

Memory and Mineness in Personal Identity

Forthcoming in *Philosophical Psychology*

Author Dr Rebecca Roache

Address Department of Politics, International Relations, and Philosophy
Royal Holloway, University of London
Egham
Surrey
TW20 0EX

Website <http://rebeccaroache.weebly.com>

Memory and Mineness in Personal Identity

Rebecca Roache

Stanley Klein and Shaun Nichols describe the case of patient R.B., whose memories (they claim) lacked the sense of “mineness” usually conveyed by memory. Klein and Nichols take R.B.’s case to show that the sense of mineness is merely a contingent feature of memory, which they see as raising two problems for memory-based accounts of personal identity. First, they see it as potentially undermining the appeal of memory-based accounts. Second, they take it to show that the conception of quasi-memory that underpins many memory-based accounts is inadequate. I argue that Klein and Nichols’ characterisation of R.B.’s experience is implausible; as a result, the problems that they describe for memory-based accounts of personal identity do not arise.

1. Introduction

When we remember, we experience a strong sense of identity with the subject of the remembered experiences. Memory has been viewed by many philosophers since John Locke as key to explaining personal identity through time (for brevity, *personal identity*, or simply *identity* when the context permits). Locke (1731) viewed memory as constitutive of personal identity; Thomas Reid (1785) and Joseph Butler (1736) took memory to be evidence of it; and David Hume (1739/2000), whilst denying that persons persist, acknowledged that memory provides a powerful illusion of such persistence. Memory-based accounts of personal identity remain popular to this day; indeed, they embody what is perhaps the dominant contemporary view.¹

According to such accounts, a person P at time t is identical to a person P^1 at a later time t^1 if P^1 at t^1 remembers P ’s experiences at t .² Since identity is transitive, it can also arise from overlapping strands of such memory links: if P^2 at t^2 does not remember P ’s experiences at t , P^2 at t^2 and P at t are nevertheless identical if P^2 at t^2 remembers P^1 ’s experiences at t^1 , and if P^1 at t^1 remembers P ’s

¹ E.g. Shoemaker, 1970; Shoemaker and Swinburne, 1984; Perry, 1975a; Lewis, 1976; Parfit, 1984; Noonan, 1989.

² P^1 must remember P ’s experiences “from the inside,” in Shoemaker’s (1984) terminology.

experiences at t .³ Since, by definition, we remember only our own experiences, memory-based accounts often replace the notion of memory with that of *quasi-memory* in order to avoid circularity. Derek Parfit, elaborating on Sydney Shoemaker's (1970) idea, defined quasi-memory as follows: I have a quasi-memory of an experience if I seem to remember having the experience, *someone* had the experience, and "my apparent memory is causally dependent, in the right kind of way, on that past experience" (Parfit, 1984, p. 220). Those who appeal to the notion of quasi-memory in accounting for identity claim that ordinary memory is a sub-category of quasi-memory. From my quasi-memory of doing X , I cannot infer that *I* did X , but I can infer that *somebody* did X . Whether I am identical with the doer of X depends on what personal identity consists in. Defenders of memory-based accounts escape the charge of circularity by arguing that, when they say that identity consists in memory, what they mean is that it consists in quasi-memory subject to certain constraints. These constraints are specified without presupposing personal identity between quasi-rememberer and subject of quasi-remembered experiences.

Having noted the centrality of memory to discussions of personal identity, Stanley Klein and Shaun Nichols introduce patient R.B., whose memories, after he was hit by a car, were left devoid of the sense of "mineness" that usually accompanies our memories. Klein and Nichols take R.B.'s case to demonstrate that this sense of mineness is "a contingent feature of memory" (Klein and Nichols, 2012, p. 689). They see this as having two important implications for the relationship between memory and the metaphysics of personal identity. First, they see it as potentially undermining memory-based accounts of personal identity by undermining the appeal of such accounts. Second, they draw on it to raise a problem for Parfit's canonical characterisation of quasi-memory. Since the plausibility of memory-based accounts of personal identity depends upon the coherence of the concept of quasi-memory, this constitutes a further blow for such accounts. The implications of Klein and Nichols' conclusions about R.B. for memory-based accounts will be my focus in this paper.

Whilst I am primarily concerned with Klein and Nichols' claims about the relevance of R.B.'s case to personal identity, it is worth also noting that their particular conception of R.B.'s experiences has been used by Klein in many later publications to support claims about the structure of memory and its role in our conceptions of ourselves. He has used it, for example, to argue for a particular view of how scientists should approach the study of the self (Klein, 2012a), for a particular

³ This rough characterisation misrepresents aspects of some memory-based accounts. E.g. Lewis (1976) takes the psychological links that constitute identity to hold between person stages, not between persons.

role that memory plays in future-oriented mental time travel (Klein, 2012b), that episodic and semantic memory are more similar than is generally supposed (Klein, 2013a), that the function of memory is not to enable us to “re-live” the past (Klein, 2013b), and that auto-noetic awareness is not intrinsically linked to episodic memory (Klein, 2014b, 2015).

I will argue that Klein and Nichols’ conception of R.B.’s case is untenable. Specifically, R.B.’s memories are not most plausibly described as ordinary memories minus the usual sense of mineness, and the claim that mineness is a contingent feature of memory is indefensible. As a result, Klein and Nichols’ account does not raise problems for memory-based accounts of personal identity. I will proceed as follows. In the next section, I review Klein and Nichols’ interpretation of R.B.’s experiences. Since Klein and Nichols do not specify precisely *how* their conclusions about R.B. might undermine memory-based accounts, I speculate about this in section 3. In section 4, I argue that Klein and Nichols’ characterisation of R.B.’s memories is implausible and that, as a result, it poses no threat to memory-based accounts. I also dismiss the problem they raise about quasi-memory. In section 5, I consider and reject a way of salvaging Klein and Nichols’ claim that mineness is a contingent feature of memory. I conclude that R.B.’s case contains little to disturb philosophers of personal identity. Moreover, it cannot play the role that Klein requires of it in his later psychological work.

2. R.B.’s “unowned” memories

Klein and Nichols explain that long-term memory comprises two memory systems: procedural (corresponding to Gilbert Ryle’s (1945-46) “knowing how”) and declarative (“knowing that”). Declarative memory is subdivided into semantic and episodic memory. Semantic memory “contains relatively generic, context-free information about the world, such as *Grapes are edible*, $2 + 2 = 4$, and *Sacramento is the capital of California*” (Klein and Nichols, 2012, p. 679). Recalling semantic memories involves recalling information without necessarily recalling the experience of acquiring it. Episodic memory, on the other hand, involves “re-experiencing” (p. 680) one’s past: it

represents the “what, where, when” of an event. As such, it is experienced as a memory that makes explicit reference to the time and place of its acquisition. Examples of episodic memory are *I remember eating chicken for supper yesterday evening*; *I recall my meeting with Judith last Monday*. (p. 679)

Klein and Nichols note that philosophical discussions of the role of memory in personal identity generally focus on episodic memory (p. 680). Semantic memory is involved in some aspects of self-knowledge (pp. 680-81), but only episodic memory conveys a sense of the self persisting.

Against this background, Klein and Nichols introduce the case of R.B. After being hit by a car, forty-three-year-old R.B. suffered physical injuries along with cognitive and memory impairments. Most startling for Klein and Nichols is that

R.B. was able to remember particular incidents from his life accompanied by temporal, spatial, and self-referential knowledge, but he did not feel the memories he experienced belonged to him. In his words, they lacked “ownership”. This particular form of memory impairment—episodic recollection absent a sense of personal ownership, is a form of memory dissociation that, to our knowledge, has not previously been documented in the neurological literature. (p. 684)

Klein and Nichols use the terms *sense of (personal) ownership*, *sense of (numerical) (personal) identity* and *sense of mineness* interchangeably. In what follows, I favour *sense of mineness*, often abbreviated to *mineness*.

Let us examine Klein and Nichols’ reasons for claiming that R.B. experienced episodic memories minus mineness. They begin by arguing that R.B.’s unusual memories were episodic rather than semantic. They base this conclusion on the “temporal, spatial, and self-referential” (p. 685) nature of these memories and on R.B.’s claim to have been “remembering scenes, not facts” (p. 687). These features are characteristic of episodic, but not of semantic, memory.

What of the claim that R.B.’s memories lacked mineness? Klein and Nichols view themselves as taking R.B.’s talk of ownership at face value: “R.B. himself initiated use of the language of ‘ownership’[.] We simply adopt his expression” (p. 685). Consequently, they offer no analysis of this “language of ‘ownership’.” They write as if, aside from lacking mineness, R.B.’s memories were intact; that is, as if the lack of mineness was the *only* way in which R.B.’s memories were unusual. They do not state this explicitly, but they make many remarks that collectively are strongly indicative of this interpretation. For example, they tell us that R.B.’s memory, “though fitting the standard criteria for episodic recollection, was not accompanied by a sense of personal ownership” (p. 685; a similar comment appears at p. 687). They also write of episodic memories being composed of two separable components, content and mineness: “In R.B., the ‘mineness’ of

episodic recollection is separated from the content” (p. 694; a similar remark appears at p. 689), and in the abstract to the paper they claim that the sense of self conveyed by memory “derives from two components, one delivering the content of the memory and the other generating the sense of mineness” (p. 677). Similarly, they refer to “mental machinery”—defective in R.B.’s case—whose role it is to “insert a distinctive *I*-tag into memories” (p. 690; a similar comment appears at p. 689). And, on a number of other occasions, they describe R.B.’s impairment in terms of episodic memory minus mineness without mentioning any other ways in which R.B.’s memories were unusual (see, for example, p. 684, three times at p. 689, p. 690). Let us call the view according to which R.B.’s memories were unusual solely in their lack of mineness the *minus mineness view*.

3. Implications of the minus mineness view for the metaphysics of personal identity

Klein and Nichols argue that, historically, the sense of mineness has been an important motivation for taking memory to be central to personal identity: “Lockeans use this sense to build a theory of personal identity; Reid and Butler appeal to this sense as evidence for a persisting self; and Humeans maintain that the sense presents us with an illusion of a persisting self” (p. 678). They quote Reid and Butler expressing the view that mineness provides evidence of identity. Reid wrote:

How do you know—what evidence have you—that there is such a permanent self which has a claim to all the thoughts, actions, and feelings which you call yours?

To this I answer, that the proper evidence I have of all this is remembrance ... my memory testifies, not only that this was done, but that it was done by me who now remember it. If it was done by me, I must have existed at that time, and continued to exist from that time to the present. (Reid, 1785, p. 318; cited at Klein and Nichols, 2012, p. 695)

Butler wrote:

[W]hen any one reflects upon a past action of his own, he is just as certain of the person who did that action, namely himself, the person who now reflects on it, as he is certain that the action was at all done. (Butler, 1736, p. 295; cited at Klein and Nichols, 2012, pp. 695–96)

Klein and Nichols stop short of specifying whether, given their observations about R.B., mineness really *is* evidence of identity, viewing this as a task for another occasion. However, their comments are suggestive of a negative answer. They claim that “the case of R.B. indicates that this sense of identity is dissociable from episodic memory itself. The sense of identity turns out to be, *pace* Reid, a contingent feature of memory” (p. 689), and that “the sense of personal identity is really a by-product of the episodic memory system. That is just how episodic memory happens to work” (p. 696). Consequently, work needs to be done to “evaluat[e] the extent to which the sense of identity can be taken to reflect the reality of identity” (p. 696).

Klein and Nichols, then, take R.B.’s case to pose a problem for memory-based accounts by raising the possibility that mineness has led defenders of those accounts to over-estimate the plausibility of the view that memory provides evidence of personal identity. However, they move very quickly—in their concluding paragraph—from the claim that mineness is a contingent by-product of memory to the claim that it may not reflect “the reality of identity.” They do not explain exactly *how* their conception of R.B.’s unusual experiences raises problems for memory-based accounts. Since those accounts do not depend upon the claim that memory is accompanied by a sense of mineness, it is not obvious that Klein and Nichols raise any such problems. It will, then, be helpful to reflect on why they might believe that they do.

To this end, let us consider Klein and Nichols’ argument in the context of their wider research. Klein is a psychologist, one of whose key research interests is self-representation. Nichols is an experimental philosopher. Experimental philosophy draws on empirical data, particularly those pertaining to ordinary people’s intuitions, to elucidate philosophical issues. Writing with fellow experimental philosopher Joshua Knobe, Nichols tells us that

[T]he first major goal of experimental philosophy ... is to determine what leads us to have the intuitions we do about free will, moral responsibility, the afterlife. The ultimate hope is that we can use this information to help determine whether the psychological sources of the beliefs undercut the warrant for the beliefs. (Knobe and Nichols, 2008a, p. 7)

If Klein and Nichols’ argument is sound, then the “psychological source” of the belief that memory provides evidence of identity is the sense of mineness, which is merely a contingent by-product of memory. If mineness is only contingently linked to memory, then it is nomologically

possible—that is, possible were our brains to work differently⁴—for mineness *never* to accompany memory.⁵ Mineness, according to Klein and Nichols, is what historically has motivated memory-based accounts of personal identity. So, unless there are reasons for linking memory with identity that do not relate to mineness, Klein and Nichols’ argument may “undercut the warrant for the belief” that memory provides evidence of identity by raising the possibility that mineness leads us to overestimate the extent to which memory is evidence for identity. This does not directly undermine the claim—central to memory-based accounts of personal identity—that identity consists in memory links, but it does challenge defenders of those accounts to show that they have not been misled in thinking memory so integral to identity. Klein and Nichols’ conclusion that mineness is contingently linked with memory, then, presses us to re-evaluate the centrality of memory to identity. It may turn out, following such a re-evaluation, that memory deserves its central role in accounting for identity. On the other hand, it may turn out to be a red herring.

Would it really matter much if Klein and Nichols show memory to be a red herring in accounting for identity? After all, whilst Locke is famously said to have viewed identity to consist solely in memory,⁶ and whilst this view remains highly influential, there are well-known problems with it. These problems include the consequence that persons cease to exist during moments not later remembered. Some contemporary memory-based accounts escape this problem by viewing memory as evidence of identity without being constitutive of it;⁷ others take identity to be constituted by memory alongside other forms of psychological continuity, such as the persistence of personality traits, beliefs, values, and so on.⁸ The latter approach, which involves replacing a memory-based account with a broader *psychological account*, can recognise identity even in the absence of memory. Indeed, Shoemaker argued that memory is important in accounting for identity because it is evidence of the sort of causal dependence of later psychological states on earlier ones that, on his view, is required for personal persistence; yet, provided that other forms of psychological

⁴ Klein and Nichols indicate that they are concerned with nomological possibility at p. 683. They refer to it again in their conclusion, where they claim that mineness is a by-product of memory in that “[t]hat is just how episodic memory happens to work” (p. 696).

⁵ In section 5, I outline two ways to interpret the claim that mineness is a contingent feature of memory, and argue that the stronger interpretation is appropriate. I have applied the stronger interpretation here.

⁶ E.g. Flew, 1951, p. 55; Mackie, 1976, pp. 178–79; Parfit, 1984, p. 205; and Noonan, 1989, p. 9. Locke never explicitly endorsed this view, and some (e.g. Gustafsson, 2010) have argued that he held a weaker view.

⁷ E.g. Slors, 2001.

⁸ E.g. Shoemaker, 1970; Shoemaker and Swinburne, 1984; Perry, 1975a; Lewis, 1976; Parfit, 1984; Noonan 1989.

continuity can be taken to involve similar relations of causal dependence, continuity of non-memory psychological states “has as good a claim to be constitutive of a fact of personal identity” as does continuity of memory (Shoemaker, 1979, p. 326; see also Shoemaker, 1984, pp. 87–88). Even if Klein and Nichols undermine the appeal of memory-based accounts, then, psychological accounts may remain tenable.

Despite this post-Lockean dilution of memory, however, it retains a starring role in psychological accounts. Shoemaker claimed that “[r]emembering is best seen as just a special, albeit very important, case of the retention of acquired mental states” (Shoemaker, 1979, p. 326; see also Shoemaker, 1984, p. 90). More recently, Marc Slors wrote:

To be sure, memory is not thought to exhaust psychological continuity, but neither is it considered to be an optional ingredient. In fact, contemporary conceptions of psychological continuity evolved by adding various kinds of psychological connections between person-stages—connections such as those established by perpetuated beliefs, values, and character traits or by relations between intentions and actions—to John Locke’s memory criterion of identity; this criterion is held to be too tight, but correct in spirit. (Slors, 2001, pp. 186–87)

Since, as Slors observed, psychological accounts historically evolved out of memory-based accounts, the appeal of the former plausibly derives from the appeal of the latter. As a result, if we must re-evaluate the appeal of memory-based accounts, then we must re-evaluate the appeal of psychological accounts. This means that the problem that Klein and Nichols raise for memory-based accounts or personal identity cannot be evaded by subscribing instead to a wider psychological account.

Further, mineness may explain why memory enjoys a special role in psychological accounts. Shoemaker notes that “the concept of memory is itself a causal concept ... a present belief or impression counts as a memory of a past action or experience only if it stands to it in an appropriate causal relationship” (Shoemaker, 1979, p. 324; see also Shoemaker, 1984, pp. 95–96). By contrast, whilst “cross-temporal similarity” of certain other psychological states generally involves causal dependence of later states on earlier ones, the mere fact of cross-temporal similarity does not *imply* causal dependence in the way that the existence of memory-links does. As such, cross-temporal similarity of non-memory psychological states is evidence of personal identity only on the additional assumption that earlier states cause later ones.

Why does *the very concept* of memory involve causality? Why do we not, instead, view memories merely as resemblances of experiences—a conception that would make memory comparable to other sorts of psychological continuity on Shoemaker’s account? Here, too, mineness provides an answer. Recall that Klein and Nichols emphasise the role played by mineness in memory-based accounts by quoting Reid’s comment that “my memory testifies, not only that this was done, but that it was done by me who now remember it,” and Butler’s remark that, when a person remembers an action, “he is just as certain of the person who did that action, namely himself, the person who now reflects on it, as he is certain that the action was at all done.” What emerges from these remarks is not merely that one’s memories resemble one’s past experiences, but that one remembers those experiences *because* one had them, and that in remembering one is aware of this causal dependence. Arguably, then, mineness explains why memory is a causal concept and why defenders of psychological accounts view memory as a special type of psychological continuity.

Whilst psychological accounts rely less heavily on memory than memory-based accounts, then, there are reasons to believe that both are motivated by mineness. Therefore, if successful, Klein and Nichols’ argument may undermine the appeal not just of memory-based accounts, but also of psychological accounts. (Hereafter, for simplicity, I shall write only of memory-based accounts. My comments about them extend to psychological accounts.) There are, however, strong reasons to reject Klein and Nichols’ conclusions. In particular, there are strong reasons to reject the minus mineness view. Let us turn to them now.

4. Challenging the minus mineness view

The most pressing reason to reject the minus mineness view is that it is not supported by R.B.’s remarks.⁹ Consider his report of recollecting a scene from his time spent as a student at MIT:

I can picture the scene perfectly clearly ... studying with my friends in our study lounge. I can “relive” it in the sense of re-running the experience of being there. But it has the feeling of imagining, [as if] re-running

⁹ As far as I can ascertain, comments by R.B. other than those reported in Klein and Nichols, 2012, are unavailable elsewhere. Klein discusses R.B. in various other places—including Klein, 2012a; Klein, 2012b; Klein, 2013a; Klein, 2013b; Klein, 2014a; Klein, 2014b; Klein, 2015—but these contain no additional relevant information about R.B., and the reader is referred to Klein and Nichols, 2012, for a full discussion of R.B.

an experience that my parents described from their college days. It did not feel like it was something that really had been a part of my life. Intellectually I suppose I never doubted that it was a part of my life. Perhaps because there was such continuity of memories that fit a pattern that lead up to the present time. But that in itself did not help change the feeling of ownership. (Klein and Nichols, 2012, p. 686)

R.B.'s experience "has the feeling of imagining," but "intellectually"—perhaps based on its coherence with other memories—he is able to identify it as a memory of an event from his past. This report does not (as Klein and Nichols assume) convey that the experience felt just like a memory except that it lacked mineness, but that it did not feel like a memory *at all*. R.B. makes similar, albeit less explicit, comments about other memories. For example, he reports that whilst his memory of his pre-injury past was "just fine," "none of it was 'me'. It was the same sort of knowledge I might have about how my parents met or the history of the Civil War" (p. 685). These memories of his life before the injury, he says, involved "remembering scenes, not facts" (p. 686)—a detail that Klein and Nichols take to indicate that the memories described by R.B. were episodic rather than semantic. All this suggests that R.B.'s episodic memories felt to him like imaginative reconstructions based on semantic memories of other people's accounts; that, after all, is how one would likely come to represent scenes from one's parents' first meeting or the Civil War.

The comment that Klein and Nichols seem to view as most supportive of their view—since they focus exclusively on it when discussing the implications of R.B.'s case (pp. 689-90, p. 693)—is:

I could clearly recall a scene of me at the beach in New London with my family as a child. But the feeling was that the scene was not my memory. As if I was looking at a photo of someone else's vacation. (p. 686)

Even this is ambiguous, however. Klein and Nichols interpret the latter half of the second sentence as "the scene was not *my* memory"—that is, as emphasising the memory's lack of mineness—yet it also bears the interpretation, "the scene was not my *memory*." The latter interpretation, which implies that the experience felt unlike a memory, echoes R.B.'s other comments in this vein. R.B.'s accounts of his unusual memories, then, do not support the claim that these memories were unusual solely in their lack of mineness; that is, R.B.'s account does not support the minus mineness view.

Whilst R.B.'s reports provide reason to resist the minus mineness view, we should be suspicious of his comparison of his memories to imaginings. He implies that, since his memories felt

like imaginings, they felt unlike ordinary memories; yet memories and imaginings often feel similar. That we can confuse them is both a commonplace of everyday life and a phenomenon that has been experimentally reproduced and observed many times.¹⁰ The neurologist Oliver Sacks describes the experience of discovering such a confusion. On relating to his brother his recollection of a bomb falling near their house during the Second World War, Sacks learns that he could not have an episodic memory of the event:

[Michael] said, “You never saw it. You weren’t there.”

I was staggered by Michael’s words. How could he dispute a memory I would not hesitate to swear on in a court of law, and had never doubted as real? “What do you mean?” I objected. “I can see the bomb in my mind’s eye now, Pa with his pump, and Marcus and David with their buckets of water. How could I see it so clearly if I wasn’t there?”

“You never saw it,” Michael repeated. “We were both away at Braefield at the time. But David [our older brother] wrote us a letter about it. A very vivid, dramatic letter. You were enthralled by it.” Clearly, I had not only been enthralled, but must have constructed the scene in my mind, from David’s words, and then appropriated it, and taken it for a memory of my own. (Sacks, 2013)

Sacks illustrates that memories and imaginings can be phenomenologically indistinguishable. Therefore, the unusual nature of R.B.’s memories cannot be adequately explained by comparing them to imaginings. How, then, should we explain it?

One possibility is that, in claiming that they felt like imaginings, R.B. meant that his memories felt like *new* imaginings. In experiments to test the extent to which misinformation is incorporated into apparent episodic memories, a longer delay between providing subjects with misinformation and testing their memory is associated with an increased likelihood of the subject claiming to have episodic memories incorporating the misinformation (Higham, 1998). Memories of imaginative reconstructions of described events, then, are increasingly likely to be mistaken for episodic memories of those events as time passes. One reason for this is that, soon after encountering a description, one is more likely to remember having done so (Loftus, 2005, p. 362). Another reason relates to the fact that memories of imaginings are associated with more information about cognitive operations and less perceptual and contextual information than memories of external (non-imagined) events (Johnson et al., 1993, p. 4). Over time, we forget information about

¹⁰ See Loftus 2005 and the references therein.

cognitive operations associated with memories of imaginings, and we also forget perceptual and contextual information associated with memories of external events, making it easier to confuse memories of long-past external events with memories of long-past imaginings (Brédart et al., 2003, p. 9).

We should pause to note that there are at least two ways of interpreting R.B.'s terminology of the *feeling* of memories and imaginings. On the one hand, R.B. might be referring to the phenomenology of these states. On the other, he might be referring loosely to whatever features of memories and imaginings usually enable him automatically and subconsciously to identify them; these features could involve phenomenology, content, metacognitive tagging and/or something else. His claim that “[my memory] has the feeling of imagining,” on the former interpretation, means “my memory has the phenomenology of imagining;” on the latter interpretation, it means “my memory has those features (whatever they are) that I usually subconsciously associate with imagining.” Adopting the former interpretation involves ascribing to R.B. conceptual tools not generally possessed by non-philosophers. Since we have no reason to believe he is a philosopher, the latter interpretation is more natural. As such, his claims about how his memories felt are compatible with the observations of the previous paragraph; that is, with the view that memories and imaginings are at least sometimes distinguished on the basis of the sort of content-based features mentioned by Loftus, Johnson et al., and Brédart et al. His claims are also compatible, on this interpretation, with the view that neither memories nor imaginings involve any proprietary phenomenology—a view to which I wish to remain open. In adopting R.B.'s terminology of the *feeling* of memory and imagining, then, I do not refer specifically to the phenomenology of these states, but in a shorthand way to whatever means we generally use automatically and subconsciously to identify them. (Specifying these means is beyond the scope of this paper.)

Since memories of long-past external events may feel like memories of long-past imaginings, that Sacks mistook his memory of an imaginative reconstruction for an episodic memory of an external event may be attributable to the fact that many years had passed since he read his brother's description of the event. On the other hand, R.B.'s memories may—for reasons that are unclear—have felt like memories of recent imaginative reconstructions (or, simply, like imaginative reconstructions), which we would expect to feel different to old episodic memories. If they indeed had this character, his experience is comparable to cryptomnesia, in which information is forgotten

then remembered without being recognised by the subject, who mistakes it for new information.¹¹ Cryptomnesia can involve episodic memory, as demonstrated by George Harrison's unwitting plagiarism of Ronald Mack's song, "He's So Fine," in his own "My Sweet Lord." Harrison was prosecuted for infringement of copyright, but the judge commented that he did not believe that Harrison had consciously plagiarised Mack.¹² Rather, Harrison had heard, forgotten, then remembered the song as if it were an original idea of his own. If, as in cryptomnesia, R.B.'s memories felt like recent imaginings, whilst what they depicted revealed them to be episodic memories, it is understandable that he claimed not to feel ownership of them.¹³

These reflections highlight the point that it may be inappropriate to view R.B.'s memories—as Klein and Nichols do—as *lacking* some characteristic of normal experience, i.e. mineness. His account provides no more support for the view that some phenomenological aspect of normal memory was absent from his experiences than for the view that his memories possessed some additional, abnormal phenomenological feature. That R.B. himself reports a lack of ownership does not justify adopting the former view. Schizophrenics often report a lack of ownership of their actions, yet the most influential view of schizophrenic experience takes it to have a phenomenological character that is absent in healthy individuals.¹⁴ A striking illustration of this is provided by the fact that some schizophrenics, unlike healthy people, are able to tickle themselves (Blakemore et al., 2000).

Not only is the minus mineness view unsupported by R.B.'s reports, it also leads Klein and Nichols to some unattractive conclusions. They claim that whilst R.B. was capable of third-person knowledge about the subject of his remembered experiences—that is, he "had no trouble representing that R.B. had experiences on a beach in New London" (p. 689)—he lacked the ability, characteristic of episodic memory, of "representing, from the first person, 'I had these experiences'" (p. 690). This claim is contradicted by R.B.'s reports. He is quoted many times expressing first-person knowledge about his remembered past, such as when he says, "I am able to re-live [the event]. I have a feeling, a sense of being there" (p. 687). Moreover, Klein and Nichols claim that

¹¹ Brown and Murphy, 1989; Brédart et al., 2003; Sacks 2013.

¹² This conclusion is endorsed by Sacks (2013).

¹³ Comparing R.B.'s experience to cryptomnesia is intended only to illustrate that old episodic memories can be mistaken for new imaginings. We have no reason to believe that R.B. experienced cryptomnesia since we have no reason to believe that his unusual memories were forgotten then rediscovered.

¹⁴ Frith, 1992; Brüne, 2005.

R.B.'s problem with first-person representation is "natural[ly]" explained by his lacking "a special kind of conceptual self-representation" (p. 690)—the "I-concept"—which they characterise with reference to John Perry's (1977) account of indexicals. However, their view that first-person indexical thought involves conceptual self-representation is controversial. The difficulties in accounting for indexical self-reference have led some to conclude that "I" does not refer at all;¹⁵ and D.H. Mellor (1988-89) has argued that whilst "I" refers, its reference does not involve conceptual self-representation. Klein and Nichols' conclusions about first-person representation, then, are both irrelevant and controversial.

Klein and Nichols invoke their view that memory is accompanied by a separable "I"-concept to raise a problem about quasi-memory. Recall that, according to Parfit, quasi-memories must be caused "in the right kind of way." Klein and Nichols argue that "quasi-memory theorists" (p. 695)—that is, defenders of memory-based accounts who replace claims about memory with claims about quasi-memory—must state whether the "I"-concept, some other form of self-representation, or none of these "is required in order for a candidate quasi-memory to count as being 'caused in the normal way'" (pp. 694-95). Failure to do so leaves quasi-memory theorists vulnerable to the charge that some aspect of memory not captured by Parfit's account of quasi-memory is involved in personal identity, which could make quasi-memory a poor replacement for memory.¹⁶

This attack on quasi-memory may be ignored by anyone who rejects Klein and Nichols' claim that mineness is separable from memory. It is, however, puzzling that Klein and Nichols should take R.B.'s case to raise a problem for Parfit's conception of quasi-memory. Parfit claimed that our memories "come with a belief that ... they are about our own experiences," and that this "separable" belief would be absent were we to quasi-remember other people's experiences (Parfit, 1984, p. 222). His view that memories are accompanied by a separable belief that they depict one's own experiences closely resembles Klein and Nichols' view that memories are accompanied by a separable "I"-concept. As such, it is difficult to see why Klein and Nichols' account of R.B.'s experiences should raise a problem for Parfit. On the contrary, their account—were it tenable—might be viewed as providing empirical support for Parfit's claim about separability.¹⁷

¹⁵ Anscombe, 1975; Diamond and Teichman, 1979.

¹⁶ Klein and Nichols do not spell out this implication for quasi-memory theorists. I infer it from their view that this issue is important because "[t]he notion of quasi-memory is supposed to undergird a theory of personal identity" (p. 694).

¹⁷ Even so, Parfit's claim that quasi-memory can be conceived as memory minus the belief that the depicted experiences are one's own, given his view that quasi-memory can replace memory in accounting for identity, is controversial. John

Since R.B.'s case does not support the minus mineness view, defenders of memory-based accounts of personal identity may ignore the problems raised by Klein and Nichols. R.B.'s case supports only the weaker claim that his memories did not feel like memories. For convenience, let us say that they lacked a *feeling of memory*, which is a shorthand way of saying that R.B.'s memories were not identifiable as such using whatever automatic, subconscious means he usually employs to identify memories. Let us consider whether this weaker claim can support Klein and Nichols' claim that mineness is a contingent feature of memory, which is required in order to sustain their attack on memory-based accounts of personal identity.

5. R.B. minus minus mineness

Can Klein and Nichols' claim that mineness is a contingent feature of memory be salvaged without invoking the minus mineness view? A defender of Klein and Nichols may argue that, far from ruling out this claim, the view that R.B.'s memories lacked a feeling of memory *entails* it. Since R.B.'s memories did not feel like memories, and since an experience that does not feel like a memory *ipso facto* does not feel like a memory of one's own experience, R.B.'s memories lacked mineness in virtue of lacking a feeling of memory. Suppose that Klein and Nichols' defender holds a weaker version of their views about mineness, such that by "R.B.'s memories lacked mineness," she means only "R.B.'s memories were such that he did not feel that they depicted his own past experiences." She does not, then, hold the (stronger) minus mineness view. To reiterate, the minus mineness view involves the claim that R.B.'s memories were unusual *solely* in their lack of mineness; by contrast, "R.B.'s memories were such that he did not feel that they depicted his own past experiences" does not entail any particular claim about what is wrong with the memories. Is the claim that mineness is a contingent feature of memory plausible on this toned-down version of Klein and Nichols' account?

It is not. The toned-down account recognises that R.B.'s case does not support the view that memories could lack mineness without also lacking a feeling of memory. As such, it makes the plausibility of the claim that mineness is a contingent feature of memory dependent on the

McDowell, Quassim Cassam, and David Wiggins have argued that this conception of quasi-memory is parasitic on the idea of memory, therefore it cannot replace memory in the way intended by Parfit. I have argued elsewhere that, whilst this objection applies to quasi-memory as Parfit introduces it, his main conclusion—that what matters in survival is not identity—demands only on a weaker, more plausible conception of quasi-memory (Roache 2006).

plausibility of the claim that the feeling of memory is a contingent feature of memory. For Klein and Nichols' attack on memory-based accounts of personal identity to work, then, it must be plausible that the feeling of memory is a contingent feature of memory. But this latter claim is not plausible. Let me explain.

There are two possible ways to interpret the claim that the feeling of memory is a contingent feature of memory. On one interpretation, for any episodic memory, that memory could lack a feeling of memory. On a stronger interpretation, it is possible for episodic memory *never* to be linked with a feeling of memory. Let's consider each of these in turn.

The weaker interpretation need entail only that one may have a single episodic memory that feels unlike a memory whilst the rest feel like memories. This need not concern defenders of memory-based accounts, who do not require memory to be perfect, and who can accept the odd confused memory provided that many other memories are preserved. Since Klein and Nichols believe the claim that mineness is a contingent feature of memory to raise a problem for memory-based accounts, it would be inappropriate to ascribe to them the weaker interpretation.

On the other hand, the stronger interpretation is implausible. Episodic memory is not merely a capacity to represent one's past experiences; it is a faculty for non-inferential knowledge about one's past.¹⁸ In normal circumstances, remembering *X* is sufficient for (fallibly) knowing that *X* occurred. These normal circumstances involve, *inter alia*, recognising our memories as such: to derive knowledge that *X* occurred from one's memory of *X*, one must be able to distinguish remembering *X* from other attitudes towards *X*, including imagining. Were our memories *routinely* to lack a feeling of memory, we would be unable to make such distinctions without drawing inferences from external information, such as the reports of others. In this case, episodic memory would not be a faculty for non-inferential knowledge about the past. Consequently, the feeling of memory is not a contingent feature of episodic memory; neither, therefore, is mineness.

I have argued that it is conceptually impossible for memory routinely to lack a feeling of memory. Since what is conceptually impossible cannot be nomologically possible, my conclusion undermines Klein and Nichols' claim that it is nomologically possible for memory routinely to lack mineness. More generally, my conclusion also cautions us against drawing generalisations about the dispensability of features of memory (and, indeed, other mental states) from cases like R.B. That a subject who spends the first forty-three years of his life with a fully functional episodic memory

¹⁸ Goldman, 1967; Cassam, 1997; Wright, 2003.

system should then, for a relatively short period, experience episodic memories that are abnormal in some respect does not entail that it is in any way possible for episodic memory *always* to take this abnormal form. At best, it may show that the episodic memory system could work differently; but this does not entail that its outputs in that event could be properly deemed memories rather than, for example, mere representations or resemblances of past events.

6. Conclusion

Klein and Nichols take R.B.'s case to raise problems for memory-based accounts of personal identity. We have seen, however, that the problems they raise presuppose their view of R.B.'s experiences. Since there are persuasive reasons to reject their view, R.B.'s case need not worry philosophers of personal identity. The implausibility of their view of R.B. also means that R.B.'s case cannot bear the argumentative weight that Klein places on it in his later psychological work. Whilst re-evaluating his conclusions—about, for example, how scientists should study the self, and the similarities between episodic and semantic memory—is a task that lies beyond the scope of this paper, these conclusions must stand independently of Klein and Nichols' conception of R.B.'s experiences if they are to be tenable.¹⁹

References

- Anscombe, G. E. M. (1975). The First Person. In S. Guttenplan, S. (Ed.) *Mind and Language* (pp. 45-65). Oxford: Clarendon Press.
- Blakemore, S., Wolpert, D., & Frith, C. (2000). Why Can't You Tickle Yourself? *NeuroReport*, 11(11), R11-R16.
- Brédart, S., Lampinen, J. M., & Defeldre, A. (2003). Phenomenal Characteristics of Cryptomnesia. *Memory*, 11(1), 1-11.

¹⁹ I am very grateful to Neil Levy (who kindly read two separate drafts of this paper) and Piers Benn for helpful comments, and to two anonymous reviewers who also provided useful feedback.

- Brown, A. S. & Murphy, D. R. (1989). Cryptomnesia: Delineating Inadvertent Plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(3), 432-42.
- Brüne, M. (2005). "Theory of Mind" in Schizophrenia: A Review of the Literature. *Schizophrenia Bulletin*, 31(1), 21-42.
- Butler, J. (1736). Of Personal Identity. In his (1736/1906). *The Analogy of Religion*. London: J. M. Dent.
- Cassam, Q. (1997). *Self and World*. Oxford: Oxford University Press.
- Diamond, C. & Teichman, J. (Eds.). (1979). *Intention and Intentionality*. Ithaca, NY: Cornell University Press.
- Flew, A. (1951). Locke and the Problem of Personal Identity, *Philosophy* 26(96), 53–68.
- Foster, J. & Robinson, H. (1985). *Essays on Berkeley: A Tercentennial Celebration*. Oxford: Oxford University Press.
- Frith, C. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hove: Lawrence Erlbaum Associates.
- Goldman, A. I. (1967). A Causal Theory of Knowing, *The Journal of Philosophy*, 64(12), 357–72.
- Gustafsson, J. E. (2010). Did Locke Defend the Memory Continuity Criterion of Personal Identity? *Locke Studies*, 10, 113–29.
- Higham, P. (1998). Believing Details Known to Have Been Suggested, *British Journal of Psychology*, 89, 265-83.
- Hume, D. (1739/2000). *A Treatise of Human Nature*, ed. D. F. Norton and M. J. Norton. Oxford: Oxford University Press.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source Monitoring, *Psychological Bulletin*, 114(1), 3-28.
- Klein, S. (2012a). The Self and Its Brain, *Social Cognition*, 30(4), 474-518.
- Klein, S. (2012b). The Complex Act of Projecting Oneself into the Future, *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(1), 63-79.
- Klein, S. (2013a). Making the case that episodic recollection is attributable to operations occurring at retrieval rather than to content stored in a dedicated subsystem of long-term memory, *Frontiers in Behavioral Neuroscience*, 7(3), 1-14.
- Klein, S. (2013b). The temporal orientation of memory: It's time for a change of direction, *Journal of Applied Research in Memory and Cognition*, 2, 222-34.
- Klein, S. (2014a). Sameness and the self: philosophical and psychological considerations, *Frontiers in Psychology*, 5(29), 1-15.

- Klein, S. (2014b). Auto-noesis and Belief in a Personal Past: An Evolutionary Theory of Episodic Memory Indices, *The Review of Philosophy and Psychology*, 5(3), 427-47.
- Klein, S. (2015). What memory is, *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(1), 1-38.
- Klein, S. & Nichols, S. (2012). Memory and the Sense of Personal Identity, *Mind*, 121(483), 677-702.
- Knobe, J. & Nichols, S. (2008a). An Experimental Philosophy Manifesto. In J. Knobe & S. Nichols (Eds.), *Experimental Philosophy* (pp. 3-14). Cary, NC: Oxford University Press.
- Lewis, D. (1976). Survival and Identity. In A. O. Rorty (Ed.), *The Identities of Persons* (pp. 17-40). Berkeley: California University Press.
- Locke, J. (1731). *An Essay Concerning Human Understanding*. London: Edmund Parker.
- Loftus, E. (2005). Planting Misinformation in the Human Mind: A 30-Year Investigation of the Malleability of Memory, *Learning & Memory*, 12, 361-66.
- Mackie, J. L. (1976). *Problems from Locke*. Oxford: Clarendon Press.
- Mellor, D. H. (1988-89). I and Now, *Proceedings of the Aristotelian Society*, 89, 79-94.
- Noonan, H. (1989). *Personal Identity*. New York: Routledge.
- Parfit, D. (1984). *Reasons and Persons*. New York, NY: Oxford University Press.
- Perry, J. (1975). Personal Identity, Memory, and the Problem of Circularity. In his *Personal Identity* (pp. 135-55). Berkeley, CA: California University Press.
- Perry, J. (1977). Frege on Demonstratives, *Philosophical Review*, 86, 474-97.
- Reid, T. (1785). *Essays on the Intellectual Powers of Man*. Edinburgh: Bell & Robison.
- Roache, R. (2006). A defence of quasi-memory, *Philosophy*, 81, 323-55.
- Ryle, G. (1945-46). Knowing How and Knowing That: The Presidential Address, *Proceedings of the Aristotelian Society*, New Series, 46, 1-16.
- Sacks, O. (2013, February 21). Speak, Memory. *The New York Review of Books*. Retrieved February 12, 2015, from <http://www.nybooks.com/articles/archives/2013/feb/21/speak-memory/>
- Shoemaker, S. (1970). Persons and Their Pasts, *American Philosophical Quarterly*, 7(4), 269-85.
- Shoemaker, S. (1979). Identity, Properties, and Causality, *Midwest Studies in Philosophy*, 4(1), 321-42.
- Shoemaker, S. (1984). Personal identity: a materialist's account. In S. Shoemaker & R. G. Swinburne (Eds.), *Personal Identity*. New York: Blackwell.
- Slors, M. (2001). Personal Identity, Memory, and Circularity: An Alternative for Q-Memory, *Journal of Philosophy*, 98(4), 186-214.

Wright, C. (2003). Some Reflections on the Acquisition of Warrant by Inference. In S. Nuccetelli (Ed.), *New Essays on Semantic Externalism, Scepticism, and Self-Knowledge* (pp. 57-78). Cambridge, MA: MIT Press.