PRESERVATIONISM IN THE EPISTEMOLOGY OF MEMORY

By Matthew Frise

Preservationism states that memory preserves the justification of the beliefs it preserves. More precisely: if S formed a justified belief that p at t₁ and retains in memory a belief that p until t₂, then S’s belief that p is prima facie justified via memory at t₂. Preservationism is an unchallenged orthodoxy in the epistemology of memory. Advocates include Sven Bernecker, Tyler Burge, Alvin Goldman, Gilbert Harman, Michael Huemer, Matthew McGrath, and Thomas Senor. I develop three dilemmas for it, in part by drawing on research in cognitive psychology. The dilemmas centre on preservationism’s implications for certain cases involving either stored beliefs, forgotten evidence, or recollection failure. Each dilemma shows that preservationism either is false or lacks key support.

Keywords: preservationism, memory, epistemic justification, problem of stored beliefs, problem of forgotten evidence, problem of forgotten defeat.

I. INTRODUCTION

I believe I have a hand. My current experience supports my belief: it appears to me that a hand is attached to what appears to be my arm, and I seem to feel myself wiggling fingers. But most of my current beliefs are merely ‘stored’ in memory. I’m not actively endorsing their contents. And my current experience supports few of them. What epistemically justifies the merely stored beliefs that seem justified? If there is no explanation, then most beliefs are unjustified. The problem of stored beliefs is roughly the problem of supplying this explanation.¹

This problem is sometimes conflated with a second: the problem of forgotten evidence. This is roughly the problem of explaining how a belief can remain

¹ Cf. Pappas (1980: 130). Frise (forthcoming b), Goldman (1999), and Senor (1993) direct this problem at internalism about epistemic justification. Internalism states, among other things, approximately that one’s mental life alone affects one’s justification (cf. Conee and Feldman 2004; the denial of this view is externalism).
justified even when all past, direct evidence for it has been lost. I can’t recall how I learned that there are such things as fainting goats, but I believe that they exist, and justifiably. But how is my belief justified? These problems differ because we keep our evidence for some stored beliefs, and we sometimes forget evidence for beliefs while they are right before our minds, outside of storage. A received view in the literature on memory justification offers a solution to both problems:

Preservationism. If $S$ formed a justified belief that $p$ at $t_1$ and retains (in memory) a belief that $p$ until $t_2$, then $S$’s belief that $p$ is prima facie justified (via memory) at $t_2$, where $t_2$ is any time after $t_1$. According to preservationism, when $S$ keeps a belief that $p$ that had been formed with justification, the belief is prima facie justified, even if the belief is stored, even if $S$ is unable to recall why she originally formed it, and even if none of $S$’s conscious mental states justify it. In the absence of a reason not to believe $p$ (a defeater), $S$’s belief is justified overall. Memory preserves belief and justification. Or, since some preservationists drop the parenthetical ‘in memory’ and ‘via memory’ from their theory: whatever preserves belief preserves justification.

Preservationism theorizes about the justification of a subject’s belief (doxastic justification) at a later time. As a result, it theorizes about the justification a subject has for believing (propositional justification) at the later time. Any subject with a justified belief that $p$ later on has justification for, or is justified in, believing that $p$ later on.

It is important to distinguish preservationism from two theories sometimes also called ‘preservationism’:

Anti-generativism. If $S$ is prima facie justified in believing $p$ via memory at $t$, then $S$ ultimately acquired all $S$’s memory justification for believing $p$ prior to $t$ via some source(s) other than memory.

Preservationism*. (1) Preservationism is true, and (2) If $S$’s belief that $p$ is unjustified at $t_1$ and $S$ retains a belief that $p$ until $t_2$, then ceteris paribus $S$’s belief that $p$ is unjustified at $t_2$.

---

2 Many philosophers credit Harman (1986: ch.4) for introducing this problem, but see Pappas (1980: 127–9). We often lose direct evidence for our beliefs, forgetting what our exact information source was, what exactly it indicated, and in what exact context we encountered it. But we have some indirect evidence for anything we believe, since we have reason to believe that we are rational agents and therefore that most of our beliefs are justified. It would be a mistake to stipulate that a rational agent has no evidence for something he believes. But for simplicity my discussion of forgotten evidence hereafter omits the ‘direct’.


Anti-generativism states that memory does not generate justification. At most, memory preserves justification acquired from other sources, and it may not even do that. Preservationism∗ states that memory preserves the positive or negative justificatory status of a preserved belief.

The reason it is important to distinguish these theses from preservationism is this. Anti-generativism has been attacked, but usually while being called ‘preservationism’. 5 Similarly, Huemer (1999: 349–50) refers to the conjunction of preservationism and anti-generativism as ‘the preservation theory’, and attacks the conjunction simply by attacking anti-generativism. Huemer’s (1999: 351) preferred, alternative theory in fact entails preservationism. And McGrath (2007: 4) calls preservationism∗ ‘preservationism’, and attacks preservationism∗ simply by attacking (2). But McGrath (2007: 14) endorses epistemic conservatism which, I will show, entails preservationism. My point is this: several theses called ‘preservationism’ or ‘the preservation theory’ have been attacked, but preservationism has not. It has been an unchallenged orthodoxy in the epistemology of memory for over forty years. 6

In this paper, I challenge preservationism. But first I will mention additional reasons for endorsing it. Preservationism is attractive not just because it offers solutions to the problems of stored beliefs and forgotten evidence, but also because it avoids the problem of forgotten defeat. Roughly, this is the problem that a theory allegedly faces when it counts a subject’s unjustified belief from the past as justified later on, simply because the subject forgets all past reasons not to hold the belief. Preservationism evades this problem since any belief which it implies is justified later on was, in the past, justified overall. Since the belief was justified overall, the subject did not have reason to abandon it. Forgetfulness does not result in an unjustified belief becoming justified, since there is no defeater to forget. The problem of forgotten defeat does not emerge. It is controversial whether the problem of forgotten defeat is a genuine problem, but nearly all defenders of preservationism think it is. 7 So, many philosophers


6 Michaelian (2011) doubts a standard preservationist assumption, namely, that memory typically stores rather than creates beliefs. In the interests of charity, I grant the assumption. Without it, preservationism loses considerable support—it attributes little justification (its antecedent is rarely satisfied), and thus trivially does not solve the problems of stored beliefs and forgotten evidence. Senor (1995: 290) almost objects to preservationism. He thinks there are problems for an improved version of epistemic conservatism, and the improved version just is preservationism. The problems are that (a) preservationism has implications that are ‘completely foreign to the spirit of evidentialism’, and that (b) a reliabilist will deny preservationism if memory is unreliable. While (a) and (b) are correct, they are not objections. Memory is reliable, so reliabilists accept preservationism as contingently true. And (a) gives no reason to prefer evidentialism to preservationism. Indeed, evidentialists such as Conee & Feldman (2011: 304–5) offer a rival theory to preservationism, but give no reason either. They don’t object to preservationism, but rather to the claim that evidentialism lacks important resources that preservationism offers. In Sections II and IV, I show that Conee and Feldman’s theory actually undermines preservationism.

endorse preservationism because it solves two important problems without introducing a third.

And some philosophers endorse it because it follows from a plausible, more general theory of justification. For instance, epistemic conservatives like Gilbert Harman (1986: ch. 4) and McGrat (2007) hold that if S believes that \( p \), then S is prima facie justified in retaining her belief that \( p \). S's believing that \( p \) a moment ago is enough for her belief that \( p \) to be prima facie justified now. It follows that S's justifiedly believing that \( p \) a moment ago is enough for her belief that \( p \) to be prima facie justified now—and this is just preservationism.

And, for instance, many process reliabilists hold that memory is a conditionally reliable belief-sustaining process—typically, when a true belief enters into memory processing, only a true belief will come out. Reliabilists also hold that if an initially justified belief is sustained by a conditionally reliable process, the belief retains its justification. The conjunction of these two reliabilist theses entails preservationism. But even Alvin Goldman, the father of reliabilism, finds preservationism more obvious than reliabilism. Goldman (2009: 322–8) uses preservationism to argue for reliabilism. He thinks preservationism handles cases of stored belief, forgotten evidence, and forgotten defeat marvelously. For Goldman (2011: 261), a good theory of justification accommodates preservationism. Interestingly, preservationism is common ground between theories as dissimilar as reliabilism and epistemic conservatism, paradigm externalist and internalist theories of justification.

I will present three dilemmas, each showing that preservationism either is false or lacks key support. By casting doubt on this orthodoxy, I also cast doubt both on epistemic conservatism and on the conjunction of reliabilist theses that entails preservationism. We should prefer a preservative-free epistemology.

II. A NEW EVIL DEMON PROBLEM

We will see my first dilemma by comparing two cases:

CASE1. Smith formed a justified belief that \( p \) a few weeks ago. By noon today his belief is stored in memory and remains undefeated, even though Smith has forgotten all past evidence for or against \( p \), and has no new direct evidence for \( p \).

CASE2. Smith\(^*\) is Smith's mental twin at noon today. But Smith\(^*\) popped into existence just at noon. At noon he and Smith share all experiences, beliefs, and dispositions to believe, but all his apparent memories are misleading.

---

8 Cf. Goldman (1979: 13) and Senor (2010: 523). Reliabilists may of course deny that preservationism is true in any possible world in which memory is not conditionally reliable. But I am evaluating whether preservationism is true even in worlds relevantly like ours.

9 Cf. Russell (1921). Huemer (1999: 349–50) uses an example like CASE2 to attack antigenerativism, but not preservationism. My case crucially differs from Russell's and Huemer's since it centers on a subject at the moment of his creation, not five minutes after.
According to preservationism Smith’s belief that $p$ is justified at noon today. Preservationism has no implications about whether Smith*’s belief that $p$ is justified at noon today. But is it? Preservationists face a dilemma. Either (a) no, Smith*’s belief that $p$ is not justified, or (b) yes, it is. Because Smith and Smith* are mental twins at noon, it seems they should be fully alike in justification; it should be that both justifiedly believe that $p$ or that neither does. But on (a), only one believes that $p$ justifiedly. So at first glance, on (a), preservationists face the new evil demon problem—they allow some mentally identical subjects to differ in justification—and there is reason to deny any theory facing this problem.\(^\text{10}\)

And at first glance, on (b), preservationism loses key support. Its main appeal is its ability to solve economically the problems of stored beliefs and forgotten evidence. On (b) Smith*’s belief that $p$ is justified at noon today, but preservationism itself cannot explain how this is so. So preservationism is at best a partial solution to these problems. What’s more, it’s initially hard to see a way of supplementing preservationism that is not ad hoc. If there is an economical, complete solution to these problems that doesn’t entail preservationism, then there is much less reason to endorse preservationism. I will consider such a solution shortly. First, let’s see if preservationism’s trouble on option (a) is genuine.

Perhaps it seems this particular new evil demon problem won’t worry preservationists. After all, many preservationists are reliabilists, and reliabilists and other externalists seem simply unmoved by the new evil demon problem. However, many reliabilists and other externalists are moved by the new evil demon problem, tailoring their theories so that they specifically avoid it.\(^\text{11}\) And we can stipulate that Smith*’s beliefs are formed by reliable processes that in his case happened to yield mostly false beliefs. Whatever processes form Smith*’s beliefs (say, the machinations of a benevolent demon) mostly form true beliefs about the past. We can even stipulate that Smith* and Smith share a world, and that similar processes form or sustain their beliefs. Reliabilists will agree that (a) incorrectly denies that Smith* has justification, and that preservationism should not be paired with it. For many philosophers, including reliabilists, it is a desideratum of a theory that it never lets mere mass deception preclude justification. The deception must involve an epistemic defect. Smith and Smith* have different mental histories, but Smith*’s history involves no epistemic defect. It is not as if Smith* had an unjustified belief that $p$ in the past. He has no past. No past unjustifiedness could explain his lack of present justification. He is simply deceived. The new evil demon problem here is indeed worrisome.

\(^{10}\) See Cohen (1984).

One might disagree, because CASE2 may seem incoherent. Following Huemer (1999: 356n.15), we might hold that, in order for a subject to have a stored belief, that belief must at some time have been occurrent. Therefore, when a victim such as Smith* pops into existence, he lacks stored beliefs altogether. Until Smith* occurrently believes that \( p \), he has no belief that \( p \) that could be justified or unjustified.

But there is reason to deny that a stored belief must at some time have been occurrent. Information that enters memory is normally altered in at least three stages of memory processing.\(^{12}\) If memory stores nearly as many beliefs as we think it does, we should allow that the content resulting from this processing has never been occurrently endorsed. So a stored belief need not have been occurrent at any time.

Perhaps (b) is the more promising option, then: Smith* justifiedly believes that \( p \). Above I argued that on (b) preservationism loses key motivation, because it is at best an incomplete solution to the problems of stored beliefs and forgotten evidence. Preservationism itself does not imply (b), so we must supplement it with some other theory in order to have a complete solution. Perhaps we have found a non-ad hoc supplement: one of the reliabilist theories that is tailored to avoid the new evil demon problem.

Let’s grant that there is a reliabilist supplement for preservationism that explains (b). It is worth noting that, first, non-reliabilist theories like epistemic conservatism cannot accept the supplement. So non-reliabilists still have not explained (b). Secondly, preservationists will be saddled with all objections its supplement already faces, so they may not ultimately be better off. And thirdly, to supplement preservationism is to admit that, contrary to what is commonly supposed, preservationism is an incomplete solution to the problems of stored beliefs and forgotten evidence.

Now, this last point does not yet give reason to doubt preservationism. It gives reason only if there is a more parsimonious solution to these problems, i.e. a solution that needn’t appeal to a complete theory of justification. Unfortunately for preservationism, there is such a solution. Conee and Feldman (2011: 304–5) defend a view that rivals preservationism. On that view, having a disposition to recall \( p \) as known can prima facie justify a stored belief that \( p \). They claim that a subject has such a disposition toward any proposition she has learned and not forgotten (where ‘having learned’ is broadly understood). These dispositions fully account for any justification in the thorny cases of stored beliefs and forgotten evidence. Conee and Feldman’s view can count Smith and Smith\(^{13}\) each as having a justified belief that \( p \) at noon today. And Conee and Feldman’s

---

\(^{12}\) See Bernecker (2008: ch. 9), Frise (forthcoming a), and Michaelian (2011).

\(^{13}\) Dokic (2014) and Michaelian (2011) help motivate an additional rival solution that does not appeal to a complete theory of justification (for elaboration see Frise forthcoming a). Michaelian argues that memory can create justification for believing content that never entered memory and content that entered but was not believed. Dokic argues that memory often generates support for
view does not entail preservationism. On their view, if one keeps a belief that was justified in the past, but one is not disposed to recall its content as known, then contrary to preservationism the belief is not thereby prima facie justified. So Conee and Feldman offer a rival solution to the problems of stored beliefs and forgotten evidence without appealing to a complete theory of justification (see footnote 13).

I am not arguing for their view. I am arguing that it is considerably simpler than a supplemented version of preservationism that explains (b), and so their view appears preferable. Perhaps it faces some problems that preservationism can avoid. But until that is shown, on option (b) preservationism is inferior to a rival.

Let’s take stock. Preservationists face a troubling new evil demon problem on (a), a problem even many reliabilists want to avoid. Preservationism does not explain (b), and so it is an incomplete solution to the problems motivating it. A complete solution to these problems that includes preservationism is insufficiently parsimonious. (a) gives us reason to doubt preservationism, and (b) robs preservationism of key motivation. Philosophers who had some reason to endorse preservationism will have less. Philosophers who had no reason to endorse now have reason to doubt. Preservationists might not think this dilemma topples their theory. Unhappily for them, however, my next two arguments knock harder.

III. UNDERCUTTING DEFEAT OF FORGOTTEN EVIDENCE

I will introduce my second dilemma for preservationism with:

CASE3. Years ago Brown read the textbook United States History, which gave her evidence for \( p \), and as a result Brown formed a justified belief that \( p \). She has retained her belief ever since. By a moment before noon today, Brown has acquired no additional direct evidence for \( p \) and has forgotten her original evidence for \( p \) and its source. However, at noon she discovers that United States History mostly contains falsehoods. Her belief that \( p \) remains stored.14

According to preservationism Brown’s belief that \( p \) is prima facie justified a moment before noon today.

Does Brown’s noontime discovery that United States History mostly contains falsehoods defeat her preservative justification for her belief that \( p \), i.e., the justification she has for believing \( p \) in virtue of her satisfying preservationism’s

---

14 Throughout I use ‘discovers’, etc., non-factively. The reader is free to understand such terms factively if it helps persuade her of my points.
antecedent? Either (a) yes, Brown’s discovery defeats her preservative justification, or (b) no, it does not. If Brown had not forgotten why she had originally formed a belief that $p$, her discovery would defeat her justification; her discovery gives her reason to doubt that *United States History* is a good source of information, thereby undercutting any support which she realizes she acquired by reading the book. A subject’s evidence for $p$ suffers undercutting defeat when she has reason to suppose the evidence does not in her case support $p$. (Contrast this with the rebutting defeat of one’s evidence for $p$, which occurs when one has counterbalancing reason to suppose not-$p$.) Preservationism in effect allows forgotten past justifiers, whatever they may be, to continue to justify. But it is not obvious whether the undercutting defeat of forgotten evidence defeats descendent preservative justification.

I argue that either way of developing preservationism—adopting option (a) or (b)—introduces significant problems for the theory. And preservationism must be developed. It is a theory about only prima facie justification, not overall justification. Yet much of its alleged support comes from its ability to solve the problems of stored beliefs and forgotten evidence, and these problems centre on overall justification. Solving these problems requires accounting for overall justification. So, preservationists must say something about when preservative justification is and is not defeated.

Some philosophers with preservationist-friendly views think that, when forgotten evidence suffers undercutting defeat, prima facie justification tracing back to it is defeated too.15 They would defend (a); Brown is not at noon overall justified in believing $p$. But now consider:

CASE 4. Brown* is Brown’s mental twin at noon today. Brown and Brown* have identical mental histories and have always had identical environments, except in this respect: Brown*’s original justified belief that $p$ was instead a result of her acquiring evidence for $p$ by reading *American History* years ago. Brown* has acquired no additional direct evidence for $p$, and has forgotten her original evidence for $p$ and its source. At noon she discovers that *United States History* mostly contains falsehoods. Her belief that $p$ remains stored. Brown*’s discovery does not defeat her preservative justification since it traces back to evidence acquired from a different book. Preservationism implies that at noon today Brown* is still justified in believing that $p$. And yet on (a), preservationism implies that Brown is at noon not justified in believing that $p$. But Brown and Brown* have all the same evidence, form beliefs in equally reliable ways, are just as competent, share all beliefs, have nearly identical pasts, and so on. Neither has ever been massively deceived. As far as Brown can tell, discovering that *United States History* mostly contains falsehoods is of no significance to $p$ or to anything else she believes. The same is true of Brown*.

15 See e.g., Naylor (2015: 380–2) and Senor (1995: 278–9).
But according to preservationism this discovery eliminates Brown’s justification and not Brown∗’s. This is implausible.

The problem here differs from the new evil demon problem discussed in Section II. That problem was that preservationism makes it too hard for certain twins to acquire justification. The problem here is that preservationism on (a) makes it too hard for certain twins to retain justification. Brown loses her justification for believing p because she unwittingly acquires a defeater for evidence she no longer possesses, yet Brown∗ remains justified despite acquiring the same information. At noon Brown could even recognize that believing p solely on the basis of reading United States History is unjustified. She could appreciate how defeat generally works and what counts as evidence for a proposition, and she could be disposed to adjust her doxastic attitudes as soon as she sees reason to. But from her perspective she lacks reason to. It seems she should count as overall justified in believing that p.

And if she lacks this justification, which doxastic attitude toward p is justified for her? It does not seem that Brown, as a result of discovering that United States History mostly contains falsehoods, should suspend judgment in p, or disbelieve that p. After all, as far as Brown can tell her discovery is unrelated to p. It does not seem reasonable for her, in response, to change her doxastic attitude toward p.16

Part of the trouble for preservationism is that studies strongly suggest that we have fairly little memory for the sources of our beliefs, and what memory we do have is highly fallible.17 We cannot recall the exact original sources of evidence (particular people, books, websites, newspapers, etc.) for most of our beliefs. Consider what happens if we generalize preservationism’s implications on (a). Any time that we discover a source is unreliable, we lose justification for any belief for which that source provided all our previous evidence. Since we frequently cannot identify our sources, discovering that a source is unreliable will often, without our noticing, destroy all our justification for many beliefs. Since we have no reason to think that our justification is lost, we will typically keep these beliefs, unjustifiably. On (a), even a subject in excellent circumstances can lose her preservative justification too opaquely from her perspective. Paired with (a), preservationism seems false.

Some preservationists think that preservative justification remains when the relevant forgotten evidence suffers undercutting defeat; on (b), Brown’s discovery does not defeat her preservative justification for believing that p.18 Epistemic conservatives and other internalists will prefer (b) to (a). So will many externalists. For many externalists, a subject’s justification is not defeated unless the subject has some mental state (e.g., an experience or belief) that works

---

16 Feldman (2005: 282–3) and McGrath (2007: 4) make a similar point in different contexts.
17 Cf. Belli et al. (1994), Lindsay (1990), Schacter (2002: chs. 4–5), and Zaragoza and Lane (1994). Our memory for sources in childhood is especially bad.
18 See e.g., Harman (1986: 38–42) and McGrath (2007: 711.13).
against her justification. What gets defeated (the ‘defeatee’) is also always mental, namely, evidence or reasons.\textsuperscript{19} When evidence is lost or forgotten, there is no defeatee. For many externalists, Brown’s preservative justification for believing \( p \) remains since she no longer has the evidence for \( p \) that her discovery threatens.

While (b) seems plausible, it fits poorly with preservationism. Suppose S forms a justified belief that \( p \), and its justification would be undercut in circumstance C, but C has not obtained. If memory maintains the belief’s justified status, why doesn’t memory maintain its susceptibility to C? Presumably, on the best explanation of why beliefs inherit past justification, beliefs also inherit defeat conditions. So it is hard to see why, on preservationism, Brown’s discovery does not undercut her preservative justification.

But I will focus on a different worry: on (b), it is too hard for preservative justification to suffer undercutting defeat. First, I will explain why it is hard. Then, why it is too hard.

On (b), Brown’s discovery does not eliminate her preservative justification. So what could eliminate it? A plausible answer is that whatever defeats preservative justification indicates to a subject that she does not satisfy a somewhat obviously necessary condition for having overall memorial support. Preservationism states no such necessary condition. What then is necessary? We’ll start with one modest proposal:

\textbf{OMJ1.} If S is overall memorially justified in believing that \( p \), then S’s total past and present memorial support for \( p \) has been adequate.

I will leave ‘memorial support’ imprecise so as not to presuppose any particular theory about what provides it. And I focus here on propositional rather than doxastic justification for the following reason. Preservationism states only how memory justifies an attitude one has. But memory can also justify having an attitude one happens to lack. What eliminates even this justification? The answer to this question is part of the complete answer to what could eliminate Brown’s preservative justification, that is, to what could indicate to Brown that she does not satisfy a fairly obvious necessary condition for having overall memorial support. Whatever undercuts all propositional memory justification undercuts preservative justification.

\textbf{OMJ1} is perhaps boringly trivial. But it helps us identify another necessary condition. It seems that, if being in some condition C is required for overall justification, then having reasonable doubts that you are in C undercuts your justification. OMJ1 identifies one such C, and so it recommends:

\textbf{OMJ2.} If S is overall memorially justified in believing that \( p \), then it is not the case that S is overall justified in believing that S’s total past and present memorial support for \( p \) has been inadequate.

Overall memorial justification requires an absence of relatively strong rational doubt about one’s memorial support. This may also be boringly trivial. Given OMJ₂, I see one type of undercutting defeater for memorial justification in general and for preservative justification in particular. Defeaters of this type can be misleading, and can themselves be defeated. Here is an account of it:

HOD. D is a higher-order defeater for any memorial justification S has for believing p iff D indicates to S that S has had a body of evidence E such that (a) E includes S’s total past evidence for p and present memorial evidence for p, and such that (b) E does not sufficiently support p.

If D is a higher-order defeater for S’s memorial justification, then D indicates the following to S: E does not support p for S (although E might normally support p), and E contains S’s present memorial evidence (e.g., an occurrent recollection) and past evidence (testimonial, perceptual, etc.) for p. D might indicate that E is in S’s case misleading, or that in S’s case a crucial element of E originated from an unreliable source, or that S’s belief that p is essentially sustained by an unreliable process. D can, but needn’t, specify what S’s original evidence for p was. D can simply indicate that, whatever it was, it does not support p. And D can misleadingly indicate what S’s past evidence for p was. S’s having both D and insufficient reason to doubt D prevents S from satisfying OMJ₂, eliminating any support for p that S’s past evidence normally gives.

Now, it may seem that a higher-order defeater is fairly easy to acquire. Brown, in relatively normal circumstances, could learn that her former history instructor was a prankster, a revisionist, or a disgruntled employee, and so he selected a textbook for Brown’s history course that mostly contained falsehoods. Brown has forgotten what this textbook was, but it can still be discredited for her with ease.

Even so, Brown would not yet have a higher-order defeater. CASE₃ stipulates that Brown has forgotten all evidence for p. An accessible memory of her learning p from a history textbook is evidence for p, so by stipulation she lacks such a memory. Merely discovering that her instructor assigned a misleading text therefore wouldn’t jeopardize her preservative justification for p—Brown has forgotten that that text contributed to her original evidence for p. It could even be that at noon Brown would not have suspected that she learned p from a history text—if, say, p were a biographical detail about a famous scientist or mathematician. Brown’s discovery tells her little more than that there are some unreliable texts. And that is something she already knows. It would not

---

20 Even if D indicates such unreliability, some preservationists (e.g. Goldman 2011: 262) would not concede that it defeats preservative justification. If they are correct they confirm my present point: it is hard to defeat preservative justification.
defeat her support for \( p \). More generally, for S to have a higher-order defeater for her preservative justification, she must have reason to believe that her total evidence for \( p \) does not support \( p \). It is hard to gain reason to believe this if she has forgotten what any of her evidence was. Even if Brown could infer that she probably gained evidence for \( p \) from her course’s text, she still has no reason to suppose she didn’t have additional evidence for \( p \). She might have heard about \( p \) from other history courses or textbooks or instructors too. She might have seen a documentary or heard in everyday conversation that \( p \). As far as she can tell, any of this could have been part of her original evidential basis for her belief that \( p \). To discredit one seemingly likely source is not enough to undercut preservative justification. Brown has not yet acquired a higher-order defeater.

HOD identifies one type of undercutting defeater of preservative justification. I see no other types worth considering here. Perhaps HOD requires refinement. Perfecting it will not affect my present, general point: on (b), undercutting defeaters for preservative justification are hard to acquire. They are not readily available via introspection, rational intuition, memory, perception, or clearly trustworthy testimony.

Now I will argue that on (b) preservative justification is for two reasons too hard to defeat. Defeating it is so hard that preservationism leads to an implausible asymmetry:

**Grandfathering.** If at \( t_2 \) S becomes overall justified in believing that \( E \) no longer sufficiently supports \( p \) for S, then for S basing a belief that \( p \) on only \( E \) at \( t_2 \) is unjustified. But if at \( t_2 \) S becomes overall justified in believing that \( E \) no longer sufficiently supports \( p \) for S, and S based belief that \( p \) at \( t_1 \) on only \( E \) and has retained belief that \( p \) until \( t_2 \), then S’s belief at \( t_2 \) is *ceteris paribus* justified.

Other things being equal, basing belief on undercut evidence is not justified, but a belief already based on undercut evidence is justified. Why ‘*ceteris paribus*’? Because in ordinary circumstances a subject lacks reason to believe that some defeated evidence was her only past evidence for a proposition, so her preservative justification that traces back to that evidence remains undefeated. Yet, once the evidence is defeated, a subject can no longer base belief on it with justification. Preservationism grandfathers past justification. Belief based on evidence typically remains justified when the evidence is defeated, yet belief can no longer be based on that evidence with justification. The asymmetry here is troubling. It is plausible that, once evidence is defeated, all beliefs based on only it *share* one justificatory status, regardless of when they were formed.

Preservative justification is on (b) too hard to defeat for a second reason. Recall the alleged problem of forgotten defeat. That is the problem a theory faces when it allows a subject to become justified in believing a proposition if she forgets all reasons not to believe it. The *fundamental idea* motivating this problem is that merely forgetting relevant information—becoming
ignorant—should not contribute to or allow justification. On (b), Brown’s preservative justification is not defeated in CASE3, merely on account of Brown’s forgetfulness of the original basis of her belief that \( p \). Her ignorance insulates her preservative justification from potential higher-order defeaters. (Indeed, the greater the ignorance of a belief’s source, the thicker the insulation. And as I noted above, studies show that we have fairly little memory for the particular sources of particular beliefs.) So preservationism is in trouble. It conflicts with the fundamental idea. In this respect, it is no better than theories that face the problem of forgotten defeat. That preservationism avoids this problem is ultimately no asset.

You might think that the fundamental idea is simply misguided. So, it is not bad if on (b) preservationism conflicts with it. Just deny the fundamental idea. I find the fundamental idea implausible, so I appreciate the appeal of this response. Unfortunately for preservationists, however, denying the fundamental idea incurs two costs. First, part of the motivation for preservationism is that it fully solves the problems of stored beliefs and forgotten evidence without introducing the problem of forgotten defeat. But the fundamental idea alone makes it seem as if there could be a problem of forgotten defeat. If the fundamental idea is misguided, the problem dissolves. Theories that seemed to face the problem are therefore not at a disadvantage. So preservationism lacks an apparent main advantage over the rivals that seemed to face it.

Secondly, if the fundamental idea is false and so the problem of forgotten defeat dissolves, we lose reason to suppose preservationism fully solves the problem of stored beliefs. If there is no problem of forgotten defeat, a stored belief that was never justified overall might become justified overall once all defeaters for it are lost. (There is no clear reason to deny this happens, if we deny the fundamental idea.) But preservationism cannot account for this justification. The belief has never been overall justified, so the antecedent of preservationism is never satisfied. Preservationism loses support over rivals that can account for justification in such cases of forgotten defeat.

On (b), preservative justification is too hard to undercut. It’s readily rebutted, rarely eliminated. Preservationism leads to grandfathering and conflicts with the fundamental idea, which it can deny only at a high cost. So preservationism lacks key motivation. On (a), preservationism seems false. It faces a new evil demon problem centred on the retaining of justification, allowing preservative justification to be lost too opaquely from the subject’s perspective. In short, preservationism has no good account of what happens to preservative justification when forgotten justifiers are undercut. One plausible conclusion is that forgotten justifiers don’t justify. If they don’t, preservationism is false.

IV. FAILED RECALL

I will introduce preservationism’s final dilemma with:

CASE 5. Black once formed a justified belief that $p$ and has retained it without acquiring relevant defeaters. A moment before noon today, he is asked: $p$ or not-$p$? At noon today he thinks about $p$, but fails to recall it. Black has no occurrent experiential support for $p$ at noon.

Preservationism implies that, a moment before noon today, Black’s belief that $p$ has preservative justification. Since it is undefeated his belief is overall justified. But at noon today he tries to recall whether $p$ but fails and has no occurrent experience supporting $p$, including no metacognitive support for $p$—he doesn’t have a feeling of knowing the answer, $p$ is not on the tip of his tongue, he isn’t recalling other propositions that support $p$, and so on. Black at noon lacks any experience that might justify believing $p$.

Now, is Black’s belief that $p$ at noon still overall justified to the same degree it was a moment before noon? Either (a) yes, Black’s belief remains justified to the same degree, or (b) no, the justification is at least partially defeated. So far I have been silent on preservationism’s implications about the strength or degree of preservative justification. Typically, its advocates are also silent on this. Some preservationists (e.g., Owens 1999: 323) think that a belief with preservative justification is justified at least to the same degree that it initially was. Other preservationists (Huemer 1999: 354) think that the strength of preservative justification eventually degrades. (I note that, if it degrades below whatever level is required for knowledge, preservationism doesn’t explain how we have so much stored knowledge.) Fortunately, we needn’t pick any particular version of preservationism here. What matters is just that preservationism implies that a subject who retains a previously justified belief that $p$ has some degree of prima facie justification for believing $p$. As long as there is some degree, we can discuss partial defeaters of preservative justification, that is, defeaters that somewhat lower the degree of overall justification, whatever that degree is. I will focus on the degree of propositional justification that preservationism attributes. This will help us appreciate the full range of potential defeaters.

Option (a) is initially attractive in part because a single failure to recall a proposition is not clearly itself epistemically significant.\(^{22}\) We all fail at recalling well-known things at times and then manage to recall them a moment later. It seems that our degree of justification during the brief failure needn’t have dropped. If, say, Black had knowledge-level justification for believing that $p$ a moment before noon, he might still have it a moment after. Nonetheless, I will

---

\(^{22}\) See Dokic (2001: 217) and Pappas (1983: 253). For Goldman (2011: 262) preservative justification is highly resistant to defeat due to the reliability of the belief-preservation process. Goldman even hesitates to concede that misleading testimony against the reliability of memory can defeat preservative justification.
argue that preservationism has an implausible consequence on (a). In order to support my argument, I will mention two final cases. First:

**CASE 6.** At all times before noon Black∗ is mentally just like Black and is in the same type of environment. Black∗ once formed a justified belief that \( p \) and has retained it ever since, without acquiring relevant defeaters. A moment before noon today, he is asked: \( p \) or not-\( p \)? At noon today he thinks about \( p \) and immediately and vividly recalls \( p \).

On preservationism, Black∗ has preservative justification for believing \( p \) a moment prior to noon. Since he successfully recalls \( p \) and has no defeaters, his preservative justification is not at all defeated at noon. He is as justified as he was a moment before. If he had knowledge-level preservative justification then, he has it at noon.

Secondly:

**CASE 7.** At all times before noon Black∗∗ is mentally just like Black and Black∗, and is in the same type of environment that they are in. Black∗∗ once formed a justified belief that \( p \) and has retained it ever since, without acquiring relevant defeaters. A moment before noon today, he is asked: \( p \) or not-\( p \)? At noon today he thinks about \( p \) and weakly, dimly recalls \( p \).

Black∗∗’s recollective experience that \( p \) is not detailed or vivid. It lacks a certain felt persistence that Black∗’s experience had, and Black∗∗ lacks other metacognitive support for \( p \) that Black∗ had. I propose that:

\[ P_1 \]: At noon Black∗∗ is less justified than Black∗ is in believing that \( p \).

Black∗ is at least somewhat justified in believing that \( p \). (\( P_1 \)) does not state how strongly he or Black∗∗ is justified in believing that \( p \), or even that Black∗∗ is justified in believing that \( p \) at all, or even that Black∗∗ is significantly less justified in believing that \( p \) than Black∗ is. It merely states that Black∗∗ is less justified than Black∗ is. (\( P_1 \)) is plausible. If that is not yet obvious, consider this.

Our **retrieval fluency**, or our ease of recall or remembering, massively influences whether and how confidently we endorse the information we retrieve from memory.\(^{23}\) It influences what, on the basis of memory, we occurrently believe. More generally our **processing fluency**, or our ease of accessing information, massively influences our occurrent believing and level of confidence.\(^{24}\) High fluency correlates with belief and higher confidence, while low fluency correlates with lower confidence.\(^{25}\)

Our level of retrieval fluency should at least roughly correlate with our level of **rational** confidence in what we recollect. If it does not, then it seems we are systematically irrational; our memory beliefs and level of confidence generally covary with something that is independent of justification and rationality. That

\(^{23}\) See e.g., Kelley and Lindsay (1999) and Schwarz (1998).

\(^{24}\) See e.g., Alter and Oppenheimer (2009).

\(^{25}\) See Kelley and Lindsay (1999).
is an unwelcome kind of actual scepticism, one that denies that we in fact have approximately the justification or rational confidence it seems we have. This scepticism poorly supplements preservationism, a theory that is supposed to account for the justification we seem to have. It is plausible that our level of retrieval fluency roughly correlates with our level of rational confidence.26

Given the relative fluency of their retrievals, Black** is rational in having lower confidence in \( p \) than Black* is. And lower rational confidence comes with less justification than higher rational confidence does. (Pt) follows. If, say, Black* has just enough justification for believing that \( p \) at noon to allow him to know that \( p \), Black** does not.

Recall that on option (a), preservationism tells us that in CASE\(_5\) Black acquires no defeaters for \( p \) at noon, the time of his failure to recall \( p \). Black’s overall justification hasn’t decreased at noon. Neither has Black*’s. Given this, I propose preservationists should hold that, although Black and Black* differ in some regards:

\[ \text{P}1. \text{At noon Black is as justified as Black* is in believing that } p. \]

But now preservationism is in trouble. For we can reason:

\[ \text{P}2. \text{At noon Black** is less justified than Black* is in believing that } p. \]

\[ \text{C. At noon Black** is less justified than Black is in believing that } p. \] (P1, P2)

Black fails at recalling \( p \), and Black** manages to recall \( p \), yet Black has more justification for believing that \( p \). To generalize, other things being equal, if a subject has retained an initially justified belief, actually recalling the believed proposition with low fluency results in less justification than failing to recall it does! This is implausible. The challenge preservationism faces, given (a), is that of either avoiding (C) or explaining away (C)’s apparent implausibility.

One might resist (P2)—thereby avoiding (C)—by proposing that Black* has more justification for believing \( p \) than Black does, because Black*’s recollective experience boosts his justification. His degree of justification goes up at noon. This proposal is appealing. But many preservationists (e.g., Huemer 1999: 353–4), especially the anti-generativists (e.g., Goldman 2009: 323, 2011: 259–60, Naylor 2015, and Senor 2007) will deny it. Preservationists typically deny that a memory experience can increase one’s justification for \( p \) beyond the preserved level. More importantly, if preservationism is true the proposal has an implausible consequence. Suppose that Black*’s preservative justification a moment before noon is just beneath the strength required for knowledge. By recalling \( p \) at noon his justification boosts to knowledge-level. So he comes to

know that \( p \) merely by recalling it. This is not a case in which he comes to know \( p \) by forming a belief that \( p \) on the basis of recollection, nor by defeating a defeater for \( p \) via recollection, nor by recalling \( p \) with unusual vivacity, nor by recalling previously unappreciated support for \( p \), nor by recalling \( p \) after having forgotten \( p \). His mere recollection that \( p \) promotes his justified, true, non-Gettiered belief that \( p \) into knowledge. But it is implausible that belief can be promoted into knowledge in this way.

One might try to explain away (C)'s apparent implausibility as follows. Difficult recall is at least a partial defeater of one's justification. A moment before noon, and at noon, Black** is prima facie justified in believing that \( p \) to the same degree. But at least some of that justification is defeated by his disfluent recalling experience, so his overall justification decreases at noon. When, like Black, one fails at recalling, one has no defeater and so the strength of one's overall justification remains. At noon, Black** and Black differ only in their overall justification, not in their prima facie justification. Defeat explains the apparent implausibility of (C).

Here are two problems for this reply. First, suppose weak recall is a defeater. Maybe it defeats because it gives one reason to doubt that one has a good memory, or to doubt that one recalls \( p \) as well as one would if \( p \) were true. But total recall failure would also introduce these doubts. Total recollective failure is defeater too, then. But (a) denies precisely this! The reply contravenes (a).

Secondly, it seems weak recall that \( p \) cannot be a defeater, since it seems to be (weak) evidence for \( p \). After all, weak recall displays \( p \)’s truth, even if the display could be better. But if weak recall that \( p \) is evidence for \( p \), it does not defeat one’s prima facie justification for believing that \( p \). Additional evidence for \( p \)—evidence that doesn’t bear on any other evidence one has for \( p \) and that introduces no relevant doubt—does not reduce one’s justification for believing \( p \). Of course, it might seem that weak recall that \( p \) does introduce some relevant doubt. But, again, if weak recall does, a total failure to recall should too, and we must then reject (a).

On option (a), preservationism incorrectly leads to (C). Denying preservationism avoids (C). Without preservationism, it is unclear that Black and Black* are equally justified in believing \( p \) at noon; (P2) lacks support.

Similarly, on option (b), preservationism avoids (C). Option (b) states that Black’s failure to recall \( p \) at least partially defeats his justification.\(^{27}\) So there is insufficient support for (P2). But what could Black’s partial defeater be? We need a plausible explanation. A partial defeater for \( p \) somewhat lowers one’s overall justification for believing \( p \). A partial rebutting defeater does this by providing some evidence that \( p \) is false. But Black’s failure to recall \( p \) provides

\(^{27}\) Cf. Martin (2001: 262), who suggests that temporary retrieval failure is incompatible with memorial knowledge.
him with no evidence that \( p \) is false. He has not acquired a partial rebutting defeater.

A partial *undercutting* defeater reduces one’s overall justification by, roughly, weakening how well one satisfies a condition necessary for having overall justification. It is hard to say exactly how this weakening occurs, but our intuitive grasp of this phenomenon will suffice here. In Section III, we considered two theses stating requirements for overall memorial justification (OMJ_1 and OMJ_2) and a type of undercutting defeater (higher order). OMJ_1 requires that the subject’s total past and present memorial support has been adequate. OMJ_2 requires that the subject is not justified in doubting the adequacy of this support. If Black’s retrieval failure reduced his overall justification in the light of these requirements, it would weaken how well he satisfies one of them. But Black at noon satisfies them just as well as he did a moment before. His total past and present memorial support for \( p \) is unchanged, and he is no closer to having justification for believing that it was poor. So Black’s retrieval failure offers no partial higher-order defeater.

A possibly tempting explanation of Black’s partial defeat is that his failure to recall \( p \) indicates to him that he doesn’t believe that \( p \). And clearly, if one lacks a belief that \( p \), then one lacks a *justified* belief that \( p \). His failure to recall, then, indicates that he lacks a justified belief that \( p \), thereby at least partially defeating his justification for believing that \( p \).

This explanation is inadequate. We will identify memorial undercutting defeaters only by identifying the necessary conditions for overall memorial justification that they prevent subjects from satisfying. Recall that it was OMJ_1 that made OMJ_2 plausible. OMJ_1 states a necessary condition for overall memory justification, and OMJ_2 states that one cannot reasonably doubt that one meets that condition, if one is overall memorially justified. In order for doubts about whether one believes \( p \) to undercut memorial justification for \( p \), something like the following must be the case:

OMJ_3. If \( S \) is overall memorially justified in believing that \( p \), then \( S \) believes that \( p \).

This would make it plausible that:

OMJ_4. If \( S \) is overall memorially justified in believing that \( p \), then it is not the case that \( S \) is overall justified in believing that \( S \) does not believe that \( p \).

The tempting explanation would then suggest that Black’s retrieval failure weakens how well he satisfies OMJ_4. That is why his retrieval failure is a partial undercutting defeater.

But OMJ_3 is false. It is not the case that overall memorial justification for believing that \( p \) requires belief that \( p \). Memory can retain information that we neglect to endorse, overall justifying us in having beliefs we unfortunately lack. We aren’t epistemically perfect! And if OMJ_3 is false, we have no support for OMJ_4.
Further, OMJ₄ is independently implausible. An overall rational doubt about whether one believes that \( p \) is not itself reason not to believe that \( p \). Sometimes when we discover that we don’t believe something, the justified response is to form belief. Finally, even if OMJ₄ is correct, it is not clear that Black has reason to doubt he believes that \( p \). Mere momentary retrieval failure does not indicate a lack of belief. It just indicates a lack of occurrent belief. It can take time to recall or to activate a stored belief. The partial undercutting defeat of Black’s justification for believing that \( p \) is not due to any reasonable doubts about whether he believes that \( p \).

A more promising explanation of Black’s partial defeat begins with:

OMJ₅. If S is overall memorially justified in believing that \( p \), then S is disposed to recall \( p \).

And this supports:

OMJ₆. If S is overall memorially justified in believing that \( p \), then it is not the case that S is overall justified in believing that S is not disposed to recall \( p \).

The explanation continues: Black’s failure to recall \( p \) indicates to Black that he lacks not a belief that \( p \), but a disposition to recall \( p \). If he had such a disposition, he would not in ordinary circumstances have failed at recalling \( p \) upon trying. Perhaps his retrieval failure weakens his satisfaction of OMJ₅. Even if it does not, he has still gained some reason to doubt that he is disposed to recall \( p \). This weakens Black’s satisfaction of OMJ₆. And that explains the partial defeat of his preservative justification for believing \( p \).

This explanation has merit. OMJ₆ is plausible. Support for it crucially traces back to OMJ₅. It’s hard to see how OMJ₆ could be true if OMJ₅ were false. It’s hard to see how doubting that you are in condition C undercuts your justification, unless being in C is required for overall justification. OMJ₅ is plausible. But preservationism appeals to it at its own peril. Preservationism is supposed to offer an economical solution to the problems of stored beliefs and forgotten evidence: in the thorny cases of stored belief and forgotten evidence, the subject has preservative justification. On OMJ₅, overall memorial justification requires a recollective disposition. So, in the thorny cases of stored belief and forgotten evidence, the subject has a recollective disposition.

But in Section II, I noted that Conee & Feldman (2011: 304–5) defend a solution to the problems of stored beliefs and forgotten evidence that rivals preservationism: having a disposition to recall \( p \) is sufficient for having prima facie justification for believing \( p \). OMJ₅ requires the preservationist solution to these problems to posit precisely what Conee and Feldman posit: in all relevant cases, the subject has a suitable recollective disposition. Conee and Feldman’s solution posits nothing more. Yet preservationism’s solution posits quite a bit more. It posits that the subject in each case formed a justified belief that \( p \) in the past and has believed \( p \) ever since, and that this is a means of retaining
justification; reliabilist preservationists will also posit that certain reliability conditions obtain. It is therefore unclear why we should not, in the interests of simplicity, just adopt Conee and Feldman’s view.

Again, this is not an argument for their view. It is an observation that, on the promising account of why failed recall introduces a partial defeater (the account that appeals to OMJ\textsuperscript{5} and OMJ\textsuperscript{6}), the preservationist solution to other problems importantly overlaps with a much more parsimonious rival. The preservationist solution is inferior; on option (b), preservationism lacks key support. It solves only problems that can be better solved. And on option (a) preservationism leads to (C). So we have reason to avoid preservationism.

V. CONCLUSION

I have presented three dilemmas for preservationism, an orthodox and unchallenged view in the epistemology of memory. Here are the results. Either preservationism faces a version of the new evil demon problem or it does not fully solve the problems that motivate it. Either preservationism faces another version of the new evil demon problem or it faces a problem it was supposed to avoid. And either preservationism lets failed recall result in more justification than successful recall does or there is a simpler rival solution to the problems that motivate it. Each dilemma shows that preservationism either is false or lacks key support. Together they give us reason to deny preservationism, or at least reason to explore other options.\textsuperscript{28}

REFERENCES


\textsuperscript{28} For helpful conversation and comments on drafts, I thank Matthew Baddorf, Sven Bernecker, Earl Conee, Richard Feldman, David Henderson, Daniel Immerman, Jonathan Matthen, Kevin McCain, Matthew McGrath, Kourken Michaelian, Andrew Moon, Thomas Senor, Joshua Smith, two anonymous referees, and audiences at the 2014 APA Central Division meeting, the 2014 Canadian Philosophical Association Annual Congress, Northwestern University, and Washington and Jefferson College. I revised this paper while supported by grants from the John Templeton Foundation and the Templeton Religion Trust. The views expressed in this paper do not necessarily reflect the views of either organization.
—— (forthcoming a) ‘Eliminating the Problem of Stored Beliefs’, *American Philosophical Quarterly*.
—— (forthcoming b) ‘Internalism and the Problem of Stored Beliefs’, *Ekeneninis*.

Baylor University, USA