The Brave Officer Rides Again

Andreas L. Mogensen^{1,2}

Received: 3 July 2015 / Accepted: 17 March 2017 / Published online: 17 April 2017
© The Author(s) 2017. This article is an open access publication

Abstract According to the *Psychological Account* of personal identity, personal identity across time is maintained by some form of psychological overlap or continuance. I show that the *Psychological Account* has trouble accommodating cases of transient retrograde amnesia. In such cases, the transitivity of psychological continuity may break down. I consider various means of responding to this problem, arguing that the best available response will undercut our ability to rely on intuitions about brain transplantation to support the *Psychological Account*. When the *Psychological Account* is re-interpreted in such a way as to overcome the problems posed by cases of transient retrograde amnesia, it turns out that standard brain transplantation cases involve details which are strictly irrelevant according to the theory. On the other hand, these features seem to drive our intuitions about survival in these cases. In this way, my argument undermines one of the key motivations for adopting the *Psychological Account*.

1 Introduction

According to one popular theory, personal identity across time is maintained by some form of psychological overlap or continuance.¹ Call this (admittedly vague) theory the *Psychological Account*. The chief advantage of the theory is that it is able

¹ For key statements of this view, see Grice (1941), Lewis (1976), Locke (1689), Noonan (2003), Parfit (1984), Perry (1972), Shoemaker (1963, 1984, 2008). According to the PhilPapers Survey, the psychological view of personal identity is nearly twice as popular as the biological view: see Bourget and Chalmers (2014).

✉ Andreas L. Mogensen
andreas.mogensen@philosophy.ox.ac.uk

¹ Jesus College, Turl Street, Oxford OX1 3DW, UK

² Faculty of Philosophy, University of Oxford, Oxford, UK

to capture the widely-shared intuition that a person could survive without her body and other organic parts, so long as she remains psychologically connected to past stages of herself: for example, if her brain is transplanted from one body to the next, with memory, personality, etc. intact.² The *Psychological Account* continues to have its detractors, especially among defenders of the *Animalist Account*.³ According to proponents of this view, human persons have the biological persistence conditions of animals, which do not include psychological continuance as necessary or sufficient.⁴ The *Animalist Account* has the apparent defect of being unable to capture those intuitions, mentioned above, which are taken to support the *Psychological Account*: the *Animalist Account* suggests that a person cannot survive without her body.⁵ Defenders of the *Animalist Account* nonetheless maintain that this view wins out on balance.

In this paper, I offer some help to Animalists. I show that the *Psychological Account* has difficulty accommodating cases of transient retrograde amnesia. In Sects. 2 and 3, I show that in cases of transient amnesia, psychological continuity can be intransitive. In Sect. 4, I argue that proponents of the *Psychological Account* who take identity to consist in some form of *non-branching* continuity may accommodate such cases by treating them as instances of branching continuity, similar to cases of ‘fusion’. In Sects. 5 and 6, I argue that in order to treat cases of transient amnesia in this way whilst avoiding results that seem strikingly counter-intuitive, the *Psychological Account* must submit to a substantial loss in terms of its chief advertised advantage over the *Animalist Account*. The modifications required in the *Psychological Account* to accommodate transient amnesia imply that the standard cases designed to show that a person could survive without her body are not adequate as test-cases for the theory. When these cases are appropriately re-described so as to exclude details which should be irrelevant according to the modified *Psychological Account*, the intuition that these cases involve survival recedes. Section 7 establishes that the same holds true if we favour a four-dimensionalist construal of the *Psychological Account* that treats cases of branching continuity as involving shared temporal parts. Section 8 wraps things up.

2 Transient Amnesia and the Intransitivity Problem

There are many psychological relations that may be taken to support the kind of overlap required for identity according to the *Psychological Account*: the retention of beliefs and personality traits; the connection between an intention and its execution; etc. Here, I focus on the old favourite: the relationship between an experience and the memory of it. In the ensuing discussion, I will operate as if we are dealing with individuals for whom psychological overlap can only be provided

² Shoemaker (1963) is the *locus classicus* for this thought experiment.

³ E.g., Olson (1994, 1997, 2007), Snowdon (1995), van Inwagen (1990, 1997).

⁴ This, at least, is the standard interpretation of the *Animalist Account*. For discussion of the view that human animals may persist by way of psychological continuance, see Bailey (2015) and Thornton (2016).

⁵ See Olson (2007, 39–44) for discussion.

by such memorial connections. This assumption is unrealistic, but yields a manageable framework in which to test the plausibility of the theory.

There is a naïve version of the *Psychological Account*, which everyone agrees is untenable. Following Parfit (1984), let us call the connection between an experience and the memory of that experience a *direct psychological connection*. Let us say that if there are a sufficiently many direct connections between x_1 at t_1 and x_2 at t_2 , then x_1 and x_2 are *strongly connected*. The naïve theory says:

The Naïve Psychological Theory:

x_1 at $t_1 = x_2$ at t_2 iff x_1 at t_1 is strongly connected to x_2 at t_2 .

This theory has a fatal flaw: strong connectedness cannot be the criterion of identity over time, because strong connectedness is not a transitive relation. The problem is famously illustrated by Reid's (1785) example of the brave officer. Imagine three different stages in a possible life: at t_1 , R_1 is a child stealing from an orchard; at t_2 , R_2 is a brave officer capturing an enemy standard; and at t_3 R_3 is a retired general. The retired general might remember a great deal from his time as a brave officer but have no memory of being a thieving child: thus, there is strong connectedness between R_3 and R_2 , but not between R_3 and R_1 . The brave officer, by contrast, remembers a great deal from his time as a thieving child: thus, there is strong connectedness between R_2 and R_1 . The *Naïve Psychological Theory* then implies that $R_3 = R_2$, that $R_2 = R_1$, and that $R_3 \neq R_1$, which is impossible.

There is a well-known move that is supposed to revise the *Psychological Account* so that it avoids this problem. Rather than taking personal identity to consist in strong connectedness, one takes personal identity to consist in *psychological continuity*, which involves "overlapping chains of *strong connectedness*" (Parfit 1984, 206. Emphasis in text). The revised psychological theory then says:

The Revised Psychological Theory:

x_1 at $t_1 = x_2$ at t_2 iff x_1 at t_1 is psychologically continuous with x_2 at t_2 .

The standard wisdom is that this revision solves the intransitivity problem. It implies that $R_3 = R_2$, that $R_2 = R_1$, and that $R_3 = R_2$, since R_3 is linked to R_1 by a chain running via R_2 . Psychological continuity is said to be a transitive relation by Parfit (1984, 206) and McMahan (2002, 40), and hence of the right form to be constitutive of identity over time.

However, psychological continuity is not in fact a transitive relation. We need only a very simple modification of Reid's example in order to bring out the problem. The key is to make R_2 rather than R_3 the one who cannot remember being R_1 .

Here is the example.⁶ Suppose that at t_1 , R_1 is a child stealing from an orchard; at t_2 , R_2 is a brave officer who has just been hit on the head by the butt of an enemy rifle and suffers transient retrograde amnesia; at t_3 R_3 is a retired general with an excellent memory of his entire life. At t_2 , R_2 has transient retrograde amnesia: he cannot remember any experience prior to the concussion, though he can still form

⁶ I'm grateful to an anonymous referee for drawing my attention to a similar case discussed by Perry (1975). My paper should perhaps have been called "The Brave Officer Rides Again (Again)".

new memories. There is therefore no connectedness or continuity linking R_2 and R_1 . However, memories gradually return, and new ones are formed. At t_3 , R_3 can remember enough about his t_1 childhood for there to be strong connectedness between R_3 and R_1 ; since a capacity for forming new long-term memories did not vanish at t_2 , R_3 is also continuous with R_2 via direct and/or ‘chained’ memory connections.

In the example, psychological continuity links R_3 to R_1 and R_3 to R_2 . However, there is no continuity between R_2 and R_1 , due to R_2 ’s transient retrograde amnesia at t_2 . Psychological continuity therefore relates R_1 to R_3 , R_3 to R_2 , but not R_1 to R_2 .⁷ Hence, psychological continuity is not in fact a transitive relation. In the described scenario, the *Revised Psychological Theory* apparently implies that $R_1 = R_3$, that $R_3 = R_2$, and $R_1 \neq R_2$, which is impossible.

3 Can Memory Vanish and Return?

The argument in Sect. 2 turns upon the possibility that memory for certain events could disappear at one time, only to reappear at a later time. It is in virtue of this phenomenon that psychological continuity is not transitive in the example described. One means of responding to the argument, therefore, is by querying whether memory can genuinely vanish and return. If not, my argument is spurious.

To develop this query, it will be helpful to consider a case of transient retrograde amnesia from the annals of neurology, and to reflect a little on what is required in order to count as having a memory of an experience. In a case described by Papagno (1998), the patient, CL, was a 26-year-old male, who fell from a scaffold whilst working in the cathedral of Milan. After the fall, “he quickly regained consciousness, but could not remember his name, what he was doing there or any autobiographical detail.” (Papagno 1998, 112) There was no associated anterograde amnesia. Over the course of 10 days, CL

started to regain more and more images of his past life, due to friends showing him photos of holidays they had spent together or of himself in some particular occasion. And progressively all his life became clear and vivid, especially, he said, with the help of a longtime friend. (Papagno 1998, 114)

It seems appropriate to say that after his fall CL had lost all memory of his past life, but that he regained these memories after 10 days. Further reflection makes this description less clearly appropriate. Cognitive psychologists divide the faculties associated with human memory into three: *encoding*, by which a memory trace (or ‘engram’) is created; *storage*, by which a durable record of this information is preserved; and *retrieval*, by which the durable record is accessed. CL suffered no noticeable anterograde amnesia: he had no difficulties encoding new information. Was his inability to recall events from his past due to a loss of stored information or to an impairment of the retrieval-process? On this, Papagno is unequivocal:

⁷ I assume here that psychological continuity is a symmetric relation.

A loss of stored engrams can be ruled out, because the patient did not just relearn, but progressively remembered his past life and world events, when provided with repeated cues. ... [T]his case suggests the possibility of a primary retrieval deficit, i.e., an inhibition (psychogenic and/or organic) of retrieval without storage deficit. (Papagno 1998, 119–120)

In other words, CL's brain apparently retained information about his past life, but this information was momentarily inaccessible. We may want to say, therefore, that he *did* remember events from his past throughout the entire period.

My linguistic intuitions on this are unclear. Let us say that according to the *maximal conception of memory*, one counts as having a memory of a past experience only if information pertaining to that experience is both stored *and* retrievable; according to the *minimal conception*, one counts as having a memory of a past experience purely in virtue of retaining information related to that experience, whether or not this information can be retrieved.

On the maximal conception, memories are clearly able to vanish and then reappear. It is an entirely different question whether 'minimal memory' can also vanish and return. One might insist, with some plausibility, that it cannot. Consider again Papagno's patient, CL. If CL had genuinely lost retention of any information pertaining to his past, he might have gone on to learn facts about that life from the reports of friends and family; and although he might thereby have come to know about his past life, he would not genuinely have come to *remember* anything from that life by this process, even if the descriptions offered by friends and relatives were so vivid as to create apparent memories of past experiences. Causal processes of this kind are not of the right kind for the resulting states to count as memory. Similar points would appear to apply to any other conceivable means by which he might gain information about his life: looking at old photographs; watching home-movies, etc. None of these methods can yield genuine memory, even if they yield knowledge. Hence, we might conclude that on the minimal conception, if CL's memories had genuinely vanished, he could not have recovered any memory for his past life. And so it may be thought that if we understand memory according to the minimal conception, this rules out the possibility that psychological continuity could fail to be transitive in the way that my revision of Reid's example was taken to suggest.⁸

However, it is in fact possible for a person to have a minimal memory of a past experience at t_1 , have no minimal memory pertaining to the experience at t_2 , and have a minimal memory pertaining to the experience again at t_3 . Consider a hypothetical individual whose brain creates and stores *two* memory-traces relating to each experience via two different causal processes, both triggered by the same token experience. These processes are otherwise exactly like the processes which create human memory-traces.⁹ Suppose that the processes proceed at different speeds due to differences in the myelination of the relevant axons. Let *P1* be the faster process and *P2* the slower. Then, *P1* might create a memory-trace, and the

⁸ Baillie (1993, 195) suggests that appealing to the preservation of minimal memory may help proponents of the *Psychological Account* to accommodate certain intuitions about amnesia, though he does not discuss the particular issue that concerns us here.

⁹ Actually, for all I know, human memory might be duplicative in this sense.

interval between the conclusion of *P1* and the conclusion of *P2* could allow sufficient time for the first memory-trace to be eliminated and then restored by *P2*. Hence, minimal memory for a past event can be lost and regained without invoking causal processes which are intrinsically dissimilar to those that underwrite ordinary memory.

Since memory can be lost and regained even on the minimal conception, I conclude that there is no possibility of undermining my argument in the previous section by suggesting that memories cannot re-appear. I conclude that psychological continuity is not in fact a transitive relation, unlike identity.

4 Branching Continuity

This need not be a decisive objection to the *Psychological Account*. I've tried to show that psychological continuity is not a transitive relation. It does have a property that is superficially similar to transitivity: it is transitive when its arguments are ordered linearly in time. Thus, if x_1 at t_1 is continuous with x_2 at t_2 and x_2 at t_2 is continuous with x_3 at t_3 , x_1 is continuous with x_3 ; and if x_3 at t_3 is continuous with x_2 at t_2 and x_2 at t_2 is continuous with x_1 at t_1 , x_3 is continuous with x_1 . However, transitivity can fail when the terms are not ordered in this way.

Although I know of no one who has made this point by reference to cases of transient amnesia, it has been noted before: for example, by Parfit (1984, 302) in his discussion of *branching continuity*.¹⁰ In this section, I consider whether the revisions in the *Psychological Account* that have been suggested in order to accommodate cases of branching continuity might yield a theory that also avoids the unpalatable implications engendered by the *Revised Psychological Theory* when it is applied to my revision of Reid's example.

Branching continuity occurs when two distinct individuals are both continuous with some individual existing at another point in time. The most famous case of branching is *fission*, which occurs when x_1 at t_1 is continuous with two distinct individuals, x_2 and x_3 , both existing at some time later than t_1 . Fission might conceivably occur if each hemisphere of a person's brain were transplanted into a different body, yielding two individuals who believe themselves to be the earlier person, remember events from her life, and so on. Fission creates obvious problems for the *Revised Psychological Account*. Since x_1 is continuous with both x_2 and x_3 , the *Revised Psychological Account* implies that $x_2 = x_1$ and that $x_1 = x_3$, which contradicts the assumption that $x_2 \neq x_3$. Consideration of fission-cases has therefore led to wide-scale abandonment of this theory. In its place, contemporary proponents of the *Psychological Account* may endorse some variant of the following:

The Non-Branching Theory:

x_1 at $t_1 = x_2$ at t_2 iff x_1 at t_1 is linked with x_2 at t_2 by non-branching psychological continuity.¹¹

¹⁰ See also Perry (1972, 468).

¹¹ For views of this kind, see Parfit (1984) and Shoemaker (1984).

Alternatively, they might favour a four-dimensionalist (re-)interpretation of the *Revised Psychological View* along the lines proposed by Lewis (1976). On this view, we assume individuals are not wholly present at each time at which they exist: rather, they have temporal parts, which we call *person-stages*. Person-stages are said to be *I-related* when they are stages in the life of the same individual person. We can then say

The 4-D Theory:

- (i) A person is a maximal I-interrelated aggregate of person-stages.
- (ii) Person-stage σ_1 at t_1 is I-related to person-stage σ_2 at t_2 iff σ_1 at t_1 is psychologically continuous with σ_2 at t_2 .

Unlike the identity relation, the I-relation need not be construed as transitive. This allows us to say that in cases of fission, prior to the point of division, there exists a person-stage that is a stage in the lives of two different people: i.e., the two distinct post-fission individuals. But it is misleading to describe these as ‘post-fission individuals’. Since they each have a stage located prior to the point of fission, both individuals existed prior to that point. At that point, they were, in effect, co-located in the same body.¹²

My own sympathies rest with the view that we endure rather than perdure, and so I regard the *Non-Branching Theory* as significantly more plausible. From here on, I will concentrate on discussing this version of the *Psychological Account*. I will return to the *4-D Theory* in Sect. 7, once I’ve made clear how the *Non-Branching View* deals with cases of transient amnesia.

As I say, I’ll focus on the *Non-Branching Theory*. Does the *Non-Branching Theory* have the same implication as the *Revised Psychological Theory* with respect to my re-description of Reid’s example? That is: does it also imply that $R_1 = R_3$, that $R_2 = R_3$, but that $R_1 \neq R_2$? Certainly, my re-description of Reid’s example involves nothing like fission: it is not a case in which one individual is continuous with two later individuals who are plausibly regarded as numerically distinct. If fission were the only kind of branching continuity, the *Non-Branching Theory* would have the same implications as regards my revision of Reid’s example as the *Revised Psychological Theory*.

But fission is not the only kind of branching continuity. A less widely discussed form of branching is *fusion*, which occurs when x_1 at t_1 is continuous with two distinct individuals, x_2 and x_3 , both existing at some time *earlier* than t_1 . Fusion cases are somewhat harder to imagine, but still apparently conceivable: two persons might come together, their bodies could grow into one, and the resulting person would wake up with psychological attributes linking her to both of the earlier individuals. In cases of fusion, the *Non-Branching Theory* implies that the two earlier individuals have ceased to exist, since there is no later person to whom they are linked by non-branching continuity. The new individual is a numerically distinct person.

¹² Lewis argues that this overpopulation problem can be mitigated by noting that we do not always count by identity. See Sider (1996) for doubts about Lewis’s views on counting and identity.

Consider a case of fusion involving two persons, x_1 and x_2 . Suppose that x_1 is older than x_2 : x_1 exists at a time, t_1 , at which point x_2 does not yet exist; x_2 begins to exist at t_2 ; at t_3 , they fuse, yielding another individual, x_3 . We should then say that x_1 at t_1 is continuous with x_3 at t_3 , that x_3 at t_3 is continuous with x_2 at t_2 , but that x_1 is not continuous with x_2 . In this way, cases involving fusion may involve just the kind of non-transitive psychological continuity exhibited in my revision of Reid's example. By contrast, there is no intransitive, *non-branching* continuity in such cases: x_1 is not linked via non-branching continuity to x_2 or x_3 ; and they are not linked by non-branching continuity to one another. There is, in fact, no non-branching continuity whatsoever.¹³

This may suggest that the *Non-Branching Theory* can readily accommodate my revision of Reid's example by treating transient amnesia as akin to fusion. Proponents of the *Non-Branching Theory* could say that R_1 and R_2 are distinct persons, since they are not linked by non-branching psychological continuity. They could then regard R_3 as an individual who is like a fusion of the two, since he is continuous with both. They could say that although my revision of Reid's example involves non-transitive relations of psychological continuity, there are no non-transitive relations of non-branching continuity: because the case is like fusion, there is simply no non-branching continuity in play. Hence, my case poses no challenge to the *Non-Branching Theory*.

I argue in the next two sections that regarding transient amnesia as akin to fusion is nonetheless very costly to the *Psychological Account*.

¹³ Consideration of fusion cases also allows us to block an objection that could have been raised against my argument in Sect. 2, noted by an anonymous referee. If we consider only R_1 at t_1 , R_2 at t_2 , and the interval t_1 – t_2 , there are no psychological links that allow us to count R_2 and R_1 as continuous. But, assuming that continuity is symmetric, R_1 is continuous with R_3 who is continuous with R_2 . Could we not say on that account that the problem I raised for the *Revised Psychological Theory* is spurious, since R_1 and R_2 are linked by overlapping chains of strong connectedness, albeit chains that take an unusual path, looping forward in time and then backward? We could say this, but if we understand psychological continuity as a relation that can be instantiated by forward–backward looping chains, the *Revised Psychological Theory* entails mistakenly that $x_1 = x_2$ in the fusion case. The referee suggests that this implication might be avoided if we adopt instead an interpretation of the I-relation discussed by Brueckner (2005): person-stages σ_1 and σ_2 are I-related iff (i) σ_1 and σ_2 are related by psychological continuity and (ii) there is no person-stage, σ_3 , such that either (a) σ_1 is psychologically continuous with σ_3 , σ_3 exists simultaneously with σ_2 , and $\sigma_3 \neq \sigma_2$; or (b) σ_2 is psychologically continuous with σ_3 , σ_3 exists simultaneously with σ_1 , and $\sigma_3 \neq \sigma_1$. But if continuity can be realized by forward–backward looping chains, this principle has equally problematic implications. Imagine two successive person-stages that are part of x_1 prior to fusion, σ_1 and σ_2 , and two successive person-stages that are part of x_2 prior to fusion, ζ_1 and ζ_2 . Assume that σ_1 and ζ_1 are simultaneous, as are σ_2 and ζ_2 . Obviously, $\sigma_1 \neq \zeta_1$ and $\sigma_2 \neq \zeta_2$. If continuity can be realized by forward–backward looping chains, σ_1 is continuous with ζ_2 via a chain that runs forward through σ_2 to the point of fusion and then back through the intervening stages comprising x_2 prior to fusion. Thus, σ_1 is continuous with a person-stage simultaneous with but distinct from σ_2 . Condition (ii) (a) then entails, mistakenly, that σ_1 and σ_2 are not stages in the life of one and the same person.

5 Transient Amnesia as Fusion?

There is one very obvious problem for the proposal. When two individuals fuse, according to the *Non-Branching Theory*, both cease to exist and a new person comes into being. Hence, if transient amnesia is just like fusion, in my revision of Reid's example there is not a single individual who suffers transient amnesia, but three distinct individuals existing successively. Consider how implausible this sounds as applied to the case described by Papagno. It implies that after his fall from the scaffold, CL ceased to exist and a new person came to inhabit his body. This new person did not live for very long. After ten days, he ceased to exist, and yet another person came to inhabit CL's body. This seems strikingly counterintuitive. The common-sense view, embodied in Papagno's description of the case, is that CL suffered temporary retrograde amnesia—which is to say that he lost his memories and then regained them, existing throughout.

Should we expect defenders of the *Non-Branching View* to concede the implausibility of reading Papagno's report as a story about three numerically distinct individuals? After all, they are willing to posit three such individuals in standard cases of fission and fusion. What's more, the psychological links that bind the three in standard fission cases are, if anything, more robust than the psychological links present in the transient amnesia case described by Papagno. Hence, positing the existence of three distinct individuals in the latter case may be thought *less* troubling overall.¹⁴

I think proponents of the *Non-Branching View* have a significantly greater problem in biting the bullet with respect to this case. Here is why. The way the *Non-Branching Theory* handles fission cases *is* counter-intuitive: after all, it makes the identity of *x* and *y* depend on things besides relations between *x* and *y*. For that matter, any other way of accommodating fission cases using the *Psychological Account* carries implications that are, in my view, no less strange.¹⁵ The theory's saving grace is that any other minimally plausible criterion of personal identity faces the same kind of problem. Any such criterion apparently ties identity to a relation that is able to instantiate a branching structure, yielding possible cases of reduplication or division—unless, that is, the theory includes an explicit 'no branching' clause.¹⁶

Positing the existence of multiple individuals in the Papagno case is far more problematic, as the difficulty faced by the *Psychological Account* here seems a problem for this theory in particular. This isn't the sort of case about which everyone ends up committed to saying something surprising and odd, where we just have to concede that some strongly held intuition has to go. The *Animalist Account*

¹⁴ I am grateful to an anonymous referee for this objection.

¹⁵ For example, the four-dimensionalist solution favoured by Lewis (1976) and Noonan (2003) requires us to say that there already existed more than one person prior to the point of fission.

¹⁶ See Perry (1976), Noonan (2003, 125–127). The *Simple View* (Lowe 1998; Swinburne 1984) does not encounter exactly this problem, but still requires us to make counter-intuitive judgments with respect to fission cases: for example, it seems to require us to accept the possibility that a person might go along with her transplanted left hemisphere as opposed to her transplanted right hemisphere simply as a matter of brute fact.

straightforwardly allows us to capture the intuition that a single individual, CL, persists throughout the events described in Papagno's report. Moreover, I know of no analogous cases in which patterns of organismic continuity and discontinuity force animalists to say we are dealing with instances of fusion, contrary to our pre-theoretic intuitions.¹⁷

Granting, then, that construal of CL's case as involving fusion would significantly lower the relative plausibility of their theory, there are two ways in which proponents of the *Psychological Account* may try to avoid the implication that Papagno's report does not describe the life of a single, persisting individual. The first is as follows. I said at the outset that I would be operating as if we were dealing with individuals for whom psychological connectedness could only be provided by memorial connectivity. This assumption applied in my revision of Reid's example. The proponent of the *Psychological Account* may point out that CL is *not* such an individual: since he is a normal human being, there are many other relations which could provide his amnesiac self with strong connectedness to former stages of himself. Hence, after his fall, CL might remain linked to past stages of himself via non-branching continuity, despite his amnesia.

A problem for this reply is that Papagno's report indicates that other forms of psychological connectedness were also disrupted after CL's accident. Semantic memory, comprising beliefs and knowledge about the world, was momentarily disrupted along with episodic memory; CL also exhibited startling temporary changes in personality and preference:

he was amazed that he had a tattoo on his shoulder, which he really disliked, and could not remember having done it two years before; he hated his very short-cut hair, but he admitted that if he was someone with such a tattoo, he could also be somebody styling his hair in that fashion and could also be somebody taking dope (which apparently was not true). People who knew him, said that his mood and behaviour were different than usual: he was now very extroverted, talkative, sociable, while he was known as timid, introverted, silent, unsure. (Papagno 1998, 113)

Unless there is some other way to avoid the implication, therefore, *the Non-Branching View* is in danger of yielding the following startling consequence: *contra* Papagno's report, no one individual fell from the scaffold, woke up with amnesia, and then regained his memories. Instead, one individual fell from the scaffold, another woke up, and yet another came to have extensive apparent memories pertaining to the life of both men. This, I take it, is revisionary of common sense.

There is, however, another way in which defenders of the *Psychological Account* might seek to avoid the implication that CL ceased to exist as a result of his accident. In Sect. 3, I discussed the possibility of understanding the notion of psychological continuity in terms of the minimal conception of memory, so as to avoid the implication that my revision of Reid's example involves non-transitive psychological continuity. That attempt failed. However, proponents of the *Non-*

¹⁷ This may, of course, reflect a failure of imagination on my part. At the very least, I feel the onus is on defenders of the *Psychological Account* to describe such cases.

Branching View might insist that their theory should be understood in such a way that in order for non-branching chains of memory-connectedness to support identity over time, the memories in question need only be minimal memories. If we wed the *Non-Branching View* to the minimal conception of memory in this way, we can say that only cases of transient amnesia involving the loss and recovery of minimal memories are to be regarded as akin to fusion. This would clearly avoid the unpalatable implication: as we recall, Papagno's conclusion was that CL's memory-loss resulted from a retrieval deficit, and not from the temporary loss of stored engrams.

Let us call this interpretation of the *Non-Branching View* the *Minimal Non-Branching View*. We may have some doubts as to whether this view *really* solves the problem. Suppose it was discovered that Papagno was treating a patient with a duplicative memory like that described in Sect. 3, and that the gradual return of memory was not due to the return of retrieval capacity, but rather to the slower memory-process, *P2*, finally completing its encoding of memories already previously encoded by *P1*. Would it then be substantially less troubling to say that there were in fact three different men—one who fell, one who could not recall anything, and yet another who had apparent memories of the life of the first?¹⁸

My intuitions on this are not very clear, though others' may be. I'm going to set aside this question in order to concentrate on a different problem for the *Minimal Non-Branching View*.

6 The Minimal Non-Branching View and Brain Transplantation

The most well-known advantage of the *Psychological Account* over the *Animalist Account* is supposed to be that only the former is able to capture the intuition that a person could survive the destruction of her body, so long as she remains psychologically connected to past stages of herself. However, the thought-experiments that are used to draw out this intuition always involve connectivity of *maximal* memory. For example, here is Shoemaker's original description of brain transplantation:

One day, to begin our story, a surgeon discovers that an assistant has made a horrible mistake. Two men, a Mr. Brown and a Mr. Robinson, had been operated on for brain tumors, and brain extractions had been performed on both of them. At the end of the operations, however, the assistant inadvertently put Brown's brain in Robinson's head, and Robinson's brain in Brown's head. One of these men immediately dies, but the other, the one with Robinson's body and Brown's brain, eventually regains consciousness. Let us call the latter "Brownson." Upon regaining consciousness Brownson exhibits great shock and surprise at the appearance of his body. Then, upon seeing Brown's body, he exclaims incredulously "That's me lying there!" Pointing to himself he says "This isn't my body; the one over there is!" When asked his name he

¹⁸ Thanks to Jens Johansson for suggesting this point to me.

automatically replies “Brown.” He recognizes Brown’s wife and family (whom Robinson had never met), and is able to describe in detail events in Brown’s life, always describing them as events in his own life. ...

What would we say if such a thing happened? There is little question that many of us would be inclined, and rather strongly inclined, to say that while Brownson has Robinson’s body he is actually Brown. (Shoemaker 1963, 23–24)

Suppose we think that the *Psychological Account* must treat retrieval capacity as inessential to persistence via memory continuity. What happens if we re-imagine Shoemaker’s case such that there is only connectivity of minimal memory? Suppose that Brownson wakes up, and is entirely unable to recall anything from Brown’s life. When asked his name, he does not know; when introduced to Brown’s family, he does not recognise them and seems never before to have met these people; when he is asked when his birthday is, he cannot recall ever having had a birthday. Everything is a blank. The surgeon concludes that the removal of the tumour from Brown’s brain severely damaged the neural circuits required for retrieval. All previous memory-traces remain encoded in the brain now inhabiting Robinson’s body, but are entirely inaccessible. Brownson dies within two days, due to complications arising from the surgery. In this case, I think, the verdict that Brown survived the operation as Brownson is far from compelling.¹⁹

Now, Shoemaker’s description of the brain transplant case actually goes on to discuss psychological relations besides memory retention, which are supposed to link Brown and Brownson: for example, we are told that they have exactly similar personality traits. In line with my stipulation in Sect. 2, I have omitted this sort of detail in my discussion above. I believe, however, that analogous problems involving intransitivity due to transient suspension can be generated for these relations: consider the many forms of transient suspension in psychological continuity involved in CL’s case. Transient suspensions in these other dimensions of psychological continuity raise problems for the *Psychological Account* analogous to those arising from transient retrograde amnesia. In addressing these problems, the dialectic will unfold similarly. Analogous minimal/maximal distinctions will be required in order to accommodate cases like CL’s, whereas retreating to the view that identity requires only minimal retention of psychological features will generate

¹⁹ An anonymous referee raises the following objection. Could proponents of the *Psychological Account* not try to explain away the problem, arguing that we lack the intuition that Brown survives as Brownson in this case simply because our epistemic situation doesn’t allow us to differentiate between total memory loss and mere loss of maximal memory? I don’t believe so. In considering the case, I’ve asked us to assume that the surgeon conducts an investigation that allows her to establish that Brown’s minimal memories are retained but are irretrievable. The evidence available to the surgeon and/or her testimony seems to ensure that our epistemic situation differentiates between the two possibilities. The force of Shoemaker’s original thought-experiment does not seem to require imagining that we have direct experience of Brownson’s post-operative state: imagining that we have a credible report from the acting surgeon seems good enough to trigger the desired intuition. Given that intuitions attributing identity can be elicited by testimony that psychological features are retained in their maximal form, it would seem they ought also to be elicited by testimony regarding minimal retention—unless it is the case that these intuitions are selectively concerned with connectivity and continuity in a maximal sense, as I am suggesting.

the same problem that we are here considering: the standard transplant case intuition is unlikely to be triggered if the case is described as involving retention of such states in a merely minimal form, as stored but inexpressible.

There is a general lesson to be learned. To the extent that they support the *Psychological Account*, our intuitions about identity seem to be guided by the retention of psychological features in their *maximal* form; but we cannot let identity hinge on the maximal retention of psychological features, unless we want to end up having to say quite implausible things about cases of transient amnesia like CL's. That is the dilemma ultimately confronting proponents of the *Psychological Account*.

7 Four-Dimensionalism

To make clear that this dilemma really does confront any proponent of the *Psychological Account*, let me briefly discuss why favouring the *4-D Theory* over the *Non-Branching Theory* offers no advantages here.

Proponents of the *4-D Theory* could also construe cases of transient amnesia as akin to fusion. They seem forced to do so when considering the revision of Reid's example that I discussed in Sects. 2 and 3. Understood as person-stages, R_1 and R_2 are not I-related, whereas R_2 and R_3 are I-related, as are R_1 and R_3 . So understood, the case looks to involve a pattern of I-relations like that which obtains in cases of fusion, as such cases are understood by the four-dimensionalist.

On the *4-D Theory*, fusion cases are in some sense both more and less crowded than on the *Non-Branching View*. When two individuals fuse, the *4-D Theory* implies that they continue to exist. Where proponents of the *Non-Branching View* would posit a new post-fusion individual, the *4-D Theory* posits a sequence of person-stages that are stages in the life of *both* pre-fusion individuals. Fusion therefore involves one less person overall. On the other hand, the *4-D Theory* implies that there are more people who exist post-fusion than are accounted for by the *Non-Branching View*.

It seems to me that any attempt to read CL's case along these lines yields implications that are no less hard to swallow than those we were able to derive from the *Non-Branching View* at the start of Sect. 5. Doing so would still require us to believe that after his fall, CL's body came to be inhabited by a new person. Unlike the *Non-Branching View*, we are not required to say that the gradual return of CL's memories brought about the extinction of this new person: instead, we should think that he and CL and now share the same body and mind. All this seems very hard to accept. To avoid these implications, we might adopt a version of the *4-D Theory* analogous to the *Minimal Non-Branching View*, insisting that the retention of psychological features in a minimal form is good enough for persistence and thereby allowing that only one person exists throughout the course of events detailed in Papagno's report. However, we thereby make ourselves vulnerable to the problem that I raised in the previous section: namely, that the intuitions taken to support the *Psychological Account* look to be selectively responsive to the retention of psychological features in their maximal form.

8 Conclusion

I have shown that the *Psychological Account* has trouble dealing with cases of transient amnesia. In order to accommodate these cases, the best move open to proponents of the *Psychological Account* would appear to be the adoption of the *Minimal Non-Branching View* or some analogous version of the *4-D Theory*. However, once the *Psychological Account* takes this form, it loses what is typically thought to be its chief advantage over the *Animalist Account*. When brain transplantation is described without the addition of details that are strictly irrelevant according to the theory, so construed, the suggestion that a person survives the operation ceases to be quite so compelling.

Acknowledgements I'm very grateful to Rory Madden, Jens Johansson, Eric Olson, and three anonymous referees as *Erkenntnis* for their helpful comments on previous drafts of this paper. I also wish to express my gratitude to Derek Parfit for his generous and incisive comments on the first draft. Parfit died shortly before this paper was accepted for publication. I cannot express how much his work has meant to me as a philosopher. He is greatly missed.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bailey, A. M. (2015). Animalism. *Philosophy Compass*, 10(12), 867–883.
- Baillie, J. (1993). Recent work on personal identity. *Philosophical Books*, 34(4), 193–206.
- Bourget, D., & Chalmers, D. (2014). What do philosophers believe? *Philosophical Studies*, 170(3), 465–500.
- Brueckner, A. L. (2005). Branching in the psychological approach to personal identity. *Analysis*, 65(288), 294–301.
- Grice, H. P. (1941). Personal identity. *Mind*, 50(200), 330–350.
- Lewis, D. K. (1976). Survival and identity. In A. O. Rorty (Ed.), *The identities of persons*. Berkeley: University of California Press.
- Locke, J. ([1689] 1975). *An essay concerning human understanding*. In P. H. Nidditch (Ed.). Oxford: Clarendon Press.
- Lowe, E. J. (1998). *The possibility of metaphysics: Substance, identity, and time*. Oxford: Oxford University Press.
- McMahan, J. (2002). *The ethics of killing: Problems at the margins of life*. Oxford: Oxford University Press.
- Noonan, H. (2003). *Personal identity* (2nd ed.). Abingdon: Routledge.
- Olson, E. T. (1994). Is psychology relevant to personal identity? *Australasian Journal of Philosophy*, 72(2), 173–186.
- Olson, E. T. (1997). *The human animal: Personal identity without psychology*. Oxford: Oxford University Press.
- Olson, E. T. (2007). *What are we?*. A study in personal ontology, Oxford: Oxford University Press.
- Papagno, C. (1998). Transient retrograde amnesia associated with impaired naming of living categories. *Cortex*, 34(1), 111–121.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Perry, J. (1972). Can the self divide? *Journal of Philosophy*, 64(7), 463–488.

- Perry, J. (1975). The problem of personal identity. In J. Perry (Ed.), *Personal identity*. Berkeley: University of California Press.
- Perry, J. (1976). Review of Williams' Problems of the self. *Journal of Philosophy*, 73(13), 416–428.
- Reid, T. ([1785] 2002). *Essays on the intellectual powers of man*. In D. R. Brookes (Ed.). Pennsylvania: Pennsylvania State University Press.
- Shoemaker, S. (1963). *Self-knowledge and self-identity*. Ithaca: Cornell University Press.
- Shoemaker, S. (1984). Personal identity: A materialist's account. In S. Shoemaker & R. Swinburne (Eds.), *Personal identity*. Oxford: Blackwell.
- Shoemaker, S. (2008). Persons, animals, and identity. *Synthese*, 162(3), 313–324.
- Sider, T. (1996). All the world's a stage. *Australasian Journal of Philosophy*, 74(3), 433–453.
- Snowdon, P. F. (1995). Persons, animals and bodies. In J. L. Bermúdez, A. J. Marcel, & N. M. Eilan (Eds.), *The body and the self*. Cambridge: MIT Press.
- Swinburne, R. (1984). Personal identity: The dualist theory. In S. Shoemaker & R. Swinburne (Eds.), *Personal identity*. Oxford: Blackwell.
- Thornton, A. K. (2016). Varieties of animalism. *Philosophy. Compass*, 11(9), 515–526.
- van Inwagen, P. (1990). *Material beings*. Ithaca: Cornell University Press.
- van Inwagen, P. (1997). Materialism and the psychological-continuity account of personal identity. *Philosophical Perspectives*, 11, 305–319.