

Meeting 2: Chapter 1, Part I (Nozick on Personal Identity)

I. Two Principles, Disputed

The topic of this part: what makes an individual at one time *the same person* as an individual at a later time?

One theme of Nozick's discussion: some of the difficulties that arise in this area have nothing special to do with *the identity of persons through time* and instead concern *the identity of any sort of entity through time*.

Another theme of Nozick's discussion: two commonly assumed principles about personal identity are false.

first principle: "If x at time t_1 is the same individual as y at later time t_2 , that can depend only upon facts about x , y , and the relationships between them. No fact about any other existing thing is relevant . . ." (p. 31).

side comment: Presumably 'relationships' here must mean 'intrinsic relationships' or something like that. (Some relationships between two things, such as the relation *being a grandparent of*, also concern the existence of something beyond those two things.)

second principle: "If y at time t_2 is (part of the same continuing individual) as x at t_1 in virtue of standing in some relationship R to x at t_1 , then there *could not* be another additional thing at t_2 also standing (along with y) in R to x at t_1 " (ibid.).

Both principles are very plausible.

And both principles are frequently employed in the personal identity literature to establish various claims.

For instance, Bernard Williams has argued that bodily continuity is necessary for personal identity by appealing to the second principle (machine ten feet to the left vs. another ten feet to the right ex.).

However, Nozick insists, "[b]oth of these principles are false" (p. 32).

(Actually, it's not clear to me that Nozick ends up disputing the second principle; more on this later.)

Nozick's argumentative strategy (in effect): we find analogues of these principles concerning *the identity of any entity (not just persons) through time* just as plausible, for the same reasons; but those analogous principles are false; so our rationale for accepting these principles when it comes to personal identity is also undermined.

Nozick's counterexample to the analogous principles:

"The Vienna Circle was driven from Austria and Germany by the Nazis. . . . Suppose there were twenty members of the Circle, of whom three ended up in Istanbul. These three keep meeting through the war years, discussing philosophy. In 1943, they hear that all of the others are dead. Carrying on its discussions, they proclaim that the Vienna Circle lives on in exile. In 1945, however, they learn that nine members of the Circle had gotten to America, where they continued to meet, discuss philosophy, and so on. That group in the United States is the Vienna Circle in exile; the group in Istanbul turns out not to be the Vienna Circle but its Istanbul offshoot" (ibid.).

Some things to think about when assessing Nozick's use of this case:

- Does the case successfully refute versions of our two principles in which 'individual' is replaced by 'group'?
- Is there any reason to think that the principles that hold for the identity *of persons* through time might be different from the principles that hold for the identity *of groups* through time?

II. The Closest Continuer Theory, Introduced

Nozick then uses the Vienna Circle case to motivate his own proposal about the nature of personal identity.

Let us start by defining a *closeness metric among persons*.

Intuitively, think of it this way: if person y is closer to person x than person y^* is, then y is a better candidate than y^* to be the same person as x .

How close person y is to person x is determined by a complex combination of various factors, including at least the following (p. 69 is helpful here):

- *psychological similarity*: y 's plans, ambitions, desires, preferences, projects, commitments, principles, etc. are qualitatively similar to x 's.
- *psychological dependence*: y 's psychological characteristics "stem from, grow out of, [or] are causally dependent on" x 's psychological characteristics, in an appropriate way (p. 37).
- *bodily similarity*: y 's body parts (including x 's brain) are qualitatively similar to x 's.
- *bodily dependence*: y 's bodily characteristics stem from, grow out of, or are causally dependent on x 's bodily characteristics, in an appropriate way.

Nozick doesn't take a stand on the comparative weights of these factors. (He even allows that some factors might have *lexical priority* over others, so that, for ex., any gain in psychological dependence, no matter how small, might beat out any gain in bodily dependence, no matter how large.)

At one point, he even provocatively suggests that the measure of closeness for a person might be determined in part by that person's "own conception of themselves," and hence varies somewhat from person to person (p. 69). (But which person's self-conception plays this role? x 's? y 's? Both's?)

Let us define *psychological continuity* as *psychological similarity* plus *psychological dependence*, and *bodily continuity* as *bodily similarity* plus *bodily dependence*.

Finally, let us say that y at t_2 is a *continuer* of x at t_1 if and only if y has psychological or bodily continuity with x .

Then Nozick's initial proposal is as follows:

the closest continuer theory (first pass): y at time t_2 is the same person as x at earlier time t_1 if and only if:

- y is a continuer of x ,
- y is close enough to x ,
- y is closer to x than any other continuer (at t_2) of x .

We can get a sense of how Nozick's proposal works by seeing how it handles various cases (many of which have been widely discussed in the personal identity literature):

- *Case 1*: A duplicate of your body, including the brain, is created based on measurements of your own body and brain. Both persons go on living, with qualitatively identical physical and psychological characteristics (at least initially).

verdict: You continue to exist, and the duplicate is not you (because the person with the old body has greater bodily continuity than the person with the new body and scores the same with regard to all other factors that make for closeness).

side comment: This shows that Nozick thinks that the sorts of causal relationships that make for psychological dependence are not the same as the sorts of causal relationships that make for bodily dependence.

- *Case 3:* As you are dying, a duplicate of your body, including the brain, is created based on measurements of your own body and brain, in a way that removes these characteristics from the old body as they are replicated in the new one. The new body goes on living, while the old body dies.

verdict: The duplicate is you, so you continue to exist (because the person with the new body scores well with regard to psychological similarity, psychological continuity, and bodily similarity, though that person lacks bodily continuity with the old you).

- *Case 4:* Half of your brain is transplanted into another body. The old body dies, and the person with the new body is fully psychologically similar to the pre-transplant you.

verdict: The person with half a brain and a new body is you (because with full psychological continuity and some bodily continuity, it is a close enough continuer to be you).

- *Case 5:* After an accident, half of your brain is removed and ceases to function apart from the body. The remaining half continues to function in the body, maintaining full psychological continuity.

verdict: The person with half a brain is still you.

- *Case 6:* Half of your brain is transplanted into another body, but the old body continues to live. The new body plus half-brain is exactly like the continuing person in Case 4, and the old body plus half-brain is exactly like the continuing person in Case 5.

verdict: The person with the old body is you, and the person with the new body is someone else (because although both have full psychological continuity and some bodily continuity, the person with the old body has more bodily continuity).

Nozick is hesitant about this result, though, and contemplates avoiding it by adding a clause to the closest continuer theory according to which y must be closer to x than all other continuers by a certain margin.

III. The Closest Continuer Theory, Further Elaborated

In addition to the notion of a closest continuer, there is the related notion of a closest predecessor.

x is a predecessor of y \equiv_{df} y is a continuer of x .

y is the closest continuer of x \equiv_{df} y is a continuer of x , and y is closer to x than any other continuer of x at that time.

x is the closest predecessor of y \equiv_{df} x is a predecessor of y , and x is closer to y than any other predecessor of y at that time.

What Nozick writes in the first full paragraph on p. 42 suggests that he thinks a more accurate statement of the closest continuer theory would be:

the closest continuer theory (second pass): y at time t_2 is the same person as x at earlier time t_1 if and only if:

- i. y is a continuer of x ,
- ii. y is close enough to x ,
- iii. y is closer to x than any other continuer (at t_2) of x ,
- iv. x is closer to y than any other predecessor (at t_1) of y .

But I'm not sure here, because most of the ensuing discussion focuses on the notion of a closest continuer and neglects the notion of a closest predecessor. (One reason this issue won't matter much: in every case Nozick discusses, a given y is the closest continuer of a given x if and only if x is the closest predecessor of y .)

Nozick goes on to define the notion of being mono-related, which it is tempting to think is defined like so:

x is mono-related to y =_{df} x is the closest predecessor of y , and y is the closest continuer of x .

He then writes, “This mono-relation need not be transitive, since neither closest continuer nor closest predecessor need be transitive” (p. 42).

I found this passage very perplexing. First, it not true that the conjunction of two non-transitive relations must itself be non-transitive. Second, the example Nozick goes on to give in endnote 10 is a case in which x is mono-related to y , and x is mono-related to w , but y is not mono-related to w , which doesn’t show that being mono-related fails to be a transitive relation. (Note: if x is mono-related to y as defined above, it doesn’t follow that y is mono-related to x .)

Here is my best attempt at rescuing Nozick’s reasoning here. Either we should revise our definition of being mono-related as follows:

x and y are mono-related* =_{df} either (x is the closest predecessor of y , and y is the closest continuer of x), or vice versa.

Or we should take the puzzling feature of mono-relatedness to be not its failure to be a transitive relation, but rather its failure to be a Euclidean relation: we can have R_{xy} and R_{xz} without R_{yz} .

Either way, here is a case that makes Nozick think he might need to revise his proposal yet again:

- *Case 8:* Half of your brain is transplanted into another body, and the old body continues to linger on for one hour, or two days, or two weeks. The new body plus half-brain is exactly like the continuing person in Case 4, and the old body plus half-brain is exactly like the continuing person in Case 5.

Let A be the person before the operation, B be the person with the lingering body, C be the person with the new body at the time at which the lingering body is still around, and D be the person with the new body after the lingering body expires.

Nozick consider four options for how to adapt the closest continuer theory to this case (he writes as if option 1 is the one that requires no revision to the theory as developed so far, but so far as I can see it is option 4 that requires no revision):

- *option 1:* A follows the path of closest continuation. So $A = B$, $A \neq C$, and $A \neq D$.
- *option 2:* A follows the path of closest continuation, unless it is a short path. If a time, t , is reached at which there is no continuer of the component of A at the immediately preceding time, then backtracking occurs to the nearest component (if any) of A for which there is a close enough closest continuer at t , and A is identical to that close enough closest continuer. So $A = B$, $A \neq C$, and $A = D$.
- *option 3:* Like the preceding, except we do not allow jumps, and instead take the backtracked path to be the continuation of A . So $A \neq B$, $A = C$, and $A = D$.
- *option 4:* The continuation, if any, of A at a given time is always the close enough closest continuer of A at that time. So $A = B$, $A \neq C$, and $A = D$. (And, bizarrely, $B \neq D$, whereas with option 2, $B = D$.)

Nozick finds none of these options acceptable.

His problem with option 2: if the old body lingers for three years, then he finds it obvious that $A \neq D$.

We could avoid this result by disallowing backtracking that goes back more than a certain amount of time, but Nozick thinks it absurd that this threshold could be sharp (p. 44) and sees no motivation for thinking it might be a vague threshold (p. 660, n. 13). (Why can’t we say the same about the close enough threshold in condition (ii) of his proposal?)

His problem with option 3: how can what happens in the future determine whether current person B is identical to past person A ?

Nozick contemplates moving to a *global mode* on which “Y is (a later stage of the same entity as) X if Y is the closest continuer of X, and if there is no even longer extending thing Z that more closely continues X than any equally large thing of which Y is a part” (p. 47).

So on this proposal, $A \neq B$, $A \neq C$, and $A = D$, so A does not exist while B and C are around.

Nozick’s objection to this proposal: “Surely we may punish D for what C does” (p. 45).

But, on this proposal (if understand it correctly), $C = D$. (Which is itself a weird result.) Maybe we can formulate the global condition in both temporal directions to avoid this verdict?

In the end, Nozick despairs at our prospects for finding a fully plausible way of handing Case 8: “[W]hen I contemplate my entering a situation of temporal overlap, my notion of self begins to dissolve. Is temporal overlap a koan for philosophers?” (p. 47).

IV. Ties and Caring [if there’s time]

Suppose y at t_2 and y^* at t_2 continue x at t_1 equally closely, and both are close enough to x .

Nozick considers four possibilities:

1. $x \neq y$, and $x \neq y^*$. (This is the verdict yielded by the closest continuer theory.)
2. $x = y$, and $x = y^*$. (But given that $y \neq y^*$, how can this be the case?)
3. In fact there are two people, x and x^* , inhabiting the same body at t_1 , and $x = y$, while $x^* = y^*$.
4. Either $x = y$, or $x = y^*$, but not both. (But what explains why one of them is identical to x but the other isn’t?)

Nozick finds the first possibility most plausible.

He tries to explain away our reluctance to accept this verdict as follows:

x cares equally, and to a large degree, about what happens to y and about what happens to y^* .

But if only y or only y^* were around, x would care even more about what happens to this person, because that person would be identical to x .

Thus insofar as we are tempted by possibility 2 above (say), it is because we are tempted to think that if we have a large degree of care for the fate of some future individual that varies with their closeness to us, it must be because we think of that individual as our future self.

Not so, insists Nozick: we do have a certain degree of care for future individuals that varies with the degree to which they closely continue us, but we have a special, even higher degree of care for those future individuals are us (because they are our unique close enough closest continuer at that time).