

Gendler on Why We Can't Trust Thought Experiments on Personal Identity¹

Kathryn Kanuck, Florida

In "Exceptional Persons," Tamar Gendler argues that we cannot make reliable judgments about imaginary cases because the conceptually necessary and sufficient conditions for application of the concept of personal identity (sameness of person) don't guide our application in everyday cases, and while acceptable in actual cases, this leads to error in considering hypothetical cases. In this paper, I explain Gendler's argument against the utility of thought experiments on personal identity and argue that the central case she uses to illustrate the problem is mishandled.

Gendler argues that even if our concept of personal identity is structured around a set of necessary and sufficient conditions, they do not "play a role in how it is that we identify candidates as falling under that concept" (Gendler 1998, 594). Ordinarily judgments about personal identity are based psychological and bodily continuity indifferently. However, Gendler holds, our indifference is not due bodily and psychological continuity each being conceptually sufficient, but rather to their being contingently reliably correlated. In thought experiments, often either psychological or bodily continuity is *absent*, and, relying *uncritically* on ordinary practice, she says, we make judgments based on the remaining cue. Gendler argues that our intuitions about imaginary cases can therefore mislead us by being based on insufficient evidence or on whatever *usual* cues are present.

If Gendler is right about judgments in hypothetical cases, there ought to be cases that trigger different cues and produce inconsistent results. Gendler's argument depends on showing that we make inconsistent identity judgments in two nominally different cases that are essentially the same. I argue that in her central example, the two descriptions are not of the same case and that one of them is of something conceptually impossible. In addition, I argue that even if her example were successful, we could easily avoid making the kind of mistake it would illustrate.

To show that thought experiments are unreliable, Gendler appeals to a thought experiment discussed by Bernard Williams in "The Self and the Future" (Williams 1975). She argues that the two cases involved in Williams' thought experiment (described below), though described slightly differently, are fundamentally the same, though our intuitions about personal identity differ in accordance with the description.

I call the two cases the 'mind-continuous' and 'body-continuous' cases. Williams and Gendler argue that our intuitions yield different criteria for personal identity in the two cases. However, since the cases are supposed to be different descriptions of the same situation, our intuitions, if they are tracking the conceptually necessary

and sufficient conditions for personal identity, should give us the *same* answer.

In both Williams' cases, we are to imagine the following situation: Two persons, person *AA*, the person associated with psychological-states-*A*, body/brain-*A* and person *BB*, the person associated with psychological-states-*B*, body/brain-*B*, are told at t_0 about an operation they will undergo at t_1 the result of which is that at t_2 brain-*B* is associated with the psychological states previously associated with *AA*, and brain-*A* is associated with the psychological states previously associated with *BB*. Two persons emerge—*AB* and *BA*.

In the 'mind-continuous' case we are also asked to imagine that at t_0 , before *AA* undergoes the procedure, she is told that either *AB* or *BA* will be given a large sum of money and the other will be tortured. *AA* is then asked to make a self-interested decision as to whether the reward should be given to *AB* or *BA* at t_2 . Plausibly at t_0 *AA* would choose that at t_2 *AB* receive the reward since *AB* will have the psychological states previously associated with *AA*, the individual making the choice. Here it seems that *AA* makes her choice based on psychological continuity and does not concern herself with body-*A*.

In the 'body-continuous' case, we are asked to imagine that *AA* is told by the surgeon at t_0 that *she* will be tortured after the procedure, at t_2 , but that during the torture *she* will have no memory of being told so. Furthermore, *she* will not have any memories of *her* past; rather, *she* will have a new set of impressions of *her* past. Williams argues that the person associated with psychological-states-*A*, body/brain-*A* at t_0 will be disturbed and concerned that both *her* body and mind will be harmed. That *BA* at t_2 has no memory of having been told the outcome of the procedure beforehand and knows at t_1 that body-*A* will be associated with different brain-states at t_2 seems not to affect the way *she* reacts. Williams argues that it is clear that at t_0 *AA* is, and should be, concerned about what is going to happen to *herself* at t_2 , which shows that *she* identifies *herself* with *her* body.

Gendler argues that what's going on is that the same situation is being described in different ways. One description makes psychological continuity salient. One makes bodily continuity salient. We respond to the salient continuity indifferently, as we do in ordinary life. But both reactions can't be right. So one or the other response must be wrong, she argues, and this shows that in at least one case we rely on a feature contingently correlated with a conceptual sufficient condition, but which is not one. And this shows that thought experiments are generally unreliable, or so the argument goes.

Clearly, Gendler's argument requires the two descriptions to be of the same situation. If the cases

¹Versions of this paper are presented at the UF-FSU Graduate Student Conference (2000), the Alabama Philosophical Society Conference (2000), the Florida Philosophical Association Conference (2001) and the Pacific Division Meetin of the American Philosophical Association (2001). I would like to thank Greg Ray, Kirk Ludwig, Gene Witmer and Tamar Gendler for useful discussion and feedback on this paper.

described are genuinely different, then it does not follow that one of the judgments is wrong. If the situations described are distinct circumstances, we can get different answers to the question of who is who because the application conditions for the concept can be sensitive to the differences between the two cases, thus removing the immediate threat to the utility of thought experiments.

Three conditions must hold for Williams' cases to do the work that Gendler wants them to do:

- (i) our intuitions produce discrepant judgments in response to the different descriptions,
- (ii) our discrepant intuitions are *not* easily dislodged, and
- (iii) the descriptions she offers are of the same situation.

Gendler needs all three conditions since her argument against the utility of thought experiments depends on claiming that we get conflicting intuitions about the same situation described differently. This is evidence for the claim that we do not make judgments about personal identity in hypothetical cases always on the basis of application conditions.

Williams' thought experiment is supposed to reveal that our intuitions can be swayed by various descriptions of the same situation, and so we cannot rely on them to reveal the necessary and sufficient application conditions for personal identity. If the two descriptions are of one and the same situation, yet, contrary to Gendler, our intuitions do not conflict with one another in the first place, then Gendler has not given us any reason to think we are misled by the thought experiment in the way she describes. Alternatively, if the two descriptions are of one and the same situation, then if *at first* we have conflicting intuitions, but upon reflection decide that Gendler may be mistaken, then we have grounds to argue that the danger of being misled is not serious enough to undermine the use of thought experiments in personal identity. Or, if the thought experiment does not contain two descriptions of the same situation, but descriptions of two genuinely different cases, then her claim is groundless. In what follows, I argue that we have reason to question whether any of these conditions hold.

First, assuming that the descriptions are of the same case, it is not clear that our intuitions conflict as Gendler suggests. In the body-continuous case, if AA were told at t_0 that tomorrow she would have none of her memories, she may be disturbed, but perhaps not because BA will not have any memories of AA's past or of being told the situation, but because at t_2 AA will no longer exist! That is, where the surgeon implies AA will undergo the torture that BA will be subjected to, even if AA buys into it at first, it seems that upon reflection, she will realize that *she* will not be subject to torture—*she* will be dead. So, although AA will be disturbed by the situation, she will be disturbed for different reasons than Williams (and Gendler) suggest. The description of the 'body-continuous' case expresses something conceptually impossible. We are gulled by an illegitimate appeal to authority.

Second, suppose, as Gendler argues, that from the two descriptions of the case we make different identity claims that are not immediately reconcilable—in the 'mind-continuous' case we make a judgment based on psychological continuity and in the 'body-continuous' case based on bodily continuity. Once it is made clear that we have made this mistake in judgment the question we should ask is, "can we safeguard ourselves from making this kind of mistake?"

If by being cautious and reflective we *can* safeguard ourselves, then Gendler's observation gives us

no general reason *not* to appeal to thought experiments. So, the fact (if it is a fact) that the conceptually necessary and sufficient application conditions for personal identity do not guide our application of it will not suffice as an argument against thought experiments. Furthermore, we should want a strong argument to support such a claim, since it has profound implications in philosophy independently of the consideration of thought experiments.

Finally, consider whether these two descriptions are of one and the same case. It seems that the first case is described so that it is not obvious who is identical with whom. Presumably when one makes a self-interested decision about who is to undergo torture at t_2 , one implicitly makes an identity judgment—when AA decides that BA should receive the reward and BA should receive the torture, we can infer (that AA thinks) that $AA=BA$ since one would not choose torture for oneself. The second case, however, is described so that it is unnecessary for us to appeal to our intuitions. We cannot *infer* an identity judgment from what the surgeon says about the torture since it is clear from the description who is to be identical with whom. We are told that the surgeon tells AA at t_0 that *she* will undergo torture at t_2 , but that *she* will not have any memories of being told so. The surgeon in the case presupposes that *she* (AA) will be *identical* with the person who will undergo torture at t_2 (BA), thus dictating that personal identity in this case is preserved by bodily continuity. By accepting (on authority) the description of the case as one that is conceptually possible, we have committed ourselves to bodily continuity being sufficient for personal identity, but not on the basis of an intuitive reaction. So, where the 'mind-continuous' description is neutral and invokes our intuitions, the 'body-continuous' description smuggles in the "intuition".

Additionally, if Gendler argues that we get inconsistent intuitions about who is identical with whom in the cases described, it must be that one of the intuitions is right and the other wrong. *Prima facie* Gendler does not have to say which one is correct for her argument to succeed. However, her argument only works if bodily continuity is sufficient for personal identity. In what follows, I show that presupposing *that* criterion for personal identity is disastrous for her argument.

Gendler is committed to bodily continuity being sufficient for personal identity by supposing that the 'mind-continuous' and 'body-continuous' cases are the same case, for the judgment is smuggled into the 'body-continuous' case. Gendler accepts the description is coherent. So, Gendler accepts bodily continuity suffices for personal identity. Since she needs one of our judgments to be wrong for there to be a problem, she cannot hold one of the hybrid views about personal identity, i.e., one where both bodily and psychological continuity are sufficient.

To see Gendler's commitment to bodily continuity, suppose psychological continuity is sufficient for personal identity and bodily continuity is not. Then our intuitions are correct in the 'mind-continuous' case. But the 'body-continuous' case could not then be a true description of any possible circumstance, because a personal identity claim is built into that description. That description could only be correct if bodily continuity were sufficient. Thus, if psychological continuity is sufficient and bodily continuity not, the two descriptions cannot be of the same possible circumstance; the example would not serve Gendler's purposes.

On the other hand, suppose bodily continuity is sufficient for personal identity and psychological continuity is not. Then, our intuitions in the 'mind-continuous' case

are wrong. So, if bodily continuity is sufficient and psychological continuity not, then these two descriptions might conceivably be of one and the same case, because our psychology-favoring intuitions could just be wrong.

This is disastrous for Gendler because her argument against the use of thought experiments relies on a commitment to a sufficient condition for personal identity, which is what is in question. Furthermore, anyone who wants to buy this argument must agree that bodily continuity is sufficient for personal identity because that is the only way the descriptions can be of the same case. However, it is not clear why we should agree or even how Gendler *knows* that bodily continuity is sufficient for personal identity. Endorsing bodily continuity as a sufficient condition is not a position Gendler should want to be in. Not only has Gendler presupposed a criterion that she was supposed to be searching for, she presupposed a controversial, perhaps even counterintuitive, answer to the question about the necessary and sufficient conditions for personal identity! The conclusion to be drawn is not that we should not appeal to thought experiments, but that they should be described in a way that doesn't beg the question if we want useful results. But *of course!*

References

- Gendler, T. 1998 "Exceptional Persons: On the Limits of Imaginary Cases", *Journal of Consciousness Studies*, 5:5-6, 592-610.
- Williams, B. 1975 "The Self and the Future", in J. Perry (ed.), *Personal Identity*, Berkeley: University of California Press, 179-198.