

Against Psychological Sequentialism

Huiyuhl Yi

Received: 18 February 2013 / Accepted: 15 July 2013 / Published online: 4 August 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Psychological Sequentialism holds that no causal constraint is necessary for the preservation of what matters in survival; rather, it is sufficient for preservation if two groups of mental states are similar enough and temporally close enough. Suppose that one's body is instantaneously dematerialized and subsequently, by an amazing coincidence, a collection of molecules is configured to form a qualitatively identical human body. According to Psychological Sequentialism, these events preserve what matters in survival. In this article, I examine some of the main arguments for the view and argue that they fail to establish that no causal constraint is necessary. I also argue that Psychological Sequentialism yields implausible consequences that render it hard to accept the view.

Keywords Causal constraint · Psychological approach to personal identity · Psychological Sequentialism · What matters in survival

The psychological approach to personal identity holds that we persist in virtue of maintaining psychological continuity of some sort. Among those who endorse this approach, it is widely accepted that the psychological connections underlying psychological continuity are essential ingredients of what matters in survival as well.¹ Moreover, defenders of this view typically hold that the psychological

¹ See, e.g., Parfit (1984: 215–217). 'What matters in survival' is a term that denotes one's special prudential concern for some person. Typically, this concern involves the attitude one has toward one's own future experience that one *anticipates*: e.g., we are concerned about our own future surgery in a way that is different from someone else's. Sometimes this concern is about one's own existence: e.g., we

H. Yi (✉)
UNIST (Ulsan National Institute of Science and Technology), 100 Banyeon-ri, Ulju-gun, Ulsan
Metropolitan City, Republic of Korea
e-mail: huiyuhl@unist.ac.kr

connections must be sustained by some appropriate kind of causal relation. The presence of such a causal bond helps to explain why some person existing in the future is the same person as me now, or at least preserves what matters in my survival. And the absence of such a causal bond explains why you and I are not the same person, and why your continued existence in the future does not preserve what matters in my survival. For even if we happen to share a number of exactly similar mental states, these qualitatively identical mental states are not causally related in the right way to one another.

However, a number of psychological theorists argue that no causal relation whatsoever needs to play a role in sustaining psychological connections (Hirsch 1982: 211–226; Kolak and Martin 1987; Siderits 1988; Elliot 1991). On their view, if two groups of mental states are qualitatively similar to each other without any causal basis, and the similarity is strong enough to be comparable to that which obtains between two successive group of mental states in a normal person, that is sufficient to preserve what matters in survival between the bearers of those mental states.² To illustrate, suppose that I die from a car accident. Soon after the accident, lightning strikes a tree in Tasmania, rearranging some of its molecules into a living human being. By an amazing coincidence, the resulting human being turns out to be a perfect duplicate of my body and brain, which exhibits my psychology just before the accident (Siderits 1988: 32). According to the aforementioned view, I ought to consider this series of events to be as good as my surviving the car accident. I shall call this view *Psychological Sequentialism* (or, in short, sequentialism).³ Psychological Sequentialism holds that no causal constraint is necessary for the preservation of what matters in survival; rather, it is sufficient for preservation if two groups of psychological states present in two distinct person-stages⁴ are qualitatively similar enough and temporally close enough—i.e., if they are “psychologically sequential” with each other.⁵

Psychological Sequentialism may be appealing to those philosophers, including Buddhists or philosophers inspired by Buddhism, who deny the existence of an enduring self and want to de-emphasize the boundary between oneself and others. In addition, those who believe that the causal basis for psychological continuity is just

Footnote 1 continued

normally wish that we would still exist after an imminent battle. Hence, it can be said that what matters to one's current self is preserved in one's future selves. It has been argued that what matters to one can also be preserved in a distinct individual, as in a case of fission where a single psychologically continuous stream of consciousness divides into two or more branches.

² It is important to note that their view does not aim to *infer* causality from mere succession of qualitatively similar mental states. The view is that causality is not necessary for maintaining personal identity or the preservation of what matters in survival; mere similarities suffice.

³ I follow Campbell (2005) in using this appellation.

⁴ Here I employ the notion of person-stages merely for convenience' sake. Those who favor a three-dimensionalist ontology should take 'a person-stage' as used in this article to mean a person wholly existing at a time.

⁵ Temporal closeness may not be a part of the definition of 'psychological sequentiality' holding between two person-stages. See, for example, Campbell's definition in his 2005: 381. However, many of the examples discussed by sequentialists suggest that temporal closeness is relevant to two person-stages' being psychologically sequential, although spatial continuity is not.

any cause (e.g., Parfit 1984: 207–209) might find that it is only a small step from their view to sequentialism (Campbell 2005: 382–383). In this paper, I will discuss several arguments in favor of sequentialism that have not been successfully refuted yet, and find problems with each of them. I will also provide an argument against sequentialism that has not been discussed in extant literature.

1 Argument from the Same Psychological Characteristics

In defense of sequentialism, Daniel Kolak and Raymond Martin utilize the idea that a psychologically sequential individual would have exactly the same psychological characteristics as the original person.⁶ For instance, if the original person was vicious and evil, so would be the sequential successor. If the original person was virtuous and honorable, so would be the sequential successor. Being psychologically sequential to each other, the two persons are bound to have exactly the same psychological, and thus moral, traits and inclinations whether or not they are one and the same individual. Furthermore, they will have exactly the same *reactive attitudes* about the actions resulting from those virtuous or vicious traits and inclinations.⁷

Kolak and Martin apply the foregoing consideration to what they call the “Random Match Example” (RME) featuring Robert Nozick as the victim of an incredible series of physical events.⁸ Imagine that the body of Nozick is instantaneously dematerialized as the result of a sudden quantum fluctuation. Immediately afterwards, a collection of molecules is rearranged in the vicinity, which, by an amazing coincidence, results in the formation of a new human body qualitatively identical to Nozick’s body as it was before the incident at the exact spot where it was located. Let us call the person before the incident occupying the old body and the person after the incident occupying the new body Nozick₁ and Nozick₂, respectively. Since Nozick₂ is exactly like Nozick₁, Kolak and Martin argue that *on pragmatic grounds*, it would be pointless to regard them as distinct people. In particular, they write:

Suppose [Nozick₁] had committed a terrible crime just before the breach.
Would that make [Nozick₂], who not only “remembers” performing the crime

⁶ I note that the perfect match required between psychologically sequential individuals is at the psychological level, and not at the neural level. How the psychologies are implemented at the neural level does not seem relevant to maintaining psychological sequentiality, so long as there is exact match between the psychologies thus implemented.

⁷ I find that having exactly the same reactive attitudes is less obviously relevant to the question of blameworthiness, because the default assumption is that the appropriateness of guilt, resentment and indignation is to be explained by the blameworthiness of the agent, and not the other way around. I take it that, on Kolak and Martin’s view, having a certain reactive attitude toward some virtuous or vicious traits is somehow indicative of being responsible for the actions related to them.

⁸ Nozick previously introduced a rudimentary version of an RME-type example in his 1981: 41, which Kolak and Martin utilize and refine. This is probably why they feature Nozick as the main character of their examples.

and suffers from the appropriate guilt, but still has the same psychological inclinations, *not* guilty? Would we, should we, not punish him? (1987: 342)

The claim here is that the moral and legal culpability of Nozick₁ is carried over to Nozick₂ because Nozick₂ has all the psychological features that Nozick₁ has. Hence, Nozick₂'s guilt should follow from Nozick₁'s culpability, which, according to Kolak and Martin, suggests that it is reasonable to regard Nozick₁ and Nozick₂ as the same person so far as the moral/legal viewpoint is concerned⁹; thus, what matters in Nozick₁'s survival is preserved in Nozick₂.¹⁰

My counter argument involves an analogous story that will be presented in two stages. In the first stage, imagine that a group of bio-engineers have invented an operative procedure for creating a human person, in much the same way that Dr. Frankenstein created the Creature. However, their technology is far more advanced, so they can inflict any apparent memories they want on the human person they are creating. Suppose they have just generated a person (call him *Jones*), and, out of evil playfulness, they planted in his mind extremely violent images of murdering someone, when no one in history has ever committed the murder in his pseudo-memories. I take it to be obvious that there is no basis for punishing Jones. He is a poor man born with potentially traumatic apparent memories. No one deserves to be penalized for having pseudo-memories of a violent crime that was never committed.

The second stage introduces a slight change in the preceding story: after the operation, it turns out that by an unpremeditated coincidence, those violent images created by evil bio-engineers exactly match what has happened in an actual murder case; in the actual course of human history, there happened to have been someone who has murdered a person in exactly the same way as it appears in Jones's apparent memories. Should this change alter the verdict regarding Jones's culpability? That is, could he now be charged with murder after we learn about the exact match between his apparent memories and an actual homicide? Noting that the engineers did not intend to match the pseudo-memories of their creature with any actual murder case, it is difficult to understand how this change in the second stage could make a *moral/legal* difference. For it runs counter to our moral/legal practices to base our judgment of someone's culpability on chancy ground. Whether or not we should convict an individual of a crime requires careful moral deliberation. The

⁹ Here Kolak and Martin seem to presuppose the dubious principle that if one's culpability is carried over to a later person, then they are numerically identical. (Notice, however, that the converse is plausible; cf. footnote 16.) Though this supposition may be cast in doubt, I will not press this point further in this article. Instead, I will try to show that in the case under discussion, the culpability of Nozick₁ does not survive the breach.

¹⁰ One might point out that this argument conflates personal identity with what matters in survival, and is therefore fallacious. However, even among those psychological theorists who believe that personal identity is not what ultimately matters in survival, it is widely accepted that personal identity happens to *preserve* what matters in survival. This is because personal identity is constituted by a non-branching form of psychological continuity, and psychological continuity underlain by the right kind of cause is responsible for preserving what matters in survival. According to this view, in general, if a later person is identical to an earlier person, then the later person preserves what matters in the earlier person's survival, but the converse does not hold. Hence, I do not think that Kolak and Martin conflate personal identity with what matters in survival; instead, their argument can be understood as inferring from the obtaining of personal identity to the preservation of what matters in survival.

decision should not hinge upon arbitrary factors such as the length of a straw or the number on a die. For similar reasons, whether or not we should hold Jones accountable for the killing cannot depend upon the sheer coincidence by which his pseudo-memories match the actual murder. If this coincidental match makes a difference in regard to the culpability of Jones, that seems to undermine our robust intuition regarding moral blameworthiness and responsibility.

The preceding discussion indicates that if Jones is taken to be innocent in the first stage, then he should be taken to be innocent in the second stage as well. If this logic is correct, then we should conclude likewise that Nozick₂ is not culpable for any crime committed by Nozick₁, because the original case of Nozick₁ and Nozick₂ is sufficiently analogous to the second stage of the aforementioned story. This consideration suggests that Kolak and Martin's argument is mistaken. In sum, my argument above can be represented as follows:

PR1. Jones in the first stage is not culpable of any crime.

PR2. If Jones in the first stage is not culpable of any crime, then Jones in the second stage is not culpable of any crime.

PR3. If Jones in the second stage is not culpable of any crime, then he is not culpable of the actual crime committed by the murderer in the past (without presupposing that Jones is/isn't identical to the murderer).

C1. Jones in the second stage is not culpable of the actual crime committed by the murderer in the past (without presupposing that Jones is/isn't identical to the murderer).

PR4. Jones in the second stage is analogous to Nozick₂ of Kolak and Martin's story.

C2. Nozick₂ is not culpable of the crime committed by Nozick₁ (without presupposing that Nozick₂ is/isn't identical to Nozick₁).

Defenders of sequentialism may dispute the relevance of my argument by analogy. They may claim that, in order for the analogy to be relevant, my two-stage story ought to mention that Jones has memories that match not just *some* but *all* (or nearly all) of the memories or experiences of an actual murderer, along with the duplication of all *other kinds* of psychological states between the two. This is because, given the larger argumentative context, my story ought to be an RME-type case involving persons with exactly similar psychologies like the original RME case, but about whom we have intuitions that (*pace* Kolak and Martin) favor their non-identity, or undermine the judgment that they stand in any relation that preserves what matters in survival. However, the objection goes, once I introduce exact and complete similarity between the psychological states of Jones and the actual murderer, our intuitions will favor the verdict that Jones is guilty of murder because he then has the actual murderer's psychological traits such as viciousness, hatred of the murdered victim, intention to murder the victim, and so on. In short, the sequentialist may object that my analogical argument is vitiated by a disanalogy—namely, it fails to provide a case of two individuals with exactly similar psychologies as in the original RME case.

In response to this objection, I would argue that merely having vicious moral traits or inclinations, or having appropriate reactive attitudes thereof, is insufficient

grounds for being subject to a charge of wrongdoing. Suppose both A and B wanted an innocent person C dead, but only B actually performed the act of murdering C. It would be absurd to claim that A is as guilty of C's death as B is on the grounds that A also bore ill will against C. It may be that A is blameworthy for *wanting* C dead. However, A is not in any way blameworthy for the particular act of *killing* C. Similarly, unless we presuppose that Nozick₂ is responsible for Nozick₁'s crime, the fact that Nozick₂ would be equally as vicious as Nozick₁, as well as the fact that he would have the same reactive attitudes toward Nozick₁'s crime as Nozick₁ himself would, does not warrant the assignment of blame to Nozick₂ for any of Nozick₁'s wrongdoings.

My argument here is based on the observation that no one should be held accountable for a crime merely because of his vicious traits or inclinations. In response, the sequentialist might point out that a remarkable exception could be found in Philip K. Dick's *The Minority Report*. This novel depicts a future society in which people are arrested because they would have committed a crime unless prevented by the Precrime System that detects their would-be crimes before they perform them. This system is justified on the grounds that in the "post-crime punitive system, (...) punishment was never much of a deterrent, and could scarcely have afforded comfort to a victim already dead" (Dick 2002: 72). The crux of the argument as it relates to Kolak and Martin is that the punitive system depicted in the novel seems to be conceivable, even though under such a legal institution, would-be criminals can be proclaimed guilty without actually committing a crime. This consideration, the argument goes, opens the possibility that one may be held accountable for a wrongdoing merely by having vicious traits or inclinations. Under such a punitive system, actually committing a crime is not necessary for deciding one's culpability.

However, this response does not strike me as tenable, because even the Precrime System would not condemn someone to imprisonment *merely* because of his vicious traits or inclinations. Punishment would follow only if these traits or inclinations *would* lead him to commit a particular crime in the future. For instance, I would *not* be caught by the Precrime System merely because I was inclined to murder my nagging neighbor. I would be caught only if I was actually going to act on my evil inclination. In the RME case involving Nozick, there is no reason to think that Nozick₂ is bound to act from his immoral traits. For we have no more reason to believe that Nozick₂ will in fact act on his immoral traits than to think that just anybody who is vicious and cruel will actually act on those traits, and we have insufficient grounds to think that such a person will actually act on those immoral traits. Hence, the Precrime System does not afford a counterexample to my claim regarding the insufficiency of vicious traits or inclinations for demonstrating culpability. In sum, then, Kolak and Martin's RME case does not provide a good reason to think that moral/legal culpability is transmitted across psychological similarities (including similar moral traits and inclinations) in the absence of any suitable causal basis for sustaining psychological continuity. Since Kolak and Martin adduced such a transmission of culpability as a supporting consideration for sequentialism, this undermines the sequentialist view that psychological sequentiality without any underlying causal relation preserves what matters in survival.

2 Argument from Equanimity

In another attempt to defend sequentialism, Kolak and Martin argue that our intuitions about how we would typically respond to the discovery that we had undergone an RME-type incident lend support to their view that one can survive the causal discontinuity figuring in such incidents. In furtherance of this argument, they ask us to imagine that 10 years ago, one underwent a sudden dematerialization, and, by an amazing coincidence, someone exactly like him immediately came into being at the exact spot where he was. Suppose that you have just now discovered that you are the person occupying the new body. Upon learning about your unusual past, would you be disturbed? That is, would you now believe that you have lived only for 10 years, and regard the life led by the person before the incident as if it were someone else's? Kolak and Martin answer these questions in the negative, and find that they would react in the same way about any such causal discontinuity located in the future. They claim that a person faced with such a causal "breach" either in prospect or retrospect would remain equanimous. Furthermore, this equanimity, according to Kolak and Martin, is evidence that the pre-breach person and the post-breach person are identical, which entails that what matters in survival to the pre-breach person is preserved in the post-breach person. This is because if they were not identical, the awareness of their non-identity would inflict severe psychological distress on the person who has just found about the incident. They write:

It would seem that not caring, *in this case*, is an indication of one's belief that identity would be preserved. The knowledge that we are not who we thought we were, or that we will soon be annihilated, would, for most of us, induce severe psychological trauma. The knowledge that breaches of the sort under discussion had occurred, or will soon occur, to "us" is not the least unsettling. (1987: 342)

Here Kolak and Martin draw a conclusion about one's identity from one's emotional stability. The core of the reasoning in their argument includes the following premise: If one believes that he is not identical to the pre- or post-breach person, then he would not remain equanimous. Employing this premise, they deduce that, since one does remain equanimous, he believes that he is identical to the pre- or post-breach person. This of course does not give us the desired conclusion that they *are* in fact identical. However, it is common to reason from our beliefs about personal identity to the nature of personal identity itself. For instance, Parfit himself famously argued that personal identity is not what matters in survival via investigating what we believe in various hypothetical or actual cases (Parfit 1984: 253–261). We can take it that Kolak and Martin are simply following this general method, and suggesting that we should revise our beliefs about the necessity of the causal constraint.

One problem with Kolak and Martin's line of argument, though, is that it fails to provide a case that is strong enough to warrant the transition from our beliefs about what is the case to what is actually the case. In general, one's feelings (including equanimity) about a possible state of affairs are feeble grounds on which to base any reliable conclusion regarding the truth about that state of affairs. For instance, we

may grant that one's awareness that he might have been a changeling may give him a shock. However, from the observation that one remains calm upon learning of his possible misplacement at his birth hospital, it would be utterly hasty to conclude that he was not misplaced. His equanimity may reflect that he does not believe that he was misplaced. But his disbelief may be based on a poor justification—e.g., he may refuse to believe that he is a changeling out of the wish that the couple who raised him are indeed his biological parents. None of this tells us anything about whether he was in fact misplaced as a baby. Similarly, the fact that one is equanimous with the causal breach does not tell us whether what he believes is *in fact* the case.

Moreover, it is not clear to me that most of us share Kolak and Martin's intuition that we would remain calm in the face of a future causal breach, though we may agree with their intuition in regard to a past causal breach. If a causal breach lies in our future, and we are informed of this future event, I think that we would most likely be perturbed and afraid. This asymmetry in attitude toward past and future causal breaches may be explained at least in part by our asymmetrical attitudes toward prenatal and postmortem nonexistence, often discussed in the philosophical literature on death. It has been well pointed out that we maintain complete indifference when considering the nonexistence prior to our origin, though we are typically wrapped up in utmost fear and dread when contemplating the nonexistence after our death.¹¹ This asymmetry in attitude toward our past and future nonexistence is natural and deep-seated.¹² Here it is important to note that our nonchalance toward our past nonexistence is entirely compatible with our own nonexistence prior to the causal breach. One can be calm in finding out about a past causal breach, not because he knew that he had survived the breach, but because past nonexistence is not much of a horror to him anyway. Hence, one's equanimity towards the past causal breach does not necessarily suggest that he is identical to, and thus preserves what matters in the survival of, the pre-breach person. Once again, one's equanimity is not a sure-fire guide for inferring that one would survive a causal breach.

3 The “Indiscernible Swap” Argument

More recently, Campbell discusses what he calls the “indiscernible swap” argument for sequentialism (Campbell 2005: 386). Suppose that an evil genius will

¹¹ This asymmetry in our attitudes has most famously been challenged by Lucretius. He writes, “Look back again—how the endless ages of time come to pass before our birth are nothing to us. This is a looking glass Nature holds up for us in which we see the time to come after we finally die. What is it there that looks so fearsome? What's so tragic? Isn't it more peaceful than any sleep” (2007: 101)? I will not address this issue here because doing so would be beyond the scope of this article. For some influential contemporary treatment on this issue, see Nagel (1970), Brueckner and Fischer (1986), Feldman (1991), Kaufman (1996), McMahan (2006).

¹² This does not mean that our asymmetrical attitudes toward past and future nonexistence cannot be justified. One might, for example, try to justify the asymmetry as follows: we would be much less distressed in retrospect about a past surgery than we would be in prospect about a future surgery; this is because the past surgery is already over and done with, but the future surgery is yet to come. Similarly, regarding the harms associated with our past and future nonexistence, one is already over but the other is yet to come—hence the asymmetry in our attitudes toward them.

reconfigure your brain based on the blueprint of the brain of a randomly chosen person (who is not you) from a randomly chosen place in the universe, which he scanned at a certain time t_1 . This procedure will take place almost instantaneously, and it will end exactly at t_2 , which is a fraction of a second later than t_1 . If you are a psychological theorist, you might take this news badly, as an announcement of your impending death. But suppose further that due to an amazing coincidence, the brain state (and thus the psychological state) of the blueprint donor at t_1 will happen to be exactly similar to how your brain state would have been at t_2 without the reconfiguration procedure. Campbell argues that, in this case, you ought to regard this incident as if it were an ordinary case of survival: what matters to you at t_1 in terms of survival is preserved in the resultant person existing at t_2 and thereafter (call that person *your psychological successor*, who may or may not be the same person as you), because the psychological state of your psychological successor at t_2 is exactly like how your psychological state would have been at t_2 if the “overwriting” had never occurred. Nevertheless, Campbell maintains, the psychological state of your psychological successor at t_2 is not causally related to your psychological state at t_1 ; rather, it is causally related to the psychological state of the blueprint donor at t_1 .¹³ Campbell concludes that this case reveals that no causal relation is necessary for the preservation of what matters in survival.¹⁴

I do not find Campbell’s argument convincing. The upshot of his argument is that (i) the psychological state of your psychological successor at t_2 is not caused by your psychological state at t_1 , although (ii) what matters to you as you are at t_1 in terms of survival is preserved in your psychological successor as he is at t_2 . In what follows, I will argue that on close examination it is difficult to elicit both results simultaneously from the story of the “indiscernible swap.” This consideration puts the sequentialist (including Campbell) in a dilemma—either he accepts (i) on pain

¹³ It should be noted that here Campbell relies on the counterfactual account of causation, which analyzes causation in terms of *counterfactual dependence*, where counterfactual dependence between events is roughly understood as follows: an event e counterfactually depends on an event c iff had c not occurred, e would not have occurred either. *Pace* Campbell, I will later claim that on a different version of the “indiscernible swap” argument, it is not clear that the occurrence of the psychological state of your psychological successor at t_2 is not caused by the occurrence of your psychological state at t_1 , on any plausible version of the counterfactual account of causation.

¹⁴ One might argue that Campbell’s argument is based on the functionalist theory of mind since the argument involves the matching of two brains existing in different substrates, and point out that functionalism is subject to some serious objections involving inverted spectra and multiple realizability, for instance. The possibility of an inverted spectrum may threaten the functionalist theory of mind because it suggests that two functionally indistinguishable individuals may have different internal phenomenology. I think sequentialists can stipulate that one’s psychological successor in RME-type cases ought to have experiences with exactly the same representational contents. As to the qualitative contents of their experiences, it may be necessary to stipulate that one’s psychological successor be able to enjoy nearly as much variety of qualitative experiences as one does (so, not too many missing qualia, though it might be okay to miss what it’s like to experience some nameless shade of grey and the like). In regard to the multiple realizability objection, I think sequentialists can point out that the default position in the philosophy of mind is that our mental properties *supervene* on properties of microphysical entities and arrangements thereof, so that if two individuals A and B are indiscernible at some appropriate microphysical level, then A and B are indiscernible at the mental level. So, sequentialists can stipulate that the exact match between two brains in Campbell’s argument occurs at the appropriate microphysical level.

of giving up (ii) or he accepts (ii) on pain of giving up (i). Neither case shows that causation is not necessary for what matters in survival.

Let us begin by clarifying what is involved in the aforementioned “overwriting” process, which admits of at least two possible interpretations. The most plausible understanding of the intended process is that the evil genius disassembles your brain shortly before t_2 and reassembles it at t_2 .¹⁵ On this interpretation, your psychological state at t_1 and that of your psychological successor at t_2 are indeed not causally related. However, thus understood, the “indiscernible swap” case is not significantly different from the RME-type cases that have been previously discussed. Hence, if I am right in claiming that those RME-type cases do not afford adequate reasons for thinking that the moral/legal culpability of the pre-breach person is transmitted to the post-breach person, then the present interpretation of the “indiscernible swap” case likewise offers no good reason for thinking that what matters to you at t_1 in terms of survival is preserved in your psychological successor at t_2 . For if what matters in one’s survival is preserved in a later person, then it is reasonable to suppose that one’s culpability is carried over to the later person.¹⁶ Since it is not reasonable to suppose this in RME-type cases, on the current interpretation of the “indiscernible swap” case what matters in your survival is not preserved in your psychological successor.¹⁷

The sequentialist might reply that Campbell’s case is significantly different from the previous RME-type stories because in this case there exists physical continuity of the rest of the organism (and arguably even of the brain), whereas in earlier cases there was no such continuity. Before t_2 , the argument goes, you might not be so disturbed to learn that the subsequent set of mental states, which perfectly match your old ones, will be instantiated in the same body that you currently have. I think this sort of response would be of no use to sequentialists because sequentialism is a version of the *psychological* approach, which holds that only psychological factors are relevant to personal identity, or to what matters in survival.

So far I have shown that, on the first horn of the dilemma where there is indeed a causal breach, Campbell fails to establish that you survive the breach. In the remainder of this section, I will address the second horn of the dilemma, introducing

¹⁵ Additionally, this could mean either that your psychological successor retains your old brain matter at t_2 (if the scientist is simply rearranging the parts of your brain just as one might shuffle a deck of cards), or that your old brain matter is replaced with new matter at t_2 (which seems likely if the scientist is using a quantum fluctuation generator to make your brain match the blueprint donor’s). The second alternative is closer to the RME-type cases discussed in the preceding sections. However, I do not think that the distinction between these two alternatives makes much difference since mere change of matter does not seem relevant with respect to what matters in survival.

¹⁶ Note that this is different from the dubious principle presupposed by Kolak and Martin in their argument involving the same psychological characteristics; cf. footnote 9.

¹⁷ There is another line of reasoning for rejecting the claim that what matters to you at t_1 in terms of survival is preserved in your psychological successor at t_2 . Campbell thinks that the relation that matters in survival holds here because the psychological state of your psychological successor would have been exactly similar at t_2 even if the swap had never occurred. In this sense, the swap is “trivial”—the world would not have been any different had it not occurred. However, it was recently pointed out that this sense of triviality (namely, making no “significant difference to the qualitative nature of the world”) does not entail that the trivial event in question does not affect what matters in survival. See Brueckner and Buford (2013: 99–101).

a possible move to ensure that you survive the breach, and show that it fails to establish that there is a causal breach in the first place.

To ensure the result that what matters to you at t_1 in terms of survival is preserved in your psychological successor at t_2 , the sequentialist may attempt to describe the story of the “indiscernible swap” in a slightly different way. On this version of the story, what happens is that the evil genius does not do anything because he anticipates that the brain state of your psychological successor at t_2 will turn out to be exactly similar to that of the blueprint donor at t_1 . However, if he were to notice in advance that the brain state of your psychological successor at t_2 is going to differ from that of the blueprint donor at t_1 , then the evil genius would reconfigure your brain so that the resulting brain state would match the donor’s. Unlike the previous interpretation, the current description of the story ensures that what matters to you at t_1 in terms of survival *is* preserved in your psychological successor at t_2 . To demonstrate that what matters in survival is preserved here, note that on this description the “indiscernible swap” case is analogous to a famous Frankfurtian counterexample against the Principle of Alternate Possibilities, which states that moral responsibility requires the freedom to do otherwise (Frankfurt 1969). In the counterexample, Black has implanted a super microchip in Jones’s brain unbeknownst to Jones whereby he can completely monitor Jones’s brain states, and thereby control Jones’s intentions and the resulting behaviors.¹⁸ Jones kills Smith by his own decision, but had he decided to do otherwise, then Black would have manipulated Jones into killing Smith anyway via the microchip he implanted in Jones’s brain (van Inwagen 1983: 162–164). It is important to note that in this case Black is considered to be a “counterfactual intervener”: though he did not actually do anything, he would have intervened had things been different from the actual course of events (i.e., had Jones decided not to kill Smith). Similarly, in the redescribed “indiscernible swap” story, the evil genius could be understood as playing the role of a counterfactual intervener: though he did not in fact do anything, he would have intervened had things been different (i.e., had the brain state of your psychological successor been different from how it actually was at t_2).

In the Frankfurtian case, it is generally conceded that Jones is responsible for what he has done. After all, Black did not actually intervene, and Jones had decided to kill Smith all on his own. Thus the mere existence of a counterfactual intervener does not impede the assignment of moral responsibility. If this is correct, then a similar line of reasoning ought to apply to the question of what matters in survival: the mere existence of a counterfactual intervener does not impede the preservation

¹⁸ I have used the term ‘brain state’ to refer to the total global state of a brain in all its parts, including all the distinct individual states that the brain is in at some given moment. I do not mean by it, for instance, an individual state of the brain in some localized part of it, say the state of some individual c-fiber in the brain. When interpreted in the second way, my description of the Frankfurtian counterexample may seem to make the dubious assumption that there is a one-to-one mapping between an action and a brain state. For example, my raising an arm probably does not bear a one-to-one causal relation to any single state or event in some localized part of the brain, such as the firing of an individual neuron—so that if that neuron fires, then my arm is raised. My understanding of the Frankfurtian story does not make this assumption. I do think that the story presupposes *some* mapping of Jones’s brain states and his actions, but that mapping is not one-to-one. (Without the mapping, Black would not be able to predict Jones’s actions from a reading of his brain states.)

of what matters in survival. Then, in the redescribed “indiscernible swap” case, it seems that what matters to you at t_1 in terms of survival is indeed preserved in your psychological successor at t_2 . After all, no one really tampered with your brain, and it is hard to believe that a merely modal fact—that the evil genius *could* have intervened and reconfigured your brain—terminates your existence in reality. (Suppose that you could have been killed by a truck, though in fact you were not, had you crossed the road a second ago. No one would think that you have just ceased to exist due to the mere possibility of your being run over.)

I have considered how the sequentialist could redescribe the “indiscernible swap” case in order to plausibly claim that what matters to you at t_1 in terms of survival is preserved in your psychological successor at t_2 , thus reviving Campbell’s argument. The problem with the new description, however, is that on this version of the story, your brain state at t_1 does seem causally related to the brain state of your psychological successor at t_2 . Let us take occurrences of brain states to be events. We shall refer to the event that your brain state occurs at t_1 as it actually does at t_1 ‘ e_1 ’, and the event that the brain state of your psychological successor occurs at t_2 as it actually does at t_2 ‘ e_2 ’. Campbell’s claim then is that e_2 is not causally related to e_1 , because e_2 would come about even if e_1 had not occurred (i.e., even if your brain state were different at t_1), insofar as the brain state of the blueprint donor would not differ from how it actually is at t_1 . This claim, however, is at odds with any tenable version of the counterfactual account of causation.

To see this point, note that the redescribed story of the “indiscernible swap” contains an instance of what Lewis calls ‘preemption’. In general, preemption can be regarded as a type of redundant causation that involves at least two causal chains—one merely potential, and one actual—running from potential causes to an effect. This type of causation is asymmetrical in the sense that the two potential causes do not have an equal claim to being the actual cause of the effect—one of them is the preempting cause that actually brings about the effect, while the other is the preempted alternative that does not fulfill its potential. The (merely potential) causal chain running from the preempted alternative is cut short in its path by the preempting causal chain that goes on to complete the causing of the effect.¹⁹ For example, when both Suzi and Billy throw rocks at the same bottle, and Suzi’s rock strikes the bottle sooner than Billy’s, we can say that Suzi’s throwing of her rock is the preempting cause of the bottle’s breaking while Billy’s throw is the preempted alternative. The causal chain running from Billy’s throw to the shattering of the bottle is cut off when Suzi’s rock completes its job. Similarly, in the redescription of the “indiscernible swap” case, we can plausibly take e_1 as the preempting cause of e_2 while taking the event that the blueprint donor’s brain state occurs at t_1 as it actually does at t_1 (call this event ‘ e_b ’) as the preempted alternative. There exists an unactualized causal process running from e_b to e_2 . And this process is cut off when e_2 , the effect, occurs in the actual scenario as a result of the causal process running from e_1 to e_2 .

¹⁹ In fact, Lewis (2000) divides preemption cases broadly into two kinds, on the one hand “cutting” cases as discussed above, and on the other hand “trumping” cases that do not involve cutting of the preempted causal chain midway in its path. In this article, it is only the former kind that I refer to by ‘preemption’.

Preemption cases have posed a serious challenge to the counterfactual theory of causation because in such a case, the effect does not counterfactually depend on the preempting cause due to the existence of the preempted alternative,²⁰ and thus it seems that the theory ought not to count the preempting cause as a cause of the effect. However, as Lewis (1986: 200) observes, “[i]t is clear what answer we want—the preempting cause is a cause, the preempted alternative is not—and any analysis that does not yield that answer is in bad trouble.” Many defenders of the counterfactual account have attempted to provide a viable solution to this problem, all aiming to derive the correct result that the preempting cause, and not the preempted alternative, causes the effect.²¹

In his discussion of the “indiscernible swap” argument, Campbell clearly appeals to the counterfactual account of causation. And, as noted in the above discussion of preemption, any plausible version of the counterfactual analysis should yield the result that it is e_1 (the preempting cause), and not e_b (the preempted alternative), that causes e_2 (the effect). Since the preempted alternative is not a cause, it is not the psychological state of the blueprint donor at t_1 that causes the psychological state of your psychological successor at t_2 . On the contrary, a plausible counterfactual account should yield the result that it is your psychological state at t_1 that does the causal work. So, although we may grant that, in the redescribed version of the “indiscernible swap” story, what matters to you at t_1 in terms of survival is preserved in your psychological successor at t_2 , we ought then to conclude, *pace* Campbell, that a causal relation does obtain between you at t_1 and your psychological successor at t_2 . Therefore, the redescribed story fails to show that no causal constraint is needed for the preservation of what matters in survival.²²

Consequently, the sequentialist fails in either one of the two possible interpretations on what happens to your brain in the story of the “indiscernible swap” case. If the interpretation favors disassembly followed by reassembly of your brain, then the sequentialist fails to show that what matters in survival holds between the pre-breach and the post-breach persons, though no causal relation holds between their psychological states. Alternatively, if the interpretation favors no actual intervention by the evil genius, then the sequentialist fails to show the

²⁰ For instance, in the preceding examples, if Suzi’s rock had been blocked at any point in its itinerary, Billy’s rock would have broken the bottle instead. Likewise, even if the causal link between e_1 and e_2 were somehow severed, e_2 would still have occurred due to the causal process running from e_b to e_2 . So the effect in each of these examples does not depend counterfactually on the preempting cause—i.e., it is not the case that if the preempting cause had not occurred, then the effect would not have occurred either.

²¹ For instance, in dealing with preemption, Lewis employs the notions of *stepwise dependence* (to handle “early” cutting) and *quasi-dependence* (to handle “late” cutting) in his 1986: 200–206. He later rejects the notion of quasi-dependence in his 2000, where he develops the concept of *influence* to handle the problems raised by the newly introduced trumping cases, as well as the “late” cutting cases discussed in his earlier article.

²² So far I have discussed the “indiscernible swap” argument only in light of the counterfactual account of causation because Campbell’s discussion was couched in terms of this account. However, other theories of causation may account for our intuitive judgment that there is a causal relation between e_1 and e_2 with greater ease. For instance, according to a simplistic regularity theory of causation, e_2 is caused by e_1 , because e_1 instantiates one event-type and e_2 instantiates another, these types being such that instances of the former are invariably followed by instances of the latter.

absence of causal relation between the psychological states of the pre-breach and the post-breach persons, though what matters in survival holds between them. In each case, the “indiscernible swap” argument fails to make the case for sequentialism.

4 Implausible Consequences of Psychological Sequentialism

Psychological Sequentialism is committed to two claims: (i) no causal relation is necessary for the preservation of what matters in survival, and (ii) it is sufficient for the preservation of what matters if a person existing at one time is psychologically sequential to a person existing at another time. So far, I have argued that sequentialists have not been successful in establishing the first claim. I will now argue that the second claim is problematic as well.

My argument is based on the observation that the second claim yields strikingly counterintuitive results. For example, it implies that we should be indifferent between our own wellbeing and the wellbeing of our psychological twin. Suppose that there exists what Hilary Putnam has envisaged as Twin Earth. Then there should be someone somewhere on Twin Earth whose psychological states are exactly like mine. I shall refer to him as *my twin*. My current person-stage must be psychologically sequential to my twin's person-stage a second later, because that person-stage is qualitatively identical to my own stage a second later, which is psychologically sequential to my current stage. If maintaining psychological sequentiality is sufficient for the preservation of what matters in survival, it follows that I ought not to care about my future selves any more than I care about the future selves of my twin. This, however, seems implausible. If I am told that I shall be tortured tomorrow, it seems that I ought to be worried about myself in a way that I would not be worried about my twin when I am told that he will be tortured tomorrow. Similarly, if either Earth or Twin Earth has to be destroyed, the prospect of Earth's destruction would worry me in a way that the prospect of Twin Earth's destruction would not. The kind of concern that I would attach to my twin seems to be no different from the kind of concern that I would attach to a remote stranger. My intuition here can be partly justified by the fact that what caused the psychological contents of my twin would be different from what caused my psychological contents. For example, when I look at a tree out of an interest for bonsai, my twin will look at a tree as well. But he would not be seeing the same tree; moreover, two (numerically) different sets of neurons will be activated when each of us looks at the trees. In general, many of the objects in his experience and consciousness would be numerically distinct from the corresponding objects in my experience and consciousness.²³ Hence, many of his memories, beliefs, desires, etc. that are central in constructing what matters to him would not refer to the persons, animals, events,

²³ Here I say ‘many’ rather than ‘all’ because, if there are such things as universals as some have argued, or if God exists, my thoughts about them will be exactly the same as the corresponding thoughts of my twin in content.

state of affairs, etc. figuring in the bulk of the psychological items that constitute what matters in my survival.

To make the implausibility of the second claim even more salient, imagine a countless number of random object generators spread out over the universe. The function of a random object generator is to arrange a collection of particles in a random fashion, thereby generating a material object only to destroy that object instantly. Each machine repeats this procedure over and over again. I presume that the vast majority of products generated by these machines will not exhibit even a single shred of consciousness. However, if there exist a sufficiently great number of those machines, then we may imagine that some of them will generate an object with a human figure whose brain is appropriately structured so that it can give rise to a complex array of conscious thoughts and feelings just as in normal human beings. Although such beings will cease to exist after just a second, if there *really* are a great number of those machines, then there may exist a being somewhere in the universe a moment from now whose psychological state will be exactly like your psychological state a moment from now, and therefore sequential with your psychological state now. In this way, the number of machines may be increased to the extent that any possible psychological state of any person at any given moment is now being instantiated by some machine somewhere in the universe. If this were indeed the case, then Psychological Sequentialism entails that so long as you are aware of the existence of those machines, you ought not to be afraid of your death. For instance, suppose you are hit by a train and killed instantly at t_1 . Then, *ex hypothesi*, there will be someone somewhere in the universe whose psychological state is exactly like how your psychological state would be a moment after t_1 had you not died, and who is thus psychologically sequential to you at t_1 . By the time that being ceases to exist, say at t_2 , there will exist another being, somewhere in the universe, psychologically sequential to that being at t_2 , and so on. Hence, there will always be someone who will be psychologically sequential to you as you are at t_1 , even after your death. If psychological sequentiality is sufficient for the preservation of what matters, you ought to consider this case as tantamount to one where you will survive the train accident. This strikes me as untenable. If I knew that I will soon be dead, the prospect that my death will be followed by the creation of a series of ephemeral beings would not provide much relief, even if their psychological states were sequential to mine.²⁴

I have argued that Psychological Sequentialism fails to establish that no causal constraint is required for the preservation of what matters in survival. I also have illustrated some counterintuitive consequences implied by Psychological Sequentialism. I suggest that these considerations are sufficient to reject the view.

²⁴ One might argue that it is not at all physically likely that a successive series of sequential psychological states is instantiated in the manner I have suggested above. However, the realism of the example is not an issue here. The point of the example is that for most of us the *prospect* of a series of psychological states merely sequential to one's own psychological states does not seem to provide the kind of relief typical to the preservation of what matters in survival.

Acknowledgments This work was supported by the 2010 Research Fund of UNIST. I also would like to thank Matt Hanser, Boram Lee, and anonymous referees for *Axiomathes* for their detailed and insightful comments on earlier drafts of this article.

References

- Brueckner A, Buford C (2013) Against Psychological Sequentialism. *Analysis* 73(1):96–101
- Brueckner A, Fischer JM (1986) Why is death bad? *Philos Stud* 50(2):213–221
- Campbell S (2005) Is causation necessary for what matters in survival? *Philos Stud* 126(3):375–396
- Dick P (2002) *The minority report and other classic stories*. Citadel Press, New York
- Elliot R (1991) Personal identity and the causal continuity requirement. *Philos Q* 41(1):55–75
- Feldman F (1991) Some puzzles about the evil of death. *Philos Rev* 100(2):205–227
- Frankfurt H (1969) Alternate possibilities and moral responsibility. *J Philos* 66(23):828–839
- Hirsch E (1982) *The concept of identity*. Oxford University Press, New York
- Kaufman F (1996) Death and deprivation: or why lucretius' symmetry argument fails. *Australas J Philos* 74(2):305–312
- Kolak D, Martin R (1987) Personal identity and causality: becoming unglued. *Am Philos Q* 24(4):339–347
- Lewis D (1986) *Philosophical papers, vol II*. Oxford University Press, Oxford
- Lewis D (2000) Causation as influence. *J Philos* 97(4):182–197
- Lucretius (2007) *The nature of things* (trans: AE Stallings, an introduction by R Jenkyns). Penguin Books, New York
- McMahan J (2006) The Lucretian argument. In: McDaniel K, Raibley JR, Feldman R, Zimmerman MJ (eds) *The good, the right, life and death: essays in honor of Fred Feldman*. Ashgate Publishing, Burlington, pp 213–226
- Nagel T (1970) Death. *Noûs* 4(1):73–80
- Nozick R (1981) *Philosophical explanations*. Harvard University Press, Cambridge
- Parfit D (1984) *Reasons and persons*. Clarendon Press, Oxford
- Siderits M (1988) Ehring on Parfit's relation R. *Analysis* 48(1):29–32
- Van Inwagen P (1983) *An essay on free will*. Oxford University Press, Oxford