

# Personal Identity, Continuity and Brain Transplants

---

**Pećnjak, Davor**

*Source / Izvornik:* **Studia hermeneutica, 2008, 5, 77 - 88**

**Journal article, Published version**

**Rad u časopisu, Objavljena verzija rada (izdavačev PDF)**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:261:507588>

*Rights / Prava:* [In copyright](#) / [Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2025-02-22**



*Repository / Repozitorij:*

[Repository of the Institute of Philosophy](#)

Davor Pećnjak

## Personal Identity, Continuity and Brain Transplants

The problem of personal identity consists in question in what sense a certain person stays the same person if various psychological and/or physical predicates about her change over time. More formally, what is it that makes a certain person A in time  $t_1$  the same as person B at some later time  $t_2$ ?

First, I'd like to say a few words what does it mean to be a person at all. What follows should suffice for our purposes. According to Locke (1690, 1959) who says that person is "a thinking intelligent being, that has reason and reflection, and consider itself as itself, the same thinking thing, in different times and places", Dennett (1976) summarized that some X must fulfill (at least) the following in order to be a person:

1. persons are intentional;
2. persons are rational;
3. persons are objects of attitudes taken towards them;
4. persons can reciprocate attitudes;
5. persons communicate verbally;
6. persons have a special kind of consciousness; to these conditions, Kathleen Wilkes (1988) adds
7. persons use tools.

According to Dennett, these would be, to a first approximation, conditions of personhood.

So, supposition is that human beings are persons because they fulfill the abovementioned conditions. Of course, other kind of beings which would satisfy the mentioned conditions can also be persons.

Various criteria are formulated to answer the question of personal identity. Two main groups of answers exist: those that can be classified as giving *psychological* criteria and those as giving *physical* criteria. Physical criterion or bodily criterion says the following in its basic form: person A in  $t_1$  is identical with person B in  $t_2$  (and  $t_1 \neq t_2$ ) if and only if there is a spatiotemporal continuity of body X from

$t_1$  to  $t_2$  and we identify A in  $t_1$  with body X and B in  $t_2$  also identify with body X.

Shoemaker (1963) invented the thought experiment which showed that the spatiotemporal continuity of the whole of the body is not necessary, but only the spatiotemporal continuity of the brain. Let me briefly present his example.

Organ transplants are nowadays widely applied in medicine. Even heart transplants is so advanced that patients can live for years after the transplantation. New organs adapt to organism in which they are inbuilt. It is at least logically possible that human body function for some time with the help of very sophisticated machines, while the brain is taken out (for medical treatment, for example). We may imagine the following story, according to these initial conditions, says Shoemaker.

Let's imagine that surgery is so advanced that a surgeon can take out the brain out of the head, examine it and operate over it and then take it back into the head. After that has been done, everything functions well and normal and life goes on normally for the patient. Let's say that two persons went on to operate their brains in such a way. These are Mr. Brown and Mr. Johnson. Everything has been done well except that the surgeon has put the brain of Mr. Brown into the head of Mr. Johnson and, *vice versa*, the brain of Mr. Johnson into the head of Mr. Brown. Because of further complications, one of them soon died – one with the body of Brown and the brain of Johnson. The other one – with the brain of Mr. Brown and the body of Mr. Johnson – functions well and is now awakened from narcosis. Let's call him "Brownson". After he regains complete consciousness, Brownson looks in the mirror and is shocked by the looks he saw. The body is Johnson's so Brownson shouts that it not his body. Physicians and surgeons ask him questions – for example what is his name. He automatically answers: "Brown"; furthermore, he knows everything about Brown's wife, dog, family and to all facts and events from Brown's life he refers as to "his". He does not know anything about the life of Johnson. In subsequent days he shows character traits, interests and characteristic behaviour as it was previous Mr. Brown's and nothing like Johnson's. What we would say in these circumstances? It seems, says Shoemaker, that there is not much doubt that we would say that though Brownson has the body of Johnson, the person in question is *not* Johnson, but it is Brown. But

this very answer would not imply the use of spatiotemporal bodily continuity criterion as a criterion of personal identity. Body in our example is certainly Johnson's body (spatiotemporal bodily continuity of Johnson's body is evident). Well, somebody could try to argue that we cannot fully claim that in this case we have Johnson's body, because it is Brown's brain in that body and not Johnson's brain. But the brain, as physical object, is part of the body, and because brains are switched, in Johnson's body one does not find the brain which originally made a physical whole, so this is not a natural whole which was complete original Johnson's body. But this kind of claim does not affect our discussion. For example, if engine is taken out from a car, we still recognise that car, a Mercedes, as a Mercedes, despite its engine being taken out (and possibly replaced). So, we recognise Johnson's body as Johnson's according to all its appearance despite the fact that the brain has been replaced. Nevertheless, in this case we would say that person who survived and continues its existence is, in fact, Mr. Brown. It was Mr. Brown's brain (in Johnson's body) that was preserved and which is responsible for right answers about previous Brown's life, remained character traits, etc. Thus, it seems that spatiotemporal continuity of the whole body is not necessary for personal identity, the spatiotemporal continuity of one part – the brain – could be enough.

Now, I'd like to present the formulation of the psychological criterion of personal identity (Parfit 1971, 1984, Shoemaker 1970). This criterion is based on the concept of memory. Just one remark: those who hold the psychological criterion do not have also to hold that human beings, i.e. persons, are something over and above body and brain.

We rely on the concept of "overlapping memory chains". "Overlapping chains of direct memory" means the following: We can remember some event E and we can have memory about it for some time and after that time we can forget about E. Before we forget event E, we can remember some event F, so for some time we can have memories about both events, E and F. After a while, we forget about E and only the memory about F remains. But, further, we can remember about an event G before we forget about event F etc. For some time, then we have an overlapping of memories between F and G and so on as time goes. This is what we mean by "overlapping chains of memory".

Let's say further that between X today and Y a year ago, there are *direct* memory links if X today can remember some experiences which Y had a year, or ten years, or so, ago. Even if there are no such direct memory links there can be *continuity* of memory between X today and of Y from the past. It can be so if there is an overlapping chain of direct memories, as defined above. Now, Parfit (1984 p.206) defines two general relations: "*Psychological connectedness* is the holding of particular direct psychological connections" and "*Psychological continuity* is the holding of overlapping chains of *strong connectedness*". Concerning connectedness, it can be realised in one or in several thousands direct psychological connections between X now and Y from the past. Parfit also says that between X now and Y yesterday, there has to be enough number of connections and that enough would be at least half the number of direct connections that hold over every day between actual people. We call a situation "strong connectedness" when there are *enough* direct connections.

But this is not a relation which could be a relation of identity. Identity, i.e. relation of equivalence, must be reflexive, symmetrical and transitive. Strong connectedness does not have a property of transitivity. Let's take an example. I am now strongly connected with myself yesterday and yesterday I was strongly connected to myself of a day before and we can go so on into the past day by day. But it does not follow from this that I now am strongly connected with myself ten years ago. Maybe I now only have just a few direct connections between myself now and myself ten years ago or maybe I have none. We have much more memories about the events from a previous day than we have memories about events ten years ago. Transitivity relation would require that if A is in f-relation with B and B is in f-relation with C, then A must be in f-relation with C also. We saw that this does not have to be a case because A does not have to be in f-relation to C if f-relation is defined as relation of strong connectedness. In order to have transitivity and to establish identity relation, we must appeal to *continuity* also. Using abovementioned definitions of *psychological connectedness* and *psychological continuity*, Parfit (1984, p.207) formulates psychological criterion of personal identity:

1. There is *psychological continuity* if and only if there are overlapping chains of strong connectedness. X today is one and the same person as Y at some past time if and only if 2. X is psychologically continuous with Y, 3. this continuity has the right kind of cause, and

4. there does not exist a different person who is also psychologically continuous with Y. 5. Personal identity over time just consists in the holding of facts like 2 to 4.

Imagine now that the following operation is going on (Robinson 1988): we have two persons – me and my twin brother. Operation will consist in replacing neurons, one by one. After one neuron will be pulled out from my brain, one neuron from the brain of my twin brother will be inserted in my brain which is, in turn, pulled out from his brain. Assume that both persons consented to this. Let's imagine that this operation goes well, neurons are being replaced, one by one. At the end, we find all the neurons which assembled the brain of my brother in my head. They are connected in the same way as they previously were in the head of my brother. They make the brain as it has been the brain of my twin brother. So, the question is – who is the person in my head – me or my brother?

If only one neuron has been replaced, we would not say that the person to whom it happened did not thereby survive. It is not at all plausible that person changes or, still less, that person ceases to exist after replacing just a single neuron. If so, then I would say that person neither changes after the second operation in which further neuron is replaced. All the same, I think that person does not change even after the third operation in which a further neuron is replaced. Now, we can iterate this situation until all the neurons are replaced, one by one. If we would like to say the opposite, that the person changes, then we would have to have some non-arbitrary point from which on we could say that person has changed. That non-arbitrary point should consist in a precise number of neurons from which it follows that person is changed or we should say that person changes after a single neuron has been replaced. It is not clear how we should establish this non-arbitrary point and it is implausible to say that a person changes after a single neuron has been replaced. But, Robinson (1988) claims that, at the end of the process, when all the neurons have been changed, we have the same case as in the case Shoemaker describes – it is the case of brain transplantation – the person who survives is my brother but not me; just as in Shoemaker's case, Mr. Brown survived. The only difference is the way in which the brain has been transplanted – in Shoemaker's case the whole brain is transplanted at once, while in the new case the neurons have been transplanted one by one. Because of that, Robinson says that it is

doubtful to conclude that it is me who survived, although the brain is in the body that has been my body. It seems also to Robinson that this kind of operation – replacing neurons one by one – demontsrates that there can be psychological continuity between two different persons. Let's see how he backs up this conclusion. We can imagine that the neurons which has been pulled out from my brain were not thrown to a threadwaste but were instead later reassembled into a structure that is the exact structure of my previous brain. So, it is again a functioning brain – exactly as was my brain. According to Robinson, if I had been informed before everything started that this will be the final outcome of the operation, then I should be concerned (before the operation) about the fate of this very brain that was to be reassembled later and not about the brain which happens to be in my body after the operation. Inside my body, according to Robinson, would be my brother and I would be outside my body – I would be that reassembled brain.

I think Robinson's conclusions are completely wrong. The difference *does* matter and it lies in the way operations are performed. Shoemaker's transplantation of the whole brain and replacing neurons one by one are two very different things. If we remove only one neuron from the brain, it seems that really nothing happens, because brain cells – neurons – die out in a natural way every day and there are naturally fewer and fewer of them alive and functioning in the brain. So, removing a single neuron would not harm the person (nor the brain) in question. Moreover, as our story goes, this removed neuron is immediately replaced by the other neuron, at the very same place where it has been. This inserted neuron adapts and enter into the biological and functional scheme of the brain in which it has been inserted and it becomes accepted from the brain and integrated in it. All is functioning as it should, with the same number of neurons. The functioning of the brain and supporting of everything *psychological* which makes me the person I am is neither affected nor altered. After the insertion of a new neuron everything is as it was immediately before. This holds, then, also for inserting the second, the third etc. etc. neuron. Each and every neuron which is implanted, comes to structure which functions and which integrates it into its functioning and that functioning structure always remains.

Concerning the brain of my brother on the other hand, things

are completely different. When one neuron is pulled out, nothing replaces it. The brain remains with one neuron less. Of course, this, when it is done for the first time, does not affect anything. After the second pulling, the brain remains with two neurons less; after the third pulling, the brain remains with three neurons less; after the three thousandth pulling, the brain remains with three thousand neurons less etc. Nothing replaces them. So, after a while, after certain number of neurons will be pulled out of the brain of my brother, death of my brother will occur because there will be no more enough neurons for supporting the proper functioning of the brain. In that way, my brother will cease to exist. On the other hand, nothing happens to me replacing one neuron by one, so I survive the operation though now I do not have any single neuron which originally constituted my brain – all neurons were constitutive originally of the brain of my brother.

Let's reinforce the plausibility of this conclusion with one detail added but which does not alter any essential feature. Imagine that this operation goes on *in vivo* without anesthesia; imagine that there is a totally painless technique for replacing neurons so patient is fully conscious during the entire operation. All other things stays the same – one neuron is pulled out from my brain and replaced by another one from the brain of my brother while his brain loses one neuron by one without replacing them. I do not see why taking out of one neuron at the time would cause ceasing of consciousness of the patient. And, in my brain, this neuron is immediately replaced by another one. The most I lack at the time is only one neuron. In generating consciousness many neurons cooperate from various parts of the brain (see, for example Greenfield 1995). If only one neuron is taken out from the brain, it does not affect and does not have an effect for generating consciousness. Granted that, it means that the subject in question stays conscious during the entire operation; the patient is conscious, for example, of his surrounding – operation hall, moving of nurses and surgeons and what they talk, his own wonderings about the progress of science, beauty of religion and what is the weather like outside; his consciousness is not interrupted nor lost for a single moment. At the end of operation, all neurons are repalced, and if I was the subject of that repalcing, then at the end, in the body which is mine we can find the brain which was originally composed of neurons which composed my brother's

brain. But, my consciousness was not at any moment interrupted, lost or ceased to exist. There was not any discontinuity of my (stream of) consciousness. We do not have any reason to claim that lacking one neuron at the time can erase consciousness or send a person to coma or to erase a person entirely – it has already been said that in every day neurons in the brain die out in natural way. At the time, there is only one neuron that is missing. It has been replaced by a new one from the brain of my brother and that neuron is then adapted to the work of existing brain and continue to work and function within the existing functioning whole of that brain. Every new neuron is adapted into continuously functioning brain whose characteristics do not change because of replacing just one neuron. The brain would continue to work and to function even if, at some point, operation would stop and no new neuron would be added. My psychological continuity was not in any way broken. According to that, though in my head there is a brain which is composed of neurons which had composed my brother's brain, it is I who survived but not my brother.

Even if neurons, which are implanted into my brain one by one from the brain of my brother, are reassembled in such a way that the original structure of my brain is changed and that the structure of brain get the structure which was originally the structure of the brain of my brother, this also does not mean that he would survive as a person but not me as Robinson thinks and tries to argue. The following will show that, again, it is I who survives. So, let's assume that little by little, replacing neuron one by one, the structure of my brain is reshaped into the structure which was my brother's brain structure. Since, speaking about the memory, the configurations of connections, nets and groups of neurons are carriers of contents of memories, these would change by changing the brain structure. At the end of operation, brain, which is in my body, has completely the same structure as it was the structure of my brother's brain. It does not have its original structure. Moreover, it is composed from all and only the neurons which originally assembled the brain of my brother when it was in his head.

During such a gradual operation, first, only a few memories will be changed; I will retain most of my original memories and acquire, in an artificial way, a few memories about what were not originally my experiences but were experiences of some other person, in this

case, my brother. Many continuing memories and even chains of my original memories will start to overlap with these few implanted memories. And now the chains will start to overlap. So there will be continuity of me in terms of Parfit's definition. After a while, more of this happens, but there is still continuity because previous memories which were not originally mine, already became a part of overlapping chains of connectedness and new memories which have just entered in memory started to become a memory which, then, started to overlap with already existing chains of memory. My original memories are being gradually erased but overlapping chain of strong connectedness is preserved; moreover, we can say that it is also because in normal cases, as we have seen, it could also happen and it happens often. In this operation, this happens only much faster. And it continues to the end of operation. Overlapping chain of strong connectedness is preserved though in the end, there are only contents of memories which originally were memories of my brother. But since the operation undoubtedly began from me, from the person that was I, continuity of *me* is preserved.

If we add and apply what is already said, that operation is made without anesthesia, i.e. that person from the beginning of operation, that is clearly me, does not lose consciousness, then we reinforce the intuition that *I* am the person who survives the operation. It is true that during the stages of the operation, memory would seem strange, if recalled, since it is possible that it will seem that for the same period of time in the past, two different events that took different places to occur are remembered or that some incoherent sequences of events, from the perspective of one person, are remembered and so on. That is true, but continuity would not be broken. And, moreover still, consciousness is not lost and it also continues. It continues from moment to moment of operation as it continues in all other cases in normal life. It continues to the end of operation unbroken and uninterrupted. This is for what we are most interested in in everyday life and if this is preserved, how it could be that at the end of operation it would be, as person, somebody else? Only the contents of memories has been changed to which that consciousness has access. But that contents of memory were changed in an artificial way. It is still *my* memory. Contents of memory, represented as chains, overlap, for at least some periods of time during which operation is going on; and because it clearly started from me, I am the one whose continuity in

terms of overlapping chains of strong connectedness is preserved. Introspective access to contents of memory and consciousness of them, when recalled, are preserved, though the very contents are changed, but it was not destroyed and it did not cease to exist and there is a clear path of overlapping chains of memories (in form of strong connectedness) which started from me.

If there is a clear continuation of consciousness during the operation, as it is as we have clear continuation of consciousness in our everyday dwellings in life, and if only the contents of memory are changed (however radically) in the operative way, there is no reason to think that person is changed or entirely replaced. That, for which we are mostly concerned clearly continues throughout the operation – our ordinary consciousness. That it can recall different contents from memory is nothing unusual. Unusual is only the way how certain contents in memory are implanted. Even this implantation in our example is such that even continuity of memory is preserved.

At the end, when all original memories will be erased and complete set of new memories, like of my brother, will be installed, I will even not notice that contents of my memory were completely changed, but it is still my introspection and my consciousness which have access to them. Even *that memory* has the right kind of continuity in terms of overlapping chains of strong connectedness.

What happens, at the same time, to my brother i.e. to his brain? In the first step, one neuron is pulled out from his brain but it is not replaced by another one. It means that brain now has one neuron less. But, it does not mean anything for the functioning of the brain and for generating consciousness. It stays the same after taking out a certain further number of neurons. Now, these neurons are integrated into *my brain* on the places which were previously occupied by my neurons but they are also adapted to the rest of my brain. In this way, they are accepted by the rest of my brain and they also support all psychological activities and consciousness (along with the other neurons). But, after a while and after a certain greater number of neurons has been taken out from the brain of my brother, he will feel change and difference, and things will definitely change for him. Let's make it clear through an example. Certain posterior parts of the brain are occupied by visual areas of the cortex. To be conscious of some visual sensation or to have a conscious visual

perception, a certain number of neurons (or group of neurons) is needed, which are at the same time activated. If a certain number of neurons is taken out such that it is not possible anymore to activate the needed number of them to produce normal visual perception, then conscious visual perception must be disordered in some way. When all neurons from all visual areas of the brain are once taken out, there is no possibility of having conscious visual perception and sensations. Proceeding further in the same manner, pulling out neurons one by one, eventually will lead to interruption of some vital process which will cause the death of the subject; in this case, of my brother. If he would still be conscious of something until that moment, this will be a moment when that consciousness will fade away and cease to exist and all other physical and psychological processes will cease to exist. At this moment, psychological continuity of my brother will be interrupted and will also cease to exist. The fact that neurons which originally composed his brain are now together again does not mean that they continue the *psychology* of my brother nor they support *his* consciousness. The exact nature of their taking out and implanting into another brain is essential for this case. These neurons support continuity of other person (in this case, me) and they support generating my consciousness.

So, it seems that Robinson is wrong when he says that in this kind of operation, my brother would survive and not I. The case shows that it is *I* as a person who survives.

## Literature

- Dennett, D. (1976): "Conditions of Personhood", in Rorty, A.O. (ed.) (1976)
- Greenfield, S. (1995): *Journey to the Centers of the Mind*, New York
- Locke, J. (1690, 1959): *An Essay Concerning Human Understanding* (ed. A. C. Fraser), London
- Parfit, D. (1971): "Personal Identity", *Philosophical Review*, vol.80. No.1
- \_\_\_ (1984): *Reasons and Persons*, Oxford
- Robinson, J. (1988): "Personal Identity and Survival", *Journal of Philosophy*, vol. 85

- Rorty, A.O. (ed.) (1976): *The Identities of Persons*, Berkeley  
 Shoemaker, S. (1963): *Self Knowledge and Self-Identity*, Ithaca, N.Y.  
 — (1970): "Persons and Their Pasts", *American Philosophical Quarterly*, vol. 7.  
 Wilkes, K. (1988): *Real People*, Oxford

Ivan Kordić

## Mensch und Gott in der technischen Welt Heideggers Ringen um den Menschen nach dem Tod Gottes

### 1. Die Identität aus dem Ab-grund

Die Frage nach dem Menschen ist in der Geschichte abendländischen Denkens mit der Frage nach Gott und mit dem Verstehen seines Wesens bzw. seiner Identität eng verbunden. Es ist aber nicht zu übersehen, daß die Identitäten von Mensch und Gott nie so fraglich und fragwürdig waren, wie es in der heutigen technischen Welt der Fall ist. Dies kommt auf eine rätselhafte und geheimnisvolle Art auch im Denken von Martin Heidegger zum Vorschein, in dem sogar der klassische Satz der Identität in Frage gestellt wird. Dieser kommt bei ihm in der Form eines Grundsatzes vor, der die Identität als einen Zug im Sein als dem Grund des Seienden voraussetzt. Aus dem Satz im Sinne einer Aussage wird dabei ein Satz im Sinne eines Sprunges, der sich vom Sein absetzt und in den Ab-grund springt. Dieser Ab-grund ist das Ereignis, in dem das Wesen dessen schwingt, das als Sprache spricht, die das Haus des Seins ist. Da wird der Satz der Identität als ein Sprung verstanden, den das Wesen der Identität verlangt, damit das Zusammengehören von Mensch und Sein in das Wesenslicht des Ereignisses gelangen kann. Damit hat sich das Denken in seinem Wesen gewandelt. Diese Wandlung zeigt sich besonders darin, daß man jetzt die Konstellation von Sein und Mensch aus dem Ereignis erblickt,<sup>1</sup> das Zeitlichkeit und Endlichkeit, d. h. Relativität des Seins und somit auch des das Sein verstehenden menschlichen Daseins anzeigt, das zum Ort der Offenbarkeit von Gott, Welt und Mensch geworden ist.

Aber obwohl sich in der heutigen Welt aus dem Ereignis als dem

1 Vgl. M. Heidegger, »Der Satz der Identität«, in: *Identität und Differenz*, 4. Aufl., Pfullingen 1957, S. 28.