

SUPPOSE TECHNOLOGY ADVANCED TO SUCH AN EXTENT THAT BRAIN TRANSPLANTS BECAME POSSIBLE AND SOMEONE ELSE'S BRAIN WAS IMPLANTED INTO YOUR BODY, WHILE YOURS WAS IMPLANTED INTO THEIRS. WOULD YOU EMERGE FROM THIS PROCESS WITH A NEW BRAIN, OR WITH A NEW BODY? OR WOULD NEITHER OF THESE DESCRIPTIONS OF THE SITUATION BE CORRECT?

In a world within which we would be able to perform successful brain transplants, it is inevitable that this would lead us to raise questions about the nature of our identities. Fundamentally, the issue that concerns us when we consider difficult cases of personal identity is not simply applying a known concept of personhood to an imaginary case, but rather a question of precisely what this concept is, that we claim to be able to apply to it. In reality, we only see continuity of personhood where the body and the brain, or consciousness, coincide: there exist no *real* cases in which we can examine personhood with a separated mind and body. This then, as Johnston (2015) argues, makes our whole approach to such cases flawed: if we rely on our intuitions to determine the true nature of personal identity in such hypotheticals, we assume that our judgement is not impaired by the fact that we never have to make such decisions in reality. Furthermore, as Parfit (1984) points out – albeit in connection to a different matter – our whole intuitive faculty is developed by evolution, and thus best suited to allow us to survive. This implies that our intuition can at most tell us what is best for the survival of the species or of our bodies, which is not necessarily what is objectively true.

These are valid criticisms; we need a different approach to simply relying on reasoned analysis of the case. We must instead take the high road: we need to develop a theory of personal identity independent of the case, and *then* apply it. There is here no fear that our flawed intuition will prove a barrier to understanding, because we only use abstract ideas, independent of hypothetical cases, to develop our concept of identity.

By this approach then, rather than analysing the case directly, we can discover the objective facts of the case. However, there is also a question of whether this objective truth is in fact best suited to our aims with regard to the hypothetical. If personhood, as I explore below, is something that only exists where consciousness – subjective self-awareness – exists, might not a *subjective* answer be more useful, that describes how the case appears to the relevant observers?

By the development of these two methods – the high road to an objective evaluation, and the analysis of subjective viewpoints – it will be found that, if my brain were to be transplanted into another's body, and their brain transplanted into mine, it does not accurately describe the case in an objective sense to claim either that I have a new body, or that I have a new brain. However, in a subjective sense, to the relevant observers, it would most probably appear that I have a new body.

In order to assess the objective truth of continuity of personal identity in any case, it is necessary to consider the “high road” approach, for the reasons noted above. We must therefore develop an overall conception of what it means for a person to be and to continue being, which we can apply to the case in question. The first question to be asked must be whether or not we can be said to have a bodily independent soul – that is, a soul that is independent of the physical body and brain. If we do, then clearly questions of what it means to be a person can be reduced simply to that, without being forced to consider the brain or body at all. It is therefore necessary first to examine the evidence behind the existence of a bodily independent soul. If a soul were to exist independently of the body, we would expect some kind of experience of thought or consciousness in absentia of sense; yet, when we are unconscious, there is no such experience; we require sensation and a conscious brain in order to think. This naturally leads us to the conclusion that, if there be a soul in which our personhood lies, it is not always conscious – as Locke (1700) argues – or, what is more, it is not bodily independent, as it *relies on the senses and the conscious brain* in order to perceive and to think. We therefore reach the conclusion that either a soul does not exist, or, if it does, it is located in any case within the body or the brain. This simplifies the matter, since whether a

soul exists or not, we still conclude that the person consists within the human, and is not bodily independent of it.

If not of a bodily independent soul, then, of what does a person consist? It seems clear that, when dealing with matters of personal identity, we require a sentient mind – one that is self-aware. This would account for why we view humans as people, but not other animals: our own Cartesian sense of self allows us to project this sentience onto others like us, and capable of communicating with us in a way that demonstrates a semblance of rational, self-aware thought. This may not be a necessary or sufficient condition on its own – somebody who has had a heart attack, and is revived by use of a defibrillator clearly lacks a sentient mind in those seconds of a death-like state; yet we would not believe their personhood to be gone, given that their resuscitation allows them to continue living.

However, regardless of whether it is necessary or sufficient, it is a good generic to use, and this may be the best we can achieve by reason alone. Similarly, I can recognise rust without knowing the precise necessary and sufficient conditions for its presence. The same could easily be true of our approach to personal identity; we can know when a person exists or does not exist without knowing at what limit this state changes.

Following this generic, then, what approach can we use to develop criteria for the *survival* of personhood? This is what we are primarily interested with regard to the case of brain transplants: in which body do I *survive*, the original or the one my brain is transplanted into? Given that we require a sentient mind, we must evidently find criteria for this survival that account for the role of the mind. Furthermore, these criteria must account for the fact that different sentient minds exist which appear to constitute different people; we need to know what allows me to continue being the *same* person, not just any person. The Neo-Lockean perspective, promoted by reductionists such as Parfit, seems to provide good objective criteria for this survival: I survive if and only if I display a sufficient degree of psychological *continuity* and *connectedness* over time. This, with reference to our case of brain transplants, implies that the body that continues to have thoughts and opinions following most naturally from mine, and which displays the closest similarity of thoughts and opinions to me before the transplant, is in fact me. This would seem to make the one with my brain my actual self.

However, there is a problem. The implication of the reductionist perspective is that my survival is a matter of degree, not absolute fact. This in turn implies that what allows me to survive is the crossing of some arbitrary point in the relations of psychological continuity and connectedness. What, then, if the person with my brain were in fact to have slightly below this degree – if, suppose, a few crucial particles of brain matter were lost in the transplant, causing me to lose a few memories that would have pushed me just over the edge on psychological connectedness? This is a phenomenon that Johnston (2015, 2016) describes as meaning that we are “ontological trash”: given the criteria that allow us to survive, there are an infinite number of beings that are ontologically on a par with us, which either do or do not survive, solely depending on where we draw the line regarding our relations of continuity and connectedness. Therefore, in any case of ontological survival on the Neo-Lockean view, it seems to be entirely arbitrary whether we survive or not. This makes the question of whether I survive the brain transplant seemingly unanswerable. Perhaps the best way to describe the case in an objective sense is not that I have a new brain or that I have a new body: instead, the only truly accurate answer is that there are here two humans, each of which has had a brain or body transplant. The only way to obtain an objectively accurate answer is to remove the concept of “I” entirely.

Where Johnston, and reductionists such as Parfit, go wrong, however, is in their assumption that objective reductionism means that there can be no further fact than these reductive relations involved in personhood. The idea that we can only determine if I survive on an arbitrary basis ought to point the way out of this conundrum of an “empty” answer: it is to

stop considering personhood as an objective relation and instead to consider the merits of a purely *subjective* answer. It has already above been stated that we believe a sentient mind to inhere in personhood, and that we view others as being people solely on the basis that we can project our own Cartesian sense of self onto them, as they are similar to us in apparent thought. Does this not make the very idea of personhood already a subjective concept, applied only from the position of the subjective observer? The idea, then, that we only survive based on an arbitrary decision of valid degrees of continuity and connectedness relations is not a problem, but rather a necessary feature of an *inherently subjective* relation.

In fact, the reason why the objective question cannot be answered in a meaningful sense is precisely because the relation is not an objective one. It is more useful to view it solely from a subjective perspective.

Our personhood inheres then simply in the views of relevant observers. Can we then find criteria – either necessary and sufficient, or generic – that describe relevant observers? The first and most obviously relevant observer of my survival – or lack thereof – is my own sentient brain. Since my persistence as a person relies on some degree of psychological continuity and connectedness, if my brain believes that in it consists the same person now as before, it is probably true, subjectively, that my personhood *has* persisted with my brain. Other relevant, though less significant, observers would be those who know me: they are those other than myself best placed to assess the actions and opinions of the thing with my brain and determine if, based on this information, it is likely to be me. The relevancy of these observers, then, seems clearly to be a matter of degree based solely on how intimately they knew the self I was before the transplant, in order to compare the psychological continuity and connectedness of this self with the being that has my brain.

The implication of this subjectivity is that the matter of whether I am the same person, considered as a subjective question, is a matter of two factors. The first is in the relation of psychological continuity and connectedness, and the second the relevancy of all observers. We would then find a best subjective answer by taking the aggregate beliefs of the observers on the first factor, with these beliefs weighted by the second. It is therefore obvious that the matter of my personal survival is just as much a matter of the objective value of the first relation as it is of the subjective values provided by the second. These subjective values can be expected to vary, furthermore: in a Western or anglophone society, such as that in which I live, it would be most likely that the body containing my brain is me after the transplant, due to the beliefs of this society at large regarding how consciousness inheres in the brain, and the person in the sentient, conscious mind. It would be our opinion that the relation of psychological continuity and connectedness in the full transplanted brain is of a sufficient degree to claim that I survive the transplant in the body containing my brain. However, it is not impossible, nor even a contradiction, that a radically different society of relevant observers could believe the opposite, if it was their view that personhood inheres in another part, such as the heart, than in the sentient mind, or in our actions: implying that such a degree of psychological continuity and connectedness is not sufficient to claim that my person survives in my brain. This does not weaken the argument in favour of the aggregate subjective answer, as it simply states again the fact that this answer is a subjective one. Indeed, if the beliefs of different societies regarding continuity of personhood do vary so radically, clearly this would only support the need for a subjective answer, since it would provide further evidence for the fact that personhood is itself a subjective relation, so cannot be fully described by objective facts.

There is, in any case, a possibility that the subjective solution can provide a universal answer that is still not objective, and thus avoids the chaos of different beliefs and societies, even though this is itself not a problem with it. This possibility inheres in the concept of an omniscient God. If God exists, and is forever and universally cognizant of all things, clearly the subjective view of our individual personhood in the eyes of God would take supreme

importance among all aggregate beliefs. Indeed, since it is a contradiction for an omniscient God to know or even believe something that is false, this would elevate the subjective view of personhood in the eyes of God to a position near objectivity. This would allow for a subjective answer that nonetheless holds universally, avoiding all concerns of nonuniformity in the subjective answer.

I will not attempt to debate the existence or non-existence of God here, since to do so would be too large a digression, nor shall I pretend to give a reasonable solution to the question of precisely which human God would believe to be me in the case of the brain transplants. It merely suffices to note that this possibility leaves open to consideration a subjective solution – avoiding the problems of arbitrariness – that simultaneously holds universally, avoiding the problems, if indeed they are problems, of chaos.

In which body, then, am I after the transplant? The Neo-Lockean solution demonstrates that the objective answer is that there is no answer. There are two bodies, and two brains. Nothing further can be claimed in the objective sense. However, the existence of the further fact of the subjective beliefs of a system of relevant observers implies that, in the *subjective* sense, there does exist an answer. I am English; my contemporaries are mostly anglophone, predominantly Western. The implication, for my personal case, based on the beliefs of Western civilisation in general, which form the majority part of the aggregate beliefs of my relevant observers, would be that my person will be located in the body that has my brain, that the brain is the seat of consciousness and that this very consciousness, manifesting itself visibly in our actions and opinions, necessarily forms our person. It seems true therefore to say that, whilst in an objective sense there is no real answer to whether I have a new body or a new brain, subjectively it would be taken as true that I have a new body, not a new brain.