinsicness Objection. That supports my earlier claim that this objection is mistaken.

5 Thomson's Ontological Argument

Thomson also argues:

(1) People are their bodies,

therefore

(2) X and Y are the same person if and only if they have the same body.

I shall call this Thomson's *Ontological Argument*.

It may seem that, if this argument is sound, we must accept Thomson's view, and reject mine. But that does not follow. (2) is only the first part of Thomson's Bodily Criterion. Thomson adds

(3) If we put one person's brain into another person's body, that would not make this a different body.

Even if we accept (2), we might question (3).

Reconsider our first case, in which my head is grafted onto the rest of Thomson's body. The resulting body, Thomson assumes, would still be hers. In my earlier discussion of this case, I accepted that assumption. I claimed that, despite having Thomson's body, the resulting person would be me. If people *are* their bodies, that claim would be undermined. If we wish to defend my view about this case, we would have to claim, instead, that the resulting person would have *my* body. Nor would it be enough to claim that, since this person would be me, what used to be Thomson's body would have *become* mine. If I am my body, our claim would have to be that, if my head were grafted onto the rest of Thomson's body, the resulting body would be the one that has *always* been mine.

To defend that conclusion, we would have to claim

(4) X and Y have the same body if and only if they have the same brain.

Williams calls (4) absurd (1973: 77). That dismissal is, I think, too quick. In trying to defend (4), we might appeal to my original version of this case. I supposed that, after

my head is disconnected from the rest of my body, it is supported by a heart-lung machine, and my brain continues to function. It is natural to believe that, if that happened, I would continue to exist. And, if we believe that people are their bodies, we might claim that the survival of my head amounts to the survival of my body, though in a severely diminished state. We might then claim that, when what is left of my body is attached to what is left of Thomson's body, my body continues to exist, though it now reassumes an expanded form. My body would now contain what were once parts of Thomson's body, just as it might in a more ordinary operation, in which I was merely given Thomson's heart and lungs. In this imagined case, much more of my expanded body would be new. But, as Thomson agrees, mere mass is not decisive. A human body continues to exist, she writes, if its most important parts continue to exist, and function. On the view we are now considering, there is one such part: the brain.

This view is not, I think, absurd. But I would agree with Williams that it can be reasonably denied. So it cannot provide a wholly satisfactory reply to Thomson's argument.

6 Are We Our Bodies?

How else might we reply? According to Thomson's argument,

(1) People are their bodies,

therefore

(2) X and Y are the same person if and only if they have the same body.

I have now agreed that, if we accept (2), we cannot convincingly reject the rest of Thomson's view. Since this argument is valid, our only other possible reply is to question (1).

Thomson says little to defend (1). She mentions only one other possibility: that we are immaterial substances, such as Cartesian Egos. Her assumption seems to be that, if we are not immaterial substances, we must be our bodies, since there is nothing else for us to be.

There are several other possibilities. Nagel suggests that what we are, essentially, is our brains (1986: 37–43). Like the view that a body's identity depends on that of its brain, Nagel's view is, I believe, plausible. We could explain how, even if we are essentially our brains, the rest of our bodies could also be regarded as parts of us. A similar view is well defended by McMahan (2002: 3–94). I shall here consider other possibilities.

If we are not our bodies, or our brains, what else could we be? I defended the unexciting view that we are persons, or human beings: entities that *have* bodies, and *have* brains. On this view, we are distinct from our bodies and our brains, though we are not separately existing entities, such as immaterial substances.

This view, I claimed, stays closer to our actual concept of a person. We often distinguish between ourselves and our bodies. If we thought we *were* our bodies, we

would not claim that we *have* bodies, since bodies cannot have themselves. And there are many other properties that we ascribe to ourselves, rather than our bodies. Thus we do not regard our bodies as being witty, or solving crossword puzzles, or remembering our childhood.

After mentioning such "facts of usage," Thomson replies that, in other ways, we seem not to distinguish between ourselves and our bodies. For example, rather than saying that our bodies weigh some amount, we say that *we* weigh that amount. This reply is not, I think, sufficient. There is an asymmetry here. Even if we distinguish between ourselves and our bodies, it is easy to explain our claim that we weigh what our bodies weigh. That could be like our claim, when our coat has been splashed with mud, that *we* have been splashed with mud. If something is true of X when and because it is true of Y, that need not be because X *is* Y. X may own Y, or consist in Y, or be related to Y in some other way. If instead we believe that we *are* our bodies, it is hard to explain why we deny that our bodies have all our properties. If we believe that Einstein was a genius, and that Einstein was his body, why do we deny that Einstein's body was a genius?

Our question, however, is not what we do believe, but what we should believe. Even if we do distinguish between ourselves and our bodies, Thomson may be right to reject that distinction.

In asking whether she is right, it will help to return to a similar but simpler question. What is the relation between a statue and the lump of matter of which it is composed? According to some writers, if we claim that Cellini's Venus *is* a lump of gold, we would not be using the "is" of identity. We would mean only that this statue consists in this lump, or that the statue and the lump are made up of the same matter. On this view, though this statue and this lump are very closely related, they are distinct entities.

As an account of what we mean, this view may be too precise. But the distinction which it draws is, I believe, defensible. If that is so, we can also defensibly distinguish between ourselves and our bodies.

To defend the first of these distinctions, I appealed to a standard argument. If X and Y were one and the same entity, we could not destroy X without destroying Y. But, if we melted down Cellini's Venus, we could destroy this statue without destroying this lump of gold. So these cannot be the same.

Some reply that, though we can destroy the statue without destroying the lump, that is not because the statue is a distinct entity. On their view, the concept of a statue is like the concepts of a child, or tadpole, or butterfly. It is what Wiggins calls a *phased sortal*, a term which applies to some persisting entity, in the present tense, only during part of its history, while it has certain properties. The statue *is* the lump of gold, but this lump is only a statue *while it is shaped like Venus*.

One problem for this view is what, when we use a phased sortal, we refer to an entity that may continue to exist even after it loses the properties that our term ascribes to it. Thus, when children grow up, they do not cease to exist; they merely cease to be children. The parallel claim would be that, when we melt down our statue into a shapeless lump, it continues to exist, and merely ceases to be a statue. But most of us would claim that, even though the lump continues, the statue ceases to exist. If the statue *is* the lump, as the phased-sortal view maintains, that would be a contradiction.

There is a greater problem for this view. Suppose that we do not melt down the whole statue. We heat the statue from within, and thereby manage to extract, in a single malleable lump, almost all the gold. This process does not affect our statue's surface, which amounts to 1 percent of the gold. When we have removed the statue's core, we give it a cheaper iron core instead. If our original lump continues to exist, it must be the gold core which we have removed, after detaching its outer layer.[23] We might now destroy this lump, by turning it into coins. But our statue still exists. With the same shape, and visible surface, the statue has not suffered at all. So the statue cannot be the lump, as the phased-sortal view maintains. Since it might continue to exist even after the lump was destroyed, it must be a distinct entity.

This argument appeals to what Wiggins calls

> *The Life Histories Principle*: X cannot be the same as Y if they could have different histories.

To apply this principle, we must produce a convincing case in which the histories of X and Y diverge. We began by claiming that the lump's history might be longer than the statue's. It was objected that the statue is merely the lump while-Venus-shaped. We therefore claimed that the statue's history might be longer than the lump's. These two examples, I believe, refute the phased-sortal view. X and Y cannot be the same if either might continue to exist after the other has been destroyed.

Our distinction might be challenged in another way. Some writers appeal to

> *The Exclusion Principle*: There cannot be two objects that are wholly composed, at the same time, of the same matter.

When applied to objects of the same kind, this principle can seem compelling. Suppose our statue left Cellini's workshop as a 50-pound lump of gold. It would be hard to defend the view that we should regard this gold as *two* 50-pound lumps. What could make these *different* lumps? Nor could a single mass of gold be easily regarded as two different statues.

What we are defending, though, is a different view. On this view, though these 50 pounds of gold constitute two different objects, there is only one lump, and one statue. And, since these are objects of different kinds, we can explain why they are distinct. They have different criteria of identity. As our second example shows, the statue and the lump might cease to share all the same matter. The statue might survive, with a new iron core, after the lump has been extracted, and sold off for turning into coins.

This example gives us, I believe, sufficient reason to reject the Exclusion Principle. When two objects are of different kinds, they can remain distinct, in a quite unpuzzling way, even while they are wholly composed of the same matter.

It is sometimes claimed that, if we reject the Exclusion Principle, we can be led to absurd conclusions. If our statue is distinct from our lump of gold, and each weighs 50 pounds, why do they not *together* weigh 100 pounds? If we scratch the statue, why are there not *two* scratches, one on the statue, and one on the lump of gold? But such objections can be answered. As Williams says, when two objects share the same matter, we should not count that matter twice. That is why the statue and

the lump do not, together, weigh more. And, if we scratch the statue, we thereby scratch the lump, and vice versa. Just as they share their matter, and their surface, they share this scratch.

Though there are other arguments for the Exclusion Principle, none, I believe, succeeds. The most that can be claimed is that, for certain purposes, it is worth using a conceptual scheme which obeys that principle. But we are asking whether, within our ordinary scheme, we can defensibly distinguish between statues and the lumps of which they are composed. And the answer, I believe, is Yes.

In the same way, I believe, we can defensibly distinguish between ourselves and our bodies. And we can thereby answer Thomson's argument. Though we and our bodies share the same matter, we are not the same as them; and that is partly because we and they might have different histories. Just as our statue might continue to exist, with a new core, after the original lump had been destroyed, we might continue to exist, with a body below the neck, after the rest of our old body had been destroyed. That is what, when we consider such a case, most of us are inclined to believe. And Thomson's argument does not, I believe, show that we are wrong.[24]

7 Persons and Human Beings

Let us now take stock. Thomson argued:

(1) People are their bodies,
(2) If we transplant a brain from one body to another, that would not affect the identity of this second body,

Therefore

(3) If we transplant my brain into Thomson's body, the resulting person would be Thomson.

If we find it hard to believe Thomson's conclusion, we must reject one of her premises. Though we might question (2), it is enough that, as I have argued, we can defensibly reject (1).

Consider next a similar argument:

(4) People are animals of a certain kind: human beings,
(5) If we transplant a brain from one animal to another, that would not affect the identity of this second animal,

Therefore

(3) If we transplant my brain into Thomson's body, the resulting person would be Thomson.

This, which I shall call the *Animalist Argument*, has been advanced by several recent writers. If this argument is sound, it provides a different route to Thomson's conclusion.

Some writers would reject this argument's first premise. Following Locke, they distinguish between persons and human beings. These Lockeans need not claim that this distinction is already part of our conceptual scheme. They can propose their distinction as a refinement of that scheme. On their proposed view, persons are human beings only in the constitutive sense in which our statue is a lump of gold. According to these Lockeans, while the identity of human beings may depend on that of their bodies, the identity of persons, when regarded as distinct from human beings, depends instead on psychological continuity, with the right kind of cause. To adapt one of Locke's examples, if a prince's head were grafted onto the rest of a cobbler's body, even if the resulting human would still be the cobbler, the resulting person would be the prince.

Animalists, however, have a powerful response. Human beings are conscious animals, who can have thoughts and other experiences. When a person has some experience, Lockeans must admit that a human being has that very same experience. So, if Lockeans distinguish persons from human beings, they must admit that, on their view, our thoughts and other experiences are each had *both* by a person *and* by a human being. And that may seem absurd. As McDowell writes: "surely there are not *two* lives being led here, the life of the human being . . . and the life of the person" (1997: 237).

To assess this *Two Lives Objection*, it is essential to be clear about the kind of question that we are discussing. If Lockeans distinguish persons from human beings, they may have to claim that there are twice as many lives being lived. But, in making that claim, they should not be thought to be advancing a different *hypothesis*. They are not *disagreeing* with those who do not draw their distinction. The question is only which concepts we should use. If we do not distinguish persons from human beings, we can truly claim

(6) In each human body, only one life is lived.

Since the Lockeans do draw this distinction, they claim

(7) In each human body, two lives are lived.

If we wish to reject (7), we cannot simply appeal to (6). It may seem that, if (6) is true, (7) must be false. But that is not so. Unlike us, Lockeans distinguish persons from human beings. Since they use these concepts in a different way, their claim does not conflict with ours.

To reject (7), we must challenge the Lockean proposal. Thus we might argue that, even if statues are distinct from lumps of gold, and persons are distinct from bodies, it is incoherent to distinguish persons from human beings. Or we might argue that, when Lockeans make their proposal, they do not provide, for the entities which they call persons, an adequate criterion of identity. Or we might argue that, in some other way, the Lockean distinction is defective.

In making these remarks, I assume

(8) If the concept of an X is coherent, and the conditions for its application are fulfilled, we cannot deny that there are Xs.

According to some writers, (8) implies absurdly that objects can be caused to exist by mere stipulation. But that is not so. We may invent the concept of an X; but, except in very special cases, that does not ensure that there are any Xs. Mere stipulation cannot cause reality to be a certain way.

Other writers object that, if we accept (8), our ontology will become incredible. Suppose that I define a squigillion as the combination of any bull frog and the planet Jupiter. According to (8), we cannot deny that squigillions exist. That, they claim, is wildly implausible.

This objection is, I think, misguided. There may be some defect in my suggested definition. And the concept that I define cannot have a serious use. But our objection cannot be that, when I say that squigillions exist, my claim is incredible. As defined, squigillions are not strange, separately existing entities, like ghosts, or flying saucers. When I claim that squigillions exist, I do not revise our view about reality. My proposed concept merely allows us to redescribe some ordinary facts. If there can be constellations, or archipelagos, or metropolitan areas, there can be squigillions.

Return now to the Lockean proposal. If Lockeans distinguish persons from human beings, does that lead to incoherence, or some other defect in their conceptual scheme?

One charge of incoherence might be this. Locke defines a person as a rational and conscious being that is aware of its own identity. Snowdon objects that, on this definition, human beings qualify as persons (1990: 90). So, if Lockeans try to distinguish persons from human beings, they must admit that, on their view, every adult human body is shared by two *persons*, the one which is also the human being, and the one which isn't. That seems to undermine the whole point of this Lockean distinction. Some would add that, with this admission, this distinction violates the more compelling form of the Exclusion Principle. As Locke himself declared, two objects of the *same* kind cannot be simultaneously composed of the same matter.

Part of this objection can be answered. Even if human beings are also persons, they must be persons of a *different* kind from the ones who are not human beings. These two kinds of person have different criteria of identity. Though each pair of persons usually share one body, either might continue to exist after the other was destroyed. Thus, on our present assumptions, if my head were grafted onto the rest of Thomson's body, the surviving *animal-person* would be Thomson, but the surviving *pure-person* would be me.

The rest of this objection stands, however, and does count against the Lockean proposal.

There are other objections. As the case of a statue shows, it is not a problem that, on the Lockean proposal, each pure-person shares its body with a human being. But can these beings be claimed to share all of their experiences? How can a single mental event have two subjects? "In the same way," the Lockeans might say, "in which a single scratch can be a scratch on both a statue and a lump of gold." But this analogy may seem insufficient. And suppose these beings jointly think, "I am a human being." On the Lockean proposal, there must be two thoughts here, one of which is false. How can a single mental event be regarded as the thinking of two different thoughts? In the same way, Johnston suggests, in which the raising of a banner, by two people, might be regarded as two assertions of the sentence written on that banner. But this analogy may again seem insufficient.

It is not clear that, with such objections, we can show the Lockean proposal to be incoherent. Though McDowell raises the Two Lives Objection, he himself suggests that, in the actual split-brain cases, single experiences are shared by two subjects.[25] It is tempting to protest that, even if not incoherent, such claims are wildly implausible. But, as I have said, plausibility is not the issue.

For different reasons, though, I reject the Lockean proposal. This conceptual revision would be both difficult and confusing; and it would serve, I think, no useful purpose.

One obvious question is: Who should we take *ourselves* to be, the pure-persons, or the human beings? It may seem that, without an answer to this question, we cannot even think about ourselves. And suppose that, as a Lockean, I claim to be the pure-person. Olson objects that I might be mistaken. How do I know that I am not the human being? My chance of being right, Olson claims, is only one in two (1997: 107–7).

These objections can be partly met. In some cases, if some statement is ambiguous but would be true on either of its meanings, it can be counted true. Suppose I point to Cellini's Venus, and say merely "That weighs 50 pounds." My use of "that" might refer to either the statue, or the lump of gold. But, since both weigh 50 pounds, my claim is true. If we draw the Lockean distinction, most of our thoughts about ourselves could have the same harmless ambiguity. They would have the same truth value whether applied to the pure-person or the human being.

The Lockean distinction cannot be useful, though, except in the cases where it makes a difference. And, in these cases, it is no use. Return to my question whether I am the pure-person or the human being. The problem is not that, if I try to guess, I might get the answer wrong. I could not get the answer right. For my question to have an answer, my use of "I" must be unambiguous. It must refer only to the pure-person, or only to the human being. And, on the Lockean proposal, that condition is not fulfilled.

It may seem that, if my question has no answer, I could decide to *give* it an answer. I could decide to which of these two entities I intend to refer. That would indeed be possible in most cases of this kind. I could decide, for example, whether to refer to the statue or the lump of gold. But, when I think about *myself*, this is no solution. Even if I accept the Lockean proposal, I am not free to decide whether, when I use "I," I intend to refer to the pure-person, or the human being. Any use of "I" must refer to its user. And, on this proposal, any use of "I" has two users. Whenever I use "I," both the pure-person and the human being thereby refer to themselves. So, when I ask "Who am I?" that is not a single question, whose ambiguity might be removed. I – or we – have asked two questions, with two answers. As Snowdon rightly says, when we try to think about ourselves, the Lockean proposal induces vertigo (1990: 95).

8 The Identity of Animals

Return, now, to our motive for considering this proposal. The Animalists argue:

(1) Persons are animals of a certain kind, human beings,
(2) If we transplant a brain from one animal to another, that would not affect the identity of this second animal,

Therefore

(3) If we transplant my brain into Thomson's body, the resulting person would be Thomson.

If we find (3) implausible, as most of us would, Lockeans tell us to question (1). We do better, I believe, to question (2).

Compared with Thomson's, the Animalist Argument is, in one way, stronger. Its first premise is harder to reject. We are already inclined to deny that people are their bodies; and, to defend that distinction, we need only claim that two different objects could share the same matter. We are less inclined to distinguish between persons and human beings; and, to defend that distinction, we would have to claim that two different conscious beings could share all of their thoughts and experiences. That is much more questionable, even if offered as a conceptual revision.

The Animalists' second premise is, in contrast, weaker than Thomson's. Thomson claims that, if we destroy some body's brain, and replace this with a different brain, the result would still be the same body. The Animalists claim that, in such a case, the result would still be the same animal. This claim is easier to reject.

We can start by imagining animals whose bodies mostly consist of their brains. Consider *the Mekon*, the imagined ruler, in *Dan Dare*, of a Martian species. The Mekon's head is about nine-tenths of its whole body. Must we believe that, if we replaced the remaining tenth, the Mekon would cease to exist?

As before, mere quantity is not decisive. But, by inverting the normal proportions, we can see more clearly what the Animalists might claim. Why might the Mekon's identity go with the small fraction of its body that lies below its neck? The Animalist answer would, I think, be this. In that part of the Mekon's body, there are many organs. The Mekon's brain, however large, is only a single organ. When we replace other single organs – such as a heart or kidney – no one doubts that the same animal continues to exist. Why believe that, in this respect, the brain is different?

Before we discuss this Animalist claim, we can make some general remarks. We are asking whether, if some part of an animal's body were destroyed and replaced, the resulting animal should be counted as one and the same. This is not a scientific question. Our imagined cases would not involve some natural process, which might follow discoverable laws. These cases would all be artificially produced. Ours is a purely conceptual question.

Of the possible answers to this question, the simplest would appeal only to *how much* of the animal's body is replaced. But this, we have said, cannot be right. Some parts of the body, though smaller, are more important. Another answer might appeal only to *how many organs* are replaced. But that would also be too crude. Just as some parts of the body matter more than others, so do some organs. What kind of importance, though, should we be considering?

Organs are important, we might say, if they are *vital*: if, without them, the animal could not survive. The appendix, for example, hardly counts. But this suggestion is not much better. First, we are not asking whether, after our imagined operations, the resulting animal could survive. Since all removed organs would be replaced, the resulting animal would be just as able to survive. We are asking the different question

of *which* animal this would be. Second, we can imagine organs that were vital, but had little relevance to a judgment of identity. These organs might be like the kind of removable device, intended to prevent theft, without which the engines of some cars cannot function. The identity of these engines would hardly be affected by the replacement of such a simple mechanism. The same seems true of many vital organs, such as the wind-pipe. It seems true, indeed, even of such organs as the heart.

We need, then, some further measure of importance. In asking what this might be, we face another general question. When they discuss animal identity, some writers assume that it should make no difference what kind of animal we consider. Olson reminds us that, as animals, we are more like oysters than angels. According to these writers, our proposed criterion should be universal. But this assumption is too simple. We are indeed, in certain ways, strikingly like oysters. For example, they and we have much chemistry and many genes in common. But, if the identity of animals partly depends on the importance of their different organs, our criterion must be sensitive to the kind of animal that we are considering. Not all animals have the same organs. And, in different kinds of animal, the same organ can have different importance.

Return now to our main question. When compared with other organs, can the brain be claimed to be overridingly important? In the case of certain animals, that claim could be reasonably denied. Many animals have hardly any brain, and some have none. But, in the case of animals like us, the brain is not merely one of several vital organs. Though it is only a single organ, it *is* special. It is the controlling organ: governing, in several ways, most of the body. Its relation to the body is like that of a nucleus to a cell. For animals like us, it might be the individuating organ, which determines our identity.

In asking whether that is so, it may help to reconsider our first case. At the start of my story, my head continues to function after the rest of my body is destroyed. When discussing Thomson's view, we asked whether that would amount to my body's continuing to exist. In considering the Animalist view, we have a different question. My surviving head is that of a conscious being, whose mental life is much like mine. Is this being an animal?

Suppose, first, that the Animalists say "No." They will then face a problem like the one with which they embarrass Lockeans. On the Lockean proposal, a person starts to exist during the early infancy of a human animal. It would be easier for Lockeans if, at that stage, the animal retired from the scene, leaving its body for sole occupancy by the person. If Lockeans could make that claim, they would avoid the Two Lives Objection. But, as the Animalists object, we have no reason to accept that claim. When the animal starts to be rational, and aware of its own identity, why should it, conveniently, cease to exist?

We are now supposing that, on the Animalist view, my surviving head is that of a conscious being, which is not an animal. Similar remarks apply. When my head is grafted onto the rest of Thomson's body, there will again be a living animal. What then happens to this conscious being? It would be easier for the Animalists if, at that stage, this being retired from the scene. But why should we think that? Why should this being cease to exist, leaving the animal as the new and sole possessor of my head? But, if this being continues to exist, the animal will have to share her thoughts with it. So, like the Lockeans, the Animalists face the Two Lives Objection.

Suppose next that, according to these Animalists, my disconnected head continues to be that of a living animal. This animal must still be me. Do I then cease to exist when my head is grafted onto the rest of Thomson's body? The answer should in part depend on the view we take of the rest of Thomson's body. We can suppose that, after Thomson's head has been destroyed, the rest of her body has its functioning maintained by artificial means. Would this headless body still be a living animal?

Animalists, I believe, should answer "No." As Van Inwagen claims, there is an important difference between the head and the rest of the body (1990: ch. 15). For my head to continue to function, all that it needs is a fairly simple pump, keeping it supplied with blood. There would then be a conscious being, who could see, hear, talk, and finish my book. For a headless body to continue to function, more assistance would be needed. There must be an artificial surrogate for a lower brain, sending down the spinal column much highly complex electrical information. This torso and these limbs are very far from being a self-sustaining, unified organism. They cannot, on their own initiative, do anything. As Animalists should agree, they do not constitute a living animal.

If Animalists accept these claims, they should change their view about the rest of our story. In the intermediate stage, we are now assuming, I survive as a living animal, though all that I have left is my head. Thomson does *not* survive in her headless body. So why believe that, when we connect my head to the rest of Thomson's ventilated corpse, the living animal dies, and the dead animal comes to life again? When my brain is attached to Thomson's heart and lungs, why should I, conveniently, cease to exist? But, if I survive, and Thomson comes to life again, the Animalists again face the Two Lives Objection. On their view, I and Thomson will both have to share all of our thoughts and other experiences. While such a claim might not be incoherent, there is a simpler, better view. We can claim that, when Thomson's brain is destroyed, that, for her, is the end.

The Animalist Argument, I conclude, does not undermine the view defended in my book. On that view, if our brains will continue to exist and function, we shall continue to exist. What happens below our necks, though it may affect the experiences that will be directly related to our brains, cannot cause these experiences not to be ours. That is what, on reflection, most of us are inclined to believe. And Animalists can believe this too.

9 Empty Questions

This defense of my view is, in a way, misleading. It suggests that, in this imagined story, there are different possibilities, any of which might be what really happens. But that, I believe, is not so. We have asked six questions:

1 If my head continued to exist and function, after it was disconnected from the rest of my body, would that amount to the continued existence of my body?
2 Would my surviving head still be the head of a living animal, or would it be the head of a conscious being that was not an animal?

3 If the latter, what would happen to this conscious being when my head is grafted onto the rest of Thomson's body?
4 After this second operation, would the resulting body be the same as my original body, or the same as Thomson's original body, or neither?
5 Would the resulting animal be the one that originally had my body, or the one that originally had Thomson's body, or neither?
6 What would happen to Thomson, and to me? Could either, or both, or neither expect to wake up again?

These questions, I believe, are all empty. None describes different possibilities. In trying to answer these questions, we learn nothing more about what, in such a case, would really happen. There is nothing more to learn.

Here is what would happen. My head would be disconnected from the rest of my body, and would then be grafted onto the rest of Thomson's body. There would be a series of thoughts and experiences, directly related to each other, and to my brain. If I had to decide whether to accept these operations, that is all I would need to know. What matters is reality, not how it is described.[26]

If that is so, it would be irrelevant whether, for me too, this would be the end. That would be merely another redescription. But, if you were in my position, and you cared what things were called, you could tell yourself that this would *not* be the end. As I have argued here, if your head were successfully transplanted, it would be correct to call the resulting person you. So you could expect to wake up again.

We may think that, in expecting that, you would be expecting something different to happen. But, if we do not have souls, what could the difference be?

10 Does Our Essence Matter?

It is worth adding some remarks about the Lockean proposal. I have argued that, to defend our view about such imagined cases, we need not try to distinguish between persons and human beings. But Lockeans have other grounds for drawing that distinction. They would claim, for example, that even if a fetus is a human being, it is not a person. Or consider someone who, because of brain damage, is irreversibly unconscious. Lockeans claim that, in such a case, even if the human being is still alive, the person has died.

Suppose that, to avoid the Two Lives Objection, Lockeans concede that persons are human beings. They must then withdraw the second of these claims. They must admit that, since the brain-damaged human being is still alive, so is the person. But they might appeal instead to a less ambitious form of the Lockean proposal. They might use "person" as a phased sortal: a term which applies to us, in the present tense, only while we have certain properties. On this use, just as we are children only while we are immature, we are persons only while we are rational and self-conscious. These Lockeans could then claim that, when a person has become irreversibly unconscious, though this person is still alive, he or she has ceased to be a person. That is not a contradiction, since it is merely like the claim that, when children have grown up, they have ceased to be children. Similar remarks apply to the Lockeans' other claim.

They could say that, just as some human beings are *no longer* persons, fetuses are *not yet* persons.

Compared with the bolder Lockean distinction between persons and human beings, this proposal stays closer to our actual use of the word "person." But these Lockeans might now insist that, though persons are human beings, they have a special moral status only while they are persons. As a claim about moral agents, that is uncontroversial. Only persons can have obligations, or be morally responsible. But, in some of its other implications, the Lockean claim is controversial. Consider, for example, the morality of killing. On this neo-Lockean view, while it is nearly always wrong to kill persons, there is much less objection to the killing of those human beings who, at the time, are not persons. Such a view could support some kinds of euthanasia, and most abortions. And, since this view implies that newborn babies are not yet persons, it could support infanticide as well.

Many of us would reject this view. In explaining where it goes wrong, we might say that, if "person" is used as a phased sortal, this concept must be of secondary importance. While we are persons for much of our lives, that is not what we *are*. As these Lockeans admit, we are human beings. Perhaps, as a newborn baby, I was not yet a person; but, in killing that baby, we would have been killing *me*. That, we may claim, is what counts.

If we make these claims, however, we must answer further questions. Consider a fertilized ovum, or a zygote. Are *these* human beings?

Given the vagueness of the term "human being," these questions are indeterminate. But we are free to adopt a tighter definition. Suppose, first, that we decide to use "human being" so that it applies to any organism of our species. On this *biological* use, a fertilized human ovum – or at least an early brainless embryo – is a human being.

If we adopt this definition, we may need to reconsider some of our moral principles. Suppose that, when rejecting infanticide, we appeal to the principle that it is seriously wrong to kill innocent human beings. On the biological definition, this principle implies that the killing of a human ovum is seriously wrong. Most of us would not accept that implication. And we cannot be forced to do so merely by our adoption of this definition. Or suppose we believe that, if we could save from death one human being, or many, we have more reason to save the many. What if the single human were an adult, and the many were one-celled organisms in a Petrie dish? Would it then be better to save the many? Most of us would answer "No."

If we give these answers, our concept of a human being will lose much of its significance. We shall no longer believe that all human beings have the same moral status. We shall give a special status to those human beings who are sufficiently developed to have brains, and other organs. To express these beliefs, it would help to use some term as a phased sortal, so that it applies to human beings only while they do have brains, or other organs. That is how some people already use the term "human being." But, if we have adopted the biological definition, we shall need some other phrase. We might say "*fully developed* human being." That would be the term which, on our view, has most moral significance. On this view, it is not wrong to kill those human beings who have not yet have developed brains.

Suppose, next, that we reject the biological definition. We adopt the other use just mentioned, which treats "human being" as a phased sortal. On this use, just as acorns

are not oak trees, and eggs are not chickens, fertilized ova are not human beings. We agree that, as human beings, we are organisms of a certain species. But just as a child is not yet an adult, and an acorn is not yet an oak tree, a fertilized ovum is not yet a human being. Given this use of "human being," we need not revise our moral principles. We can still claim it to be seriously wrong to kill any innocent human being, while it is a human being.

The two views just described are, in substance, the same. The difference is only that, where one says "human being," the other says "*fully developed* human being."

Neo-Lockeans may now smile. These views are, in their structure, just like the neo-Lockean proposal. On all three views, though we are essentially organisms of a certain species, that is not what counts. These organisms have a special status only *while* they have certain properties. The difference between these views is only in the properties they single out.

On the two views just sketched, these properties are having a brain, and other organs. Human organisms acquire these properties by about the fourth month of pregnancy; and, once acquired, these properties are hard to lose. According to the neo-Lockeans, the relevant properties are those of being rational and self-conscious. These properties are more restrictive, and are easier to lose. They are not had either by newborn babies, or by some brain-damaged human beings.

In choosing between these views, we should distinguish different moral questions. Thus, if we are considering the infliction or relief of suffering, we may think it irrelevant whether the sufferer is now a person, or even a human being. But in other contexts, such as those involving autonomy, or responsibility, it is personhood which counts. Questions of life and death are harder to decide. If some animal is not a person, or not yet a person, it is not, in a strong sense, living a life, so it may seem less important whether its life is cut short. But, in the presence of a newborn baby, that may be hard to believe.

I am not concerned, here, with the choice between these views. My point is only that, as I have said, they are views of the same kind. On the neo-Lockean proposal, though what we *are* is human beings, we have a special moral status only while we are persons. If we take either of the views just sketched, we cannot reject this proposal in the way described above. We cannot claim that, since we are essentially human beings, the concept of a person must have only secondary importance. On these other views, though what we *are* is human organisms, we have a special status only while we are human beings, or full human beings. As on the neo-Lockean proposal, what matters is not what we are, but whether, at the time, we have certain properties.

This provides a different sense in which our identity is not what matters. On all three views, I was once a fertilized ovum. If doctors had destroyed that ovum, they would have been killing *me*. But that, we believe, is not what counts.[27]

Notes

1 Derek Parfit, *Reasons and Persons*, henceforth *RP* (Oxford: Oxford University Press, 1984), Part III.

2 I also took seriously a Cartesian Criterion, which appeals to the continued existence of the same soul. But I shall assume here that we can reject that view, if only for lack of evidence.

3 I corrected this and related mistakes in the reprinted versions of *RP* (as noted on page x).

4 Those who don't think otherwise include, however, some other excellent philosophers, such as Bernard Williams, Paul Snowdon, Quassim Cassam, and Michael Ayers.

5 See this vol., p. 171. Thomson says that she wants her body to "continue to function in the ways in which living human bodies function when they support consciousness."

6 For answers to some objections here, see Unger (1990: ch. 8).

7 There may be some exceptions. For example, if there would be two such future people, it may be determinate that neither would be me.

8 Or, in a mirror-image of (1), which need not be separately listed, my body, but someone else's brain and psychology.

9 We are now considering five criteria. Three we can call *pure*, since they appeal to only one of our three elements. These are Thomson's Bodily Criterion, the Brain Criterion, and the Wide Psychological Criterion (which can be taken to include the widest version). The Narrow Psychological Criterion and what I have called my view both appeal to a pair of elements: psychological continuity with, as its cause, sameness of brain.

10 For this to be implied, we must assume that there is no other living person who is also psychologically continuous with Williams.

11 This part of my view is *not* supported by our intuitions. But that is irrelevant here.

12 On my view, it is not determinate that, in (11), the resulting person is me, nor determinate that, in (13), this person is not me.

13 Except as one element in our judgment about the sameness of the body.

14 On this view, if some future person will have both my brain and psychology, that person will be me; and, if that person will have either a different brain, or a different psychology, that person will not be me.

15 In defending her claim that the method of causation cannot matter, Thomson imagines a case where transplanting a liver carries psychological continuity. "Was it not good enough on the ground that transplanting a liver doesn't *normally* cause psychology-transplanting. (Why should that matter?)" (this vol., p. 159). To assess Thomson's imagined case, we must know what kind of causal story she has in mind. She supposes that the drugs injected in the donor caused his brain to imprint on his liver in such a way that, when transplanted, the liver caused reprogramming of the recipient's brain. That makes the liver the equivalent of my Scanner and my Replicator.

16 In Thomson's example, my view implies that it is indeterminate whether the resulting person would be Brown, or Robinson, or a new person. Thomson might call that an "odd idea." But, as I have said, given how extreme such remodeling would have to be, it is not intuitively odd that there is indeterminacy here. (I abstract from the fact that, intuitively, indeterminacy is hard to accept. Thomson and I agree that, however counterintuitive this may be, personal identity could be indeterminate.)

17 That does not follow directly from the Sufficient Condition. We need the further claim that, in the Two-Sided Case, it cannot be true that each of the resulting people would be me, nor can we plausibly suppose that one of them might be me. It doesn't quite follow that neither would be me, since we might claim that this is a case of indeterminacy: that there is no answer to the question whether, in the Two-Sided Case, either person would be me.

18 Wiggins also talks of "the impossibility of conceiving of an entity's not being identical with that which it *is* in fact identical."

19 This vol., p. 175 (n. 11).

20 The most contentious detail here is that the lower brain stem may not be straightforwardly divisible. But that is not deeply impossible, as would be needed to challenge the relevance of this imagined case.

21 If, as she claims, I could survive with only half my brain, and I could survive with less than half my body, why could I not survive with half my brain and half my body? I am not appealing to transitivity here, in a questionable way. I am merely putting together the claims that, for me to survive, I only need half my brain, and I don't even need half my body.

22 For some needed refinements here, again see Unger (1990: ch. 8).

23 When some lump has been unevenly divided, we might claim that the larger part is the original lump, or we might claim that neither part is that lump. If the larger part is 99 percent of the original lump, the first of these claims would be more plausible. But we could not claim that the smaller part, which is a mere 1 percent, is the original lump.

24 As this analogy also shows, in order to distinguish between ourselves and our bodies, we need not reject Physicalism. Even if we had no purely mental properties, we could still be distinct from our bodies. Cellini's Venus, though distinct from the lump of gold, is wholly physical.

25 In such cases, McDowell believes, a single person has two series of thoughts and experiences, in each of which there is no awareness of the other series. But McDowell suggests that, in such a case, there are three different subjects of experience. There is the person, who has the thoughts in both streams. But there are also two subjects of experience, each of which has the thoughts in one of the streams. So each of these thoughts, on McDowell's view, is thought by two thinkers: one of these sub-personal subjects, and the person. (See McDowell 1990.)

26 For a brief defense of that slogan, see my paper "The Unimportance of Identity" (in Harris 1995).

27 Though published here for the first time, this paper was written in around 1990, and was intended to be a response to Judith Thomson's contribution to *Reading Parfit* (ed. Jonathan Dancy, Blackwell, 1997), reprinted in this volume as chapter 4.3. I apologize to Thomson for the way in which my writing of this and other responses held up the original publication of her paper. I have not revised this paper to take into account any later publications (though I have added a few references.).

References

Harris, Henry, ed. 1995. *Identity* (Oxford: Oxford University Press).

McDowell, John. 1990. "Peacocke and Evans on Demonstrative Content," *Mind* 99: 255–66.

—. 1997. "Reductionism and the First Person," in Jonathan Dancy, ed., *Reading Parfit* (Oxford: Blackwell Publishing).

McMahan, Jeff. 2002. *The Ethics of Killing* (Oxford: Oxford University Press).

Nagel, Thomas. 1986. *The View from Nowhere* (Oxford: Oxford University Press).

Olson, Eric. 1997. *The Human Animal* (Oxford: Oxford University Press).

Shoemaker, Sydney, and Swinburne, Richard. 1984. *Personal Identity* (Oxford: Blackwell Publishing).

Snowdon, Paul. 1990. "Persons, Animals, and Ourselves," in Christopher Gill, ed., *The Person and the Human Mind* (Oxford: Oxford University Press).

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

van Inwagen, Peter. 1990. *Material Beings* (Ithaca: Cornell University Press).
Wiggins, David. 1980. *Sameness and Substance* (Oxford: Blackwell Publishing).
Williams, Bernard. 1970. "The Self and the Future," *Philosophical Review* 79(2); repr. in, *Problems of the Self* (Cambridge: Cambridge University Press, 1973).
—. 1973. "Are Persons Bodies?," in *Problems of the Self* (Cambridge: Cambridge University Press).