

Succession, Reflection, and Resemblance

Max Rothman

What are we? What constitutes my identity as myself? I find that this case reveals some of our intuitions about such operations, and also may point us in the right direction. Consider:

Case (1): Brown is put to sleep¹, and his brain is taken out of his body. His body is cloned, the original is destroyed, and the brain is transplanted into this new body. Brown wakes up, and goes about his day. He may not even know that this operation happened to him. His experience is psychologically continuous.

Thomson (1997) posits the physical thesis: people are their bodies. In this case, then, Brown has ceased to exist along with his original body². What is left, this new being, is a new being, according to Thomson. She appeals to this theory as ‘the ordinary view’; I cannot help but find this absurd. This “new being” looks like Brown, talks like Brown, desires like Brown, maintains all of Brown’s relationships, etc. Any ordinary conception of what it means to be an individual person will almost inevitably appeal to these “traits”, which are ‘what that person is like’ in the world. Thomson calls the physical thesis, but it seems to be anything but. I posit that this example is enough to reject the physical thesis: my intuition says that any account of ‘what we are’ must take into account these cases of exact bodily replacement.

We might think, then, that psychological continuity is the hallmark of personal identity: that Brown continues to be himself in Case (1) because his experience of the world is uninterrupted, and he continues to think of himself as such. We might also like to appeal to the fact that Brown’s brain is preserved in this case, and that this fact is what guarantees his continuity of identity. We might differentiate here between the Brain Criterion, and the Psychological Criterion. However, making a distinction here seems like a confusion of biology here, along with a problem of naming. If we mean that someone truly “has the same brain” in its full literal sense, then necessarily they will have the same psychology- the “psychology” of a person at time t is essentially determined by the structure and properties of the brain at that time t . This is not to say that one could give person X’s brain to person Y’s body, and then rewrite (i.e. change neural connections and weights) that brain to match the brain state of person Z; however, in no sense would this brain be person X’s brain anymore: it is identical with person Z’s brain. The “rewriting” of the brain, and thus the psychology, is a fully physical operation that changes what the brain is. Say I take my window out of its frame, melt the glass down, add other materials to it, make a vase out of it, and then give it to my friend. Does she now have my window? Maybe in

¹ Simply to ensure psychological continuity. We can also imagine this all happens instantaneously, to the same effect, but that is perhaps less practical.

² A confusion, perhaps: is the brain not part of the body? How much of the body are we? Is “the body” our external physical appearance? Are we no longer us after significant plastic surgery? Problems abound...

some loose sense, but it is really more correct to say “she has a vase that was once my window”³. We have changed the physical and chemical structure of the thing- to still call it the same thing seems tenuous at best. An example:

Case (2): Brown is put to sleep. His body is cloned, and the original is destroyed. Brown wakes up, and goes about his day. He may not even know that this operation happened to him. His experience is psychologically continuous.

Here, his original brain is not preserved, but the brain that is in the body immediately before and after the operation is identical- this is what guarantees his psychological continuity. And, my intuition says, we would still like to call this person Brown. So, we can reject a Brain Criterion, both for its insufficient results and for its superfluity. It seems that we are most interested in a psychological criterion.

The main objection to this formulation of the psychological criterion follows from Case (3):

Case (3): Brown is put to sleep. His body is cloned into an identical copy. Each Brown wakes up, and goes about his day. They may not even know that this operation happened to him. Their experience is psychologically continuous.

Which of these 2 beings is Brown? Answering both seems to violate the principle of identity of indiscernibles, at least at the instant of their creation. And answering neither, which might seem plausible, runs into a practical problem: asking both people which one is Brown, both respond, “Me!”. This seems to give an undesirable outcome to the psychological criterion.

One response to this objection is to append a “no branching” clause to the formulation of the psychological criterion, that as soon as we have a branch in psychological continuity those beings are no longer identical with the being from before the branching. However, the branching criterion is subject to the extrinsicness objection. This is a rejection of the fact that identity can depend on extrinsic conditions. The only difference between Cases (2) and (3) is the number of clones that are made; each of the 3 beings is identical, and each thinks they are Brown. In Case (2), we are sure that the survivor is Brown, in Case (3) it is unclear. This problem could also arise from a case like Case (3), only one of the cloning operations goes wrong and only 1 being is alive after the operation. Thus, the identity of the survivor, or rather the possibility of the survivor having a determinate identity, is dependent on the extrinsic fact of the number of clones that are made. And this seems tenuous.

Consider the following:

³ Of course, if we believed in temporal parts then we would say that it is indeed the same thing. I reject temporal parts throughout.

Case (4): Brown is put to sleep, and his brain is removed. A clone of Brown's body is made, and Brown's original brain is put into the clone's body. An identical copy of the brain is fabricated, and placed into Brown's original body.

Here, all possible cases are in front of us when we ask the question: which one is Brown? Do we follow his body? His brain? If we follow psychological continuity, both would seem to be Brown. If we've rejected the physical thesis and the brain criterion, and do not think that two things can be numerically identical with one thing, we thus must say: either it is indeterminate, or neither is Brown. Again, when asked, both respond, "I am Brown". To say they are both Brown, while seemingly practical, violates what we might think of as the "sanctity of personal identity". To say that neither of them is Brown is seemingly to deny that they veritably think they are Brown, to deny that both beings have experiences of going to sleep and waking up without interruption, with all of the memories and feelings that one who we would ordinarily think was Brown would have. If we commit to saying neither is Brown, we might call these things Brown-like, but say that they are not identical to Brown before the cloning. We will come back to the consequences of this. And to think that it is indeterminate seems confused to me: how could it be that there is one that is Brown, we just can't tell which? This is to affirm that there is still, and always has been, something that Brown *is*. That would imply that there is something that makes Brown persist through the operation. What would this thing be? We seemingly have already rejected all of the usual candidates of identity.

If it is indeterminate which one is Brown, then clearly Brown has survived the cloning operation in some way. This is an appeal to a metaphysical 'identity pointer' that hangs across one of the beings. If it is true that one of the beings in a cloning case survives, as is the case if survival is determinate yet identity is indeterminate, then the original being should be happy for the cloning operation to occur: after all, he will survive. Yet consider the following case:

Case (5): We clone Brown, and let them both walk around for 10 minutes without letting them see each other. After this time, we kill the original.

Should Brown be happy in this case⁴? Should he consent to the cloning operation? Well, if it is indeterminate which survivor is him, then his happiness to consent would depend on if he thinks he is going to die or not. And, clearly, we can say that there is at least one thing that thinks it is Brown that is going to die. Telling Brown, "It depends" when he asks "Will I die or not?" seems

⁴ I have received some confusion and criticism for this phrasing of the problem: Why phrase it as a question of consent, or of happiness? This is a practical choice: grounding the metaphysical question of personal identity in a real choice makes it more tangible and consequential. If you really think that a clone would "be you", in whatever sense you like, then you should consent to the operation, since it would effectively be as if nothing had happened at all.

small consolation. To say that Brown's identity is indeterminate seems to be just as bad as saying that he survives as both beings.

This case seems to reveal quite a lot about our intuitions regarding cloning, and how we might feel about it. I agree with Parfit that trying to find "some criterion that would make identity coincide with what matters" in our concept of what a person is is a confused effort. What matters, then? David Lewis gives us a fairly plausible answer:

"I find that what I mostly want in wanting survival is that my mental life should flow on. My present experiences, thoughts, beliefs, desires, and traits of character should have appropriate future successors. My total present mental state should be but one momentary state in a continuing succession of mental states." (Lewis, 'Survival and Identity')

Yet, our interest in a future self goes beyond simply desiring that 'there is something in the future that thinks it is the same thing that I think I am'; to see this, look at Case (5) again. Even after the original Brown is killed, there is something that thinks it is Brown that persists and is psychologically continuous with the pre-cloning being. His life would go on, entirely as before: if he was never told he was a clone, he would never know, and neither would anyone else. And yet, there is also a thing that thinks it is Brown that dies. If you ask Brown before the cloning if he should consent to Case (y), he might say yes, thinking that there is a being that will persist as him through time, who will have no knowledge of the death of the original. But ask him again (the same being) after the cloning, perhaps right before you give him a lethal injection, and he will not want to die. Why did he change his mind?

I think it is rather simple: he does not want to face death, for his experience to stop. It is not as if there is a thing that is Brown that is 'housed' in the original body, that is all of the sudden ported over to the clone when the original is killed. Something akin to such a belief is what might motivate Brown to initially consent to the process of cloning. Yet, the metaphysical identity pointer would not move after the death of the original Brown, nor would "Brown" "wake up" in the clone's body after being killed: he is simply dead.

When the clones coincide in time, they are no longer psychologically continuous with each other, even though they are with the original Brown: their perceptual experience does not overlap. The result of the cloning is that there are two experiencing beings that think they are, and always have been, Brown: and, justifiably, each would be upset to die. For the Brown that is killed, Brown's life is over, his continuous perceptual experience has stopped. This idea of indeterminacy, then, seems insufficient to answer the question of "should I be happy to be cloned" when faced with a case like this. It cannot answer what seems to be "the important questions" in these situations of persistence of being.

So we might think that both are Brown, that his consent and subsequent hesitation in Case (y) derive from the fact that we will indeed both live and die. We might already find contradiction here, but consider further:

Case (6): Brown is put to sleep. His body is cloned into 2 identical copies, and the original is vaporized. One is condemned to eternal torture, and one will experience eternal pleasure.

If we think that Brown is both of them, then how could you sufficiently respond to his question: “But what will happen *to me*?” Clearly, he should like to *be* one and not the other, but how could he say which one *he is*, or which one *he will be*? It seems like Brown could not both be tortured and pampered at the same time.

I think this response is perfectly correct, but not for the reasons the objector might think. We are left with the case that neither is Brown, just something simply incredibly Brown-like; this, I believe, is correct, although not in the usual way. The objection found in Case (z) is actually confused: Brown cannot be tortured and pampered at the same time, not because he cannot be in both states at the same time, but rather because Brown cannot “be” in either state. The contradiction is not found in being two things at once, but rather in the formulation of the case. He cannot be anything at all, and certainly not something that persists in such a way that would lead to a contraction such as in Case (z).

This is to say: Brown cannot be otherwise than what he is by expressing “I exist”. You can only “be” the current continuous perceptual experience of yourself. The concept of personal identity, I claim, is indexical. We might even want to do away with the concept of “identity” here, for sake of clarity. Each of the beings we’ve referenced thus far, the original and the clones, are all correct in saying “I am Brown”; but this does not refer to the ‘metaphysical pointer’ sense of identity, a statement of numerical identity across time. All this statement means is an indexical: “I exist, and I think that I am Brown”. For the clone to say “I am Brown”, and for us to interpret it as saying “I am Brown (the one that existed before the cloning)” is for us to interpret the statement “I am here” as referring to a place that the speaker is not actually in. I cannot mean otherwise by “I am here” than the place I am actually located; if I try to do otherwise, I am being misleading. The same is true of the persistence of beings: “I exist” is only a statement of the continuous perceptual experience of being me.

To say that “Brown will be cloned and tortured” and “Brown will be cloned and pampered” is therefore not contradictory, because these two instances of Brown do not “point” to the same thing; indeed, they do not point to anything at all. Each Brown can make a true statement about their existence and their fate, as can we, but these statements cannot mean otherwise than exactly what they do: that such a thing is true *to them*.

Instead of Brown ceasing to exist at the moment of cloning, and 2 new beings that are not Brown taking his place, it follows that there was nothing that Brown was in the first place, and thus nothing for either of the new beings to be identical with. There is no such thing as Brown at any point in time: only a continuous perceptual experience of a thing that thinks he is Brown. It is not that “you” would all of the sudden cease to exist after cloning, or “you” would somehow exist in both beings at once: both would simply *think they are a thing that is you*, just in the same way that *you think you are a thing that is you*, and this is all we can say.

We see that this immediately reveals the confusion found in asking Brown, “Should you like to be cloned?” and receiving different answers before the operation and right before his death: you have set him up for contradiction. It is true that a thing that thinks it is Brown survives, just as it does now; but it is also true that another thing that thinks it is Brown dies. Brown, despite what he might think, before he does so, will not live or die. At the moment of his question, all he can say is, “Right now, I exist. I am Brown”.

One still might object that there is a difference between a “real (physically continuous) Brown” and a “clone Brown”, appealing to the fact that the experience that the clone has memories of did not actually happen to it. But to say that “the clone is not the real Brown” or that “the clone's memories are fake memories” is only to say something about the origin of the being. If the memory is true, then there was a time in the past when some perceptual experience that is continuous with the one having the memory experienced that thing, or “made that memory”. We can certainly have false memories, those that are either implanted by some evil genius or simply made up in the way we can ordinarily have false memories; but these are false because they were not actually created from experiences of a perceptual experience continuous with our own. That is to say, the trueness or falsity of a memory is not its origin, but rather its content and relation to our continuous perceptual experience. Further, objecting that the clone “did not actually have the experience contained in the memory, so it is a false perception” quickly creates problems for physicalists. What is meant by such an objection? That the body of the clone did not experience the event in the same way that the body of the actual being did? Every 7-10 years, all of the molecules in our body have been replaced: therefore things that happened to me as a 5 year old, things that I have true memories of, did not happen to the same body as I currently have; yet, no one will deny that they happened to me. Why, then, are we so concerned that the clone has memories of things that “did not happen to their body”; if they are real memories, they at some point were generated from experiences by a thing in a body that thinks it is the same thing as you.

There are, after the cloning, two continuous perceptual experiences that think they are Brown. All the previous discussion of there being “one being” or “two beings”, then, is merely paraphrase, as are references to “a thing that thinks it is Brown”. There actually are never any

instances of a persisting identity, or any “things that think”; there are only a collection of perceptions that think of itself as such⁵. Importantly, this is not to say that what we are, the marker of our personal identity, is psychological continuity; that what we *are* is this continuous perceptual experience. I am denying that there is anything that is you at all. There is just a perceptual experience that persists over time, thinking of itself as a continuing being. I leave open if there are organisms or beings at all, as this is not the point: there is just nothing that is you.

So, the only case in which we would generally like to be cloned is one in which we have pure psychological continuity: cases equivalent to:

Case (7): While you are sleeping, you are vaporized, and then cloned.

Such a case would be identical to nothing happening at all. I believe that it is not “our identity” that is at stake in these instances of cloning: it is the continuity of what we perceive to be ourselves. We should not like to die, to have our experiences ended; and this is all. We are an illusion of succession, reflection, and resemblance.

⁵ I acknowledge that this is perhaps still a little tenuous. Are perceptions things? Do they exist? If not, then I am fine. If they do, this will need to be reformulated.

References

- Lewis, David (1976). Survival and identity. In Amelie Oksenberg Rorty (ed.), *The Identities of Persons*. University of California Press. pp. 17-40.
- Parfit, Derek (2008). Persons, bodies, and human beings. In Theodore Sider, John Hawthorne & Dean W. Zimmerman (eds.), *Contemporary Debates in Metaphysics*. Blackwell.
- Thomson, Judith Jarvis (1997). People and their bodies. In Theodore Sider, John Hawthorne & Dean W. Zimmerman (eds.), *Contemporary Debates in Metaphysics*. Blackwell.