

# 11 Forgetting

*Matthew Frise*

## 1 Introduction

This chapter is about how you will forget this chapter. It's about what forgetting anything is. We forget often, and psychologists research why. But neither they nor philosophers have tried much to unearth the *nature* of forgetting. The little shoveling in the area has turned topsoil only.

This is odd, since forgetting is philosophically important. It looks essentially connected to, and as important as, *remembering*—a popular topic in the philosophy of memory. Fifty years ago, the *Philosophical Review* published C.B. Martin and Max Deutscher's landmark article "Remembering," which proposed an analysis of remembering. The literature on their paper and topic is now sizable. Commonsense puts remembering on one pole of a spectrum, and forgetting on the opposite pole. If commonsense is correct here, then a handle on forgetting looks crucial to a handle on remembering. We fully grip the poles together or not at all. Yet "forget" appears only three times in Martin and Deutscher's paper.

Forgetting also matters in epistemology. The fact that ordinary humans forget introduces an evaluative challenge. We forget evidence, as well as counterevidence or "defeaters." This forgetting can be a kind of losing. But losing evidence or counterevidence can affect what it is reasonable to believe. Since epistemology is concerned with theories of reasonable belief, understanding what counts as forgetting will be crucial to evaluating these theories.<sup>1</sup> Without this understanding, it will be hard to tell whether an apparent case involving forgetting is a counterexample to a theory.

What's more, the pattern of ordinary human forgetting looks epistemically significant. Kourken Michaelian (2011: 400) argues that, contrary to a standard view in epistemology, normal human forgetting "approximates a virtue" rather than a vice. It is a mean between too much forgetting and too much remembering. If Michaelian or the standard view is right, then forgetting underlies a normative epistemic status.

Forgetting has no small significance in metaphysics either. Consider personal identity, that is, how a person persists or is identical to something at another time. One take on this is that personal identity consists in having

certain memory connections.<sup>2</sup> And forgetting looks like a wearing down of memory connections. Too much forgetting might destroy personal identity, on some theories. Testing these theories requires clarity on when subjects do and don't forget.

There are other areas in philosophy where forgetting matters, but turn now to psychology. Most often, when psychologists talk of forgetting, they talk about its causes, not its nature. Research focuses on the trauma, disorders, and commonplace memory processing that makes forgetting more likely, not on what forgetting is. When psychologists do say what forgetting is, their accounts are often incomplete and superficially incompatible with one another. Further reflection on forgetting will help reveal whether they really disagree with each other, whether they have ultimately distinct phenomena in mind, and whether some accounts are simply mistaken.

For reasons like these, I develop and defend a theory of forgetting in this chapter. First, I survey some viable choice-points for theorizing about forgetting, points where forgetting is somewhat murky. Then I state the theories of forgetting that flit in the literature. I introduce two tests for a theory of forgetting, and show that the best theories from the literature fail both tests. Finally, I introduce a new theory of forgetting, the *learning, access failure, dispositional (LEAD) theory*. I argue that the LEAD Theory of Forgetting is the lead theory of forgetting. The causes and normative status of forgetting don't get further airtime.

## 2 Dimensions of Forgetting

The dimensions of forgetting I will cover are notable, but not exhaustive. Keeping them in mind helps us articulate a more exact theory of forgetting. It turns out that what we say about some dimensions of forgetting limits what we can sensibly say about others. As we'll see, that introduces problems for some theories of forgetting. Some dimensions to consider are:

### *Ontological Category*

At a most general level, what kind of thing is forgetting? Maybe forgetting is a *mental state*—a state of mind. It is something that a subject is in at a time, like the state of being in pain, or feeling bliss, or desiring the truth. Another option is that forgetting is a *mental process*. It is something mental that unfolds over time, like the process of calculating a sum, or forgiving someone, or coming up with a joke. Some ordinary forgetting-attributions mask these options. When I tell you merely that “Maria is forgetting her password,” I may be saying something about her current mindset, or instead about how she is changing. It could be that these options aren't exclusive. Some forgetting may be a state, and some may be a process. Within any of its dimensions, forgetting may show some diversity.

*Content*

Set aside what forgetting itself is. What kind of thing is it that we forget? The content of forgetting could be *propositional*. To forget is to forget that  $p$ . Your friend forgets that Plato taught Aristotle, and you are forgetting that Tuesday is your mother's birthday. Or the content of forgetting could be *objectual*. To forget is to forget  $o$ , where  $o$  is an event, experience, or other object; I have forgotten my first birthday party, and I am forgetting my childhood neighbor. Or, the content of forgetting could be *procedural*. To forget is roughly to forget how to  $\phi$ , where  $\phi$ -ing is the performing of an action. The unfortunate magician forgets how to perform his magic trick, and the hungry fisher is forgetting how to fillet a fish. Different memory systems would naturally correspond to forgetting with different content types. Propositional forgetting would likely have a special relation to semantic memory, that is, memory for propositions; objectual forgetting would likely have a special relation to episodic memory, our memory for experienced events; and procedural forgetting would have a special relation to procedural memory, our memory for how to do things and for skills. If some forgetting has one type of content and other forgetting has another type, more than one memory system is likely responsible for human forgetting.<sup>3</sup>

*Relation to Content*

Forgetting appears to be a mental relation between a subject and a content. But a subject can bear many mental relations to a content. Suppose the content is propositional. A subject forgets that  $p$  is related to  $p$  in a certain way. What distinguishes that relation from others the subject bears to  $p$ ? Instead of forgetting that  $p$  a subject may, say, desire that  $p$  or imagine that  $p$ . Something characterizes forgetting, setting it apart from these other relations. Perhaps *loss* characterizes it. To forget is, in some sense, to have something no longer. What exactly is lost is a further question. Michaelian (2011: 403) says a "record" or mental representation is lost, and others (e.g., McGrath, 2007, p. 1; Naylor, 2015, p. 377) suggest a belief is lost. Another characterization of the forgetting relation is that of *failing to retrieve* (Arango-Muñoz, 2013). Forgetting is not a matter of losing something, but rather of failing to access it.

Some relation characterizes forgetting. Whatever it is, another question concerns its minimal duration: how long must the relation endure to count as forgetting rather than as something else? The minimum may be short, such that even *temporary* loss or retrieval failure could count as forgetting. Alternatively, the relation may have to be effectively *permanent*. If something is forgotten it is locked away in, or eliminated from, the mind. It would have to be relearned altogether in order to be mentally accessed. Some philosophers (e.g., Pappas, 1987, p. 153) think there is more than one kind of forgetting, and they individuate the kinds by minimal relation duration.

When the relation is permanent, it is one circumstance in which we hear the subject described as “completely” forgetting.

### *Scale*

The last dimension I’ll cover is the magnitude-range of forgetting. Can forgetting vary in intensity? The simplest answer is that it cannot. Forgetting is *binary*. You either forget or you do not. There’s no middle ground, no partial forgetting, no milder or sharper forgetting. Then again, it could be that forgetting is *gradable*. It admits of degrees or occurs to varying extents. As Daniel Schacter (2001: 33) puts it, there is “incomplete rather than total forgetting that leaves in its wake scattered shards of experience. Vague impressions of familiarity, general knowledge of what happened, or fragmentary details of experience.” When forgetting intensity is high enough, it is another circumstance in which we hear the subject described as “completely” forgetting.

### 3 Theories of Forgetting

AuQ3 A theory of forgetting fills in some dimensions of forgetting. A more complete theory completes more dimensions. The peeps in the literature about forgetting’s nature express stances on some dimensions I’ve covered. Some philosophers and psychologists state in passing simple views about forgetting. It may be most charitable to understand these views as incomplete theories, stating just a necessary or sufficient condition for forgetting. Even so, I will show that the simple views face immediate problems. Then I will go over more developed theories of forgetting.

One simple view is that you are forgetting whatever it is you aren’t recalling. That is:

*Simple Access Failure Theory.* If  $S$  does not recall  $x$ , then  $S$  forgets  $x$ .<sup>4</sup>

The Simple Access Failure Theory classifies forgetting as a state. You are in the state of forgetting anything that you aren’t recalling. But this is not quite right. At any given time, we aren’t recalling most things that we remember quite well. This is in part because we are making no attempt to recall any of this. But it would be incorrect to count us at any given time as forgetting most of what we remember! A moment ago, for example, you weren’t recalling what you did last night. You weren’t trying recall it. Still, we shouldn’t conclude that you were forgetting it.

There is a simple fix. You aren’t forgetting just anything you aren’t recalling. Rather, you are forgetting only what you aren’t recalling *in a certain situation*. What’s the situation? It could be free recall, a period where you try to recall select items (in no particular order, using any strategy). Given this fix, we have:

*Less Simple Access Failure Theory.* If  $S$  does not recall  $x$  during free recall targeting  $x$ , then  $S$  forgets  $x$ .<sup>5</sup>

This theory makes an improvement. It doesn't imply that at any given time you are forgetting most of what you remember. This is because you aren't in free recall at just any given time, and because when you are in free recall you aren't trying to recall most of what you remember. Unfortunately, a problem remains. Suppose that right now, during free recall, I am trying to recall what it is you did last night. If I am unsuccessful, we should not conclude that I am forgetting what you did last night. This is because at no point have I had any idea what you did last night. I never *learned* what you did. The Less Simple Access Failure Theory, however, oddly implies that I am forgetting what you did last night. What's more, the Less Simple Access Failure Theory is far from complete, giving just a narrow sufficient condition for forgetting. It doesn't tell us anything about the forgetting that occurs outside of free recall. And, as we'll see, forgetting outside of free recall is common.

If, during free recall, you do not recollect something that you learned and that you're trying to recall, your memory in some sense fails. The final simple view is that forgetting is just any sort of memory failure:

*Simple Memory Failure Theory.* If S's memory fails (with respect to  $x$ ), then S forgets ( $x$ ).<sup>6</sup>

On the Simple Memory Failure Theory, forgetting looks once again like a state. It at least sometimes is a state of having memory failure with respect to something. Compared to the previous theory, this theory makes a step forward. On this theory, memory failures other than those occurring during free recall count as forgetting. However, this theory also makes a step backward. It avoids being too narrow about forgetting simply by being too broad. Memory can fail in a number of ways that don't involve forgetting. It can fail by leaving us open to confabulating, adding incorrect details into what is recollected. It can fail by leading to various biases, or by continually activating unwelcome memories.<sup>7</sup> These memory failures are not automatically instances of forgetting. Of course, we could refine the idea behind the Simple Memory Failure Theory. Not just any memory failure is forgetting. Rather, forgetting is a special kind of memory failure. Just what kind is it? As we noted with the Less Simple Access Failure Theory, our focus should not be limited to failures during free recall. Also, the relevant memory failure should exclude failures to access information one never learned. I don't think there is a good, simple way to meet all the criteria here, so I won't say more about the Simple Memory Failure Theory.

The simple theories proved to be too simple. I will introduce a pair of more sophisticated theories, and then test them together. The simple theories all identify forgetting as something that can be a state. The first sophisticated theory departs notably from this, however, focusing on forgetting as a process. Timothy Williamson (2000: 34) takes this to an extreme, saying, "Not all factive attitudes constitute states; forgetting is a process." A factive

attitude is an attitude that guarantees the truth of its content. *Seeing* that  $p$  is often thought to be factive. If you see that Brian has a mullet, Brian indeed has a mullet. Williamson thinks forgetting is factive—you can forget only truths, not falsehoods. For Williamson, it is critical that all forgetting is a process, never a state. He (2000: 34) thinks “knowing is the most general factive stative attitude,” that is, anyone in a factive attitude state that  $p$  *knows* that  $p$ . (This, he thinks, is evidence that the concept of knowledge plays a central role in our thinking.) But at first glance, it doesn’t look like your forgetting that  $p$  guarantees your knowing that  $p$ . On the contrary, forgetting seems to explain a loss of knowledge. So, if forgetting is factive, it had better turn out that forgetting isn’t a state. For if it is a factive state, knowledge isn’t the most general factive stative attitude, and we lose evidence that the concept of knowledge is so special.

Williamson suggests, then:

*Process Theory.*  $S$  forgets that  $p$  only if and because  $S$  has a factive process attitude toward  $p$ .<sup>8</sup>

As formulated, the Process Theory says just this: forgetting *guarantees* there is some process involving an attitude toward a true proposition, and forgetting is explained by this process. It doesn’t say that forgetting *is* the process attitude, nor that forgetting cannot be a state. So, the Process Theory is weaker than Williamson’s view. Also, friends of the Process Theory may offer a different account for nonpropositional (objectual and procedural) forgetting. None of this will matter for my evaluations of the Process Theory in the next section. At any rate, if the Process Theory is incorrect then Williamson’s view is incorrect, since Williamson’s view entails the Process Theory. And if the Process Theory is incorrect, there is less evidence that knowledge plays the special role that Williamson says it plays. I’ll return to this point in Section 6.

The other sophisticated, and most promising, theory of forgetting currently in the literature, understands propositional and objectual forgetting as a loss of information. The information is lost from long-term memory, which includes short-term memory but not working memory. This gives us:

*Information Loss Theory.*  $S$  forgets  $x$  only if and because  $S$  loses a record of  $x$  from  $S$ ’s long-term memory.<sup>9</sup>

Understand *loses* here broadly. Sometimes loss is the *elimination* of a record from long-term memory—the information fades or gets deleted—so the loss is all else being equal permanent. But more often it is an *inaccessibility* of information despite appropriate stimuli. That is, even though the subject is presented with relevant retrieval cues, the record remains at least temporarily inaccessible.<sup>10</sup> On the Information Loss Theory, forgetting looks like a state, though it could also be a process involving this

state. The state is that of having dropped something from long-term memory, or of having something in long-term memory that cannot be retrieved. The Information Loss Theory skirts the snares that tripped the Less Simple Access Failure Theory—it theorizes about all propositional and objectual forgetting, and it makes forgetting require learning (because anything lost from long term memory was learned). And the Information Loss Theory does not conflate just any memory failure with forgetting. So it beats, or perhaps simply develops, the Simple Memory Failure Theory. In fact it's looking shipshape.

#### 4 The Metacognition Test and Prospection Test

The Process Theory and Information Loss Theory fashion up forgetting in quite different ways. Still, they have similar flaws. This section offers two tests for an adequate theory of forgetting. Each checks whether a theory can accommodate certain intuitive data about forgetting. It is bad to fail either test. Failure generates an argument against the theory. I claim that the Process Theory and Information Loss Theory fail each test, and so they are in trouble. They may face separate challenges as well, but for simplicity I consider just those they share.

First, there is the metacognition test. Not only do we cognitively process information, but we also monitor and control this processing. This is metacognition. One thing we monitor, typically unconsciously, is our own production of information (from, for example, memory). This results in a metacognitive feeling, which gives phenomenological feedback about our cognitive processing. Any of a broad range of metacognitive feelings may arise, depending on whether information is indeed produced, on the details of any produced information, and on how it is produced (e.g., quickly, slowly). One common metacognitive feeling is the feeling of forgetting. It can be a frustratingly un-specific feeling. As you are about to leave your home for the airport, you feel you are forgetting . . . *something*. You can't tell what, despite trying. You are failing to retrieve some target information. This metacognitive feeling is not only common, but it is also generally accurate evidence of forgetting.<sup>11</sup> The feeling may sometimes arise when there is no relevant forgetting, but that is the exception. A theory of forgetting should not conflict with any of this. What's more, it should make sense of cases in which, intuitively, the feeling of forgetting is accurate; the theory should imply that they are indeed cases of forgetting. These are the standards of the metacognition test.

The Process Theory does not meet these standards. It could be that processes altogether lack phenomenology, but that is not why the Process Theory fails the metacognition test. For even if no process is phenomenological, it could be that the process of forgetting involves a succession of states, and that one of these can be a phenomenological state like the feeling of forgetting. The Process Theory could then explain any accurate feeling of forgetting as occurring in connection with a state in the process of forgetting.

No, the Process Theory fails the metacognition test elsewhere. Suppose forgetting can indeed be a process. The process is not in each case ongoing. It ends. And sometimes, afterward, a relevant feeling of forgetting arises. And sometimes, this feeling is accurate—there is in fact forgetting after the process of forgetting has finished. Take a simple example. I study a list of obscure words having to do with astrophysics. Much later, you ask me to reproduce the list. I experience the feeling of forgetting. I am in fact forgetting what was on the list. So, my feeling is accurate. But the Process Theory does not make sense of this. The process of forgetting what was on the list began and ended some while ago. Once the process ends, that forgetting ends. The feeling of forgetting is a feeling having to do at least partly with how things currently stand; it is a feeling of current forgetting, not purely past. So on the Process Theory, any feeling of forgetting after the process of forgetting ends is inaccurate. According to it, I am not forgetting when having a feeling of forgetting upon trying and failing to recollect the astrophysics terms that I had learned. But that is wrong.

Despite its merits the Information Loss Theory stumbles on the metacognition test too. The Information Loss Theory explains forgetting in terms of lost target information. A feeling of forgetting will be accurate only if the target information either has been eliminated or is inaccessible when the feeling occurs. In many cases the target information really is lost, but not so in many others. These other cases are those where the target information has not been eliminated and remains *accessible* in the face of relevant cues, but the information simply is not *accessed*. Suppose you supply helpful retrieval cues, reminding me that the words on the studied list were on astrophysics, that some words start with this or that letter, and that others rhyme with these or those words. On this occasion I am unsuccessful in recollecting what was on the list, but I could succeed on another occasion. Yet I am indeed forgetting, and there is no inaccessibility or elimination of information here that explains this.

The problem is not just in the eyes of pedants; the difference between a failure to access and inaccessibility is not trivial.<sup>12</sup> I fail to access my destination if something halts me en route to it. But that is not to say my destination is inaccessible. If it were inaccessible, all my routes to it would be blocked, or I would not have the power to access it. Inaccessibility is a strong dispositional relation. Forgetting can result from a one-off disconnect, a nondispositional and weaker relation. Sometimes this weaker relation explains why a feeling of forgetting is accurate, even when the target information is accessible. That is why the Information Loss Theory doesn't pass the metacognition test.

The second test, the prospection test, focuses on certain cases of prospective memory failure. Roughly, prospective memory is memory for intended future action. Some prospective memory failure counts as forgetting.<sup>13</sup> Suppose I decide I will call you at noon, but when noon comes I am so preoccupied that it never occurs to me to pick up my phone. I needn't have a



feeling of forgetting at noon, nor any other relevant metacognitive feeling. I simply fail to perform my intended act when I am in the appropriate circumstance. The failure to recollect may result from a shortage of suitable retrieval cues—I didn't set an alarm as a reminder—or from suitable retrieval cues being available yet ineffective—my alarm went off, but I didn't recall what it was for. Regardless, it seems I am forgetting, as a direct result of some failure of prospective memory. Call this *prospective forgetting*.

The Process Theory poorly explains prospective forgetting. Prospective forgetting can happen in particular circumstances or at specific times. At *noon* I forget to call you. I am not forgetting before or after noon. But processes take time. They occur over time, not at times. So the Process Theory requires that forgetting is temporally extended. It can describe me, at noon, as being somewhere in the process of forgetting, but it cannot describe my forgetting as entirely localized to noon. But it is natural to describe some prospective memory failure as occurring just at one time. This would be the time of the intended action. Since processes are temporally bloated, the Process Theory can't explain all prospective memory failure.

The Information Loss Theory struggles on the prospection test for the same reason it struggled on the metacognition test. The theory explains forgetting in terms of eliminated or inaccessible information. Now, in prospective forgetting, an intended act is not performed at the chosen time or circumstance. Something or other is not accessed then. Suppose what is not accessed is the intention to act. This is not to say that anything is inaccessible or has been eliminated at the time of forgetting. The intention can remain accessible, but just not accessed. Better retrieval cues, or better luck with retrieval cues, could have led to access. I could have accessed my intention to call you, and would have accessed it if I had set my alarm, or if I recollected what it was for when it went off. The Information Loss Theory does not account for all prospective forgetting.

On a different theory of what *loss* consists in, the Information Loss Theory could do better on the metacognition and prospection tests. The alternative theory counts not just elimination and inaccessibility as loss, but also access failure. Advocates of the Information Loss Theory do not yet count access failure as loss, but perhaps the metacognition and prospection tests will be their gadfly. However, this alternative theory does not seem promising—it seems to imply that there are losses where, intuitively, there aren't any. Suppose you are playing a trivia game, and an answer to a trivia question is stuck “on the tip of your tongue.” You are failing to access it. This new theory apparently counts this answer as lost. This tip of the tongue state does seem like forgetting of a sort, but it needn't involve loss. It is precisely because you have not lost the answer that it feels like it is on the tip of your tongue. Because you have not lost the answer, you can reasonably rule out many possible answers (“Let's see, the fifth U.S. president wasn't James Madison . . .”), and why you can in some cases detect linguistic features (“But his name is very similar to ‘James Madison’ . . .”). The revised theory

of loss looks incorrect. Advocates of the Information Loss Theory gain little if they adopt it in order to pass the metacognition and prospection tests.

## 5 The LEAD Theory of Forgetting

No theory of forgetting in the literature is fully adequate. Each fails to account for some forgetting. I offer a new theory. It aims to be maximally general, leaving open as many dimensions of forgetting as possible. So, it aims to be a maximally unified account of forgetting, showing what the various types of forgetting fundamentally have in common. Insofar as it meets this aim, it has an asset. (We'll see it has other assets too.) Perhaps nothing unites the various forgetting phenomena beyond a family resemblance or even coincidence—English-speakers just happen to call some unrelated things “forgetting.” But if there is a sturdier connection, we should look for it.

I will build up the necessary and sufficient conditions for forgetting one by one. To start:

S forgets  $x$  to extent  $e$  at  $t$  iff . . .

Here,  $x$  can be propositional, objectual, or procedural. So the theory is general with respect to content type. It is not general with respect to ontological category, however. It indexes forgetting to a time  $t$ , and so it is not about forgetting *over* time—that is, not about the process of forgetting. For now I will say what just the state of forgetting is, and later will explain the process in terms of the state. Finally, forgetting can be gradable on this theory. One forgets to an extent,  $e$ . If forgetting can be gradable, it is a further and unanswered question just how many grades there are. The grades could be few (slight, moderate, strong) or many (extremely slight, very slight, slight. . .). But the theory is also compatible with some and even all forgetting being binary. When the magnitude range of forgetting is binary, then one forgets to the greatest extent if at all. Since the theory accounts for both binary and gradable forgetting, it is general in scale.

The first necessary condition for forgetting is:

(1) S has learned  $x$  by  $t$ , and. . .

You cannot forget what you never learned. According to (1), forgetting requires learning by the time of forgetting. Normal human learning results from experience and from memory processing information that originates in experience. Learning can also have unusual causes, such as brain-tampering or trauma.<sup>14</sup> At any rate, not all learning will be conscious. It is an open question whether we count as learning information that enters only working memory and not long-term memory, or information that is a part of experience but never attended to.<sup>15</sup> In short, (1) ensures that we forget only what we have gotten.

The next necessary condition for forgetting is:

- (2) S fails to extent  $e$  to internally access  $x$  at  $t$ , and. . .

At the time of forgetting, the subject fails to access internally whatever is being forgotten. If you are internally accessing something—that Plato taught Aristotle, your experience of your breakfast—you aren't altogether forgetting it. Whatever you internally access, you *have*, in a way that forgetting precludes. The *internally* qualification here is imprecise but important. Suppose I ask you what you were doing exactly one year ago. You try to recall, without success. I then pull up some footage of what you were doing. Merely by watching the footage, you are in a sense accessing what you were doing.<sup>16</sup> Still, you forget what you were doing, even as you watch. The kind of access forgetting cannot abide is strictly internal. For simplicity I leave 'internally' implicit from here on.

Like forgetting, access failure appears to occur to greater and lesser extents. I can bring to mind just a few details of a past experience, or a host of details, or a smattering in between. According to (2), a subject forgets something only to the extent that she fails to access it. If she completely forgets, it must be that she completely fails to access. Accessibility is another matter, however. For all (2) states, something can be completely forgotten and yet accessible, even highly accessible. So we diverge here from the Information Loss Theory.

Conditions (1) and (2) say forgetting requires learning and access failure. Forgetting requires more, though. At any given time, I am failing to access most of what I have learned—even most of what I still know—but I am not forgetting all of this bulk. The final element for forgetting is a dispositional relation to what is forgotten. But it turns out that either of two dispositional relations will do. The last necessary condition for forgetting, then, is disjunctive:

- (3) (a) at  $t$ , S intends to internally access  $x$  under some description by  $t$ , or  
 (b)  $x$  is internally inaccessible to S to extent  $e$  at  $t$ .

Forgetting involves either intending to internally access, or internal inaccessibility. I will explain these dispositional relations in turn.

According to (a), one last condition for forgetting can be intending to access. Together with the first two conditions, this tells us that *forgetting can be failing to access something that was learned and is intended to be accessed*. There is more than one way that a subject can be, at the time of forgetting, intending to access something by then. One obvious option: the subject is at the time of forgetting in fact *attempting* to access what is forgotten. If, for example, I am trying to recollect your last name right now, I am thereby intending to access it right now. If my efforts are unsuccessful, and I have indeed learned your last name, then right now I am forgetting

it. Now, if I am forgetting your last name, I am failing to access it. So the content of my current attempt to recollect your last name cannot be directly identifying it (for example, as *Nieves*)—direct identification looks like access. Rather, the content of my recollective attempt identifies it indirectly, by appropriately describing it (for example, as *Maria's last name*). That is why (a) says the subject who forgets  $x$  intends to access  $x$  *under some description*.

We now have a good explanation of why the feeling of forgetting is accurate, when accurate. While a subject has the feeling of forgetting and is indeed forgetting, she is attempting (perhaps involuntarily) to access something she had learned, satisfying (1). In virtue of attempting, she is intending to access what she had learned, satisfying (a). But she is unsuccessful, satisfying (2). These three conditions jointly suffice for forgetting. Metacognition test: passed.

There is a second option for how a subject can be intending to access what she is at the time forgetting: the subject formed this intention in the past and has kept it. That is, the subject has a suitable intention stored or standing when forgetting. A closer look at prospective memory will help us make sense of this.

Prospective memory, again, is memory for intended future action. This morning I decide I will call you at noon. Something sustains this plan, allowing me to follow through on it. That is prospective memory. Although we commonly talk as though a mere act is the content of prospective remembering and forgetting, the content really is an intention. I don't remember to call you, I remember my intention to call you. Now, since I don't intend to call you just yet, for now I file away my intention to call you. I don't want it filed away forever. I want to retrieve it at the right time, noon. In fact, normally, I *intend* to retrieve it at the right time—I have an intention to access my intention to act. To see this, note how it is possible yet unusual for me to intend a future act but not intend to access the intention to act. This might occur if, for example, I suspend or lose my intention to access my intention to call you, even while keeping the intention to call. To my relief, my wife tells me she'll remind me to call you; I cease trying to remember to call you, though I still intend to call. Prospective memory commonly involves an intention to retrieve an intention for future action.

I suggest that in prospective forgetting, the subject still intends to retrieve, but at the desired time fails to retrieve.<sup>17</sup> Suppose I haven't outsourced the reminder task to my wife. I intend to access at noon my intention to call you. Noon comes, and I do not access the intention to call, but I still have the intention to access it. I am forgetting to call you. More exactly, I am forgetting my intention to call you. If, however, I have totally discharged the reminder task, I cannot be at noon forgetting my intention to call you. I no longer intend to access my intention to call, so I am not forgetting that intention. I can of course *not remember* at noon to call you, but that is not the same as forgetting.

Conditions (1), (2), and (a) are satisfied in any prospective forgetting case. At noon I fail to access my intention to call—fulfilling (2)—yet at noon I intend to access the intention to call—fulfilling (a). I formed my intention to call you by the time of forgetting it, and this acquisition of an intention is learning of a sort, a taking in. This fulfills (1). Prospection test: passed.

Now for (b), the second disjunct in condition (3). It places an internal inaccessibility condition on forgetting. Given (b) and conditions (1) and (2), we see forgetting can be something different from what accounts for the accurate feeling of forgetting and for prospective forgetting. *Forgetting can be failing to access something that was learned and is inaccessible.* This explains cases where what is forgotten is something that has been eliminated from memory or rendered at least temporarily irretrievable. The forgetter needn't have any metacognitive feeling or standing intention to access what is forgotten. With respect to your old locker combination you have no feeling of forgetting, nor any plan to retrieve it. Yet you forget it, and will continue to forget it, and have been forgetting it for some while. It's inaccessible to you.

Inaccessibility can be partial. This is vague, but intuitive enough. If many details of a previously experienced event are now inaccessible, yet other details remain accessible, then the event itself is partially inaccessible. Condition (b) links the extent of forgetting with the extent of the inaccessibility of what is forgotten. Forgetting is greater or lesser as inaccessibility is greater or lesser. Of course, condition (2) links the extent of forgetting with something else—the extent of the failure to access what is forgotten. This raises a question: if both (2) and (b) are satisfied, and the extents of the access failure and inaccessibility are unequal, to what extent does the subject forget? Is it to the extent of the access failure, or to the extent of the inaccessibility?

Answer: the lower of the two. The access failure and inaccessibility are each at least that high. Now, the inaccessibility will always be equal to or less than the access failure. That is, something will be at least as accessible as it is accessed. Sometimes what is forgotten is highly accessible while simply not accessed. What's forgotten, though, is never highly accessed while inaccessible. So, since the extent of inaccessibility is the lower of the two when there is inequality, it determines the extent of forgetting. This way, you don't count as forgetting your current and well-remembered locker combination when *not* trying to retrieve it. When not trying to retrieve it, you are failing to a maximal extent to access it. Still, it is highly accessible to you—it is inaccessible to no extent—so you aren't forgetting it to any extent.

There are no other conditions on forgetting. I propose, then, the *learning, access failure, and dispositional*—for short, LEAD—theory of forgetting:

*LEAD Theory.* S forgets  $x$  to extent  $e$  at  $t$  iff

- (1) S has learned  $x$  by  $t$ , and

- (2) S fails to extent  $e$  to internally access  $x$  at  $t$ , and
- (3) (a) at  $t$ , S intends to internally access  $x$  under some description by  $t$ , or
  - (b)  $x$  is internally inaccessible to S to extent  $e$  at  $t$ .

The LEAD Theory says that a subject who forgets is failing to access something both learned and either inaccessible or intended to be accessed. I have been all too brief about what learning, internal access, and inaccessibility consist in. Those elements of the LEAD Theory deserve exclusive analysis on another day. It is enough work for now to establish just the structure of forgetting.

Even by its structure we can see the LEAD Theory has a number of virtues. Jointly, they are strong evidence for the theory. I have laid out two virtues already: the LEAD Theory passes both the metacognition test and the prospection test. What's more, the LEAD Theory accounts for all forgetting that the Information Loss theory accounts for, where the target information is irretrievable. It is a further virtue of the LEAD Theory that it unites various kinds of forgetting. It shows for example what prospective forgetting has in common with the forgetting that comes with a metacognitive feeling. It is general with respect to content type, connecting propositional, objectual, and procedural forgetting. And so on. There are many species of forgetting, and the LEAD Theory picks out their genus.

Another virtue of the LEAD Theory is that it is not in jeopardy if human memory turns out to be generative. Empirical research suggests our memory is not a simple warehouse that shelves the items we deposit and later withdraw. Instead, it looks like memory disassembles, discriminatingly discards, redesigns, and reassembles, even in cases where we end up withdrawing a fairly faithful representation of the past.<sup>18</sup> Memory constructs, not just a little, and not just when malfunctioning. This presents no problems for the LEAD Theory. If memory meddles with what we give it, it doesn't follow on the LEAD Theory that we are forgetting, since the meddling needn't affect accessibility. Accessing could be a generative process; accessibility could be a matter of generative power, and inaccessibility could result from generational troubles and not just storage troubles.

On the LEAD Theory, forgetting is compatible with a few things that might surprise us. This is a virtue of the theory because the compatibility is correct but hard to achieve. Forgetting is compatible with knowing, for example. Some ways of forgetting leave room for knowledge. Only certain ways preclude it.<sup>19</sup> Suppose in a trivia game, you are asked who the fifth U.S. president was. You know it was James Monroe. But this knowledge remains dispositional, even after you are asked—it is standing, not before your mind. You have a feeling of forgetting, and are trying but failing to access this information you learned. So, you are forgetting something you simultaneously know, namely, that James Monroe was the fifth U.S. president. While having this feeling of forgetting, you might even have a

second metacognitive feeling about the same information—the feeling of knowing! Sometimes both feelings are accurate.

Similarly, on the LEAD Theory, forgetting is compatible with remembering. When a subject is forgetting  $x$  because she is trying but failing to access  $x$ , she can on many theories still be remembering  $x$ .<sup>20</sup> To be remembering and forgetting the same thing at once is not rare. I am forgetting your last name. Nonetheless I am confident I still remember it, and so I continue to try to retrieve it. Eventually I successfully retrieve it, and this is because I remembered it all along. The remembering had just been dispositional or standing.

In light of its virtues, I conclude that the LEAD Theory is the one to beat.

## 6 Process Forgetting

The LEAD Theory does not, however, tell the full story of forgetting. It accounts for forgetting with a variety of content types, relations to content, and scales. But it is just a theory of *stative forgetting*, forgetting as a state. It is not, in other words, general with respect to ontological category. A final chapter in the full story of forgetting says what the *process* of forgetting is, and says what relation the process bears to the state.

I'll sketch a reductionist theory of the process of forgetting (or *process forgetting*). It takes a simple view of what a process in general is: a sequence of states. If a process of any kind is just a sequence of states, then the process of forgetting is just a sequence of states. This suggests a formula: build up a view of process forgetting by identifying its constitutive states and their relation. I propose that the constitutive states are states of forgetting, and that their relation is one of increasing strength. That is, *the process of forgetting is that of going from a state of lesser forgetting to a state of greater forgetting*. More formally:

*Reductionist Process Theory.*  $S$  is in the process of forgetting  $x$  from  $t$  to  $t+n$  iff for every pair of sequential times  $t_m$  and  $t_{m+1}$  in the interval from  $t$  to  $t+n$ ,

- (1)  $S$  forgets  $x$  to extent  $e_1$  at  $t_m$ ,
- (2)  $S$  forgets  $x$  to extent  $e_2$  at  $t_{m+1}$ , and
- (3)  $e_1 < e_2$ .

Since a process is a sequence of states on this theory, a process is extended through time. So, the left-hand side indexes the process of forgetting to spans of time. Whatever the potential extents of forgetting are, they increase over every time during the process of forgetting. This respects our intuitive judgments about process forgetting. When it seems someone is in the process of forgetting, it seems the subject ends up in a greater state of forgetting than before. If the extent of forgetting seems fixed over a stretch of time, the

subject does not seem to be in the process of forgetting during it. The subject could, however, be in a persisting *state* of forgetting over those times.

This points the spotlight at a potential ambiguity in our forgetting-attributions. That someone is forgetting over time underdetermines whether the person is in a persisting state of forgetting, or is instead in the process of forgetting. On the LEAD Theory, someone could count as being in the state of forgetting over a long stretch of time—the subject could continue to forget something long forgotten. This seems odd if taken to suggest that as time passes the subject is repeating a process of forgetting, or is forgetting more severely something already completely forgotten. But a continuing state of forgetting does not suffice for a continuing process of forgetting. Continuously forgetting can be like continuously knowing, rather than continuously coming to know. There are two ways to keep forgetting.

The Reductionist Process Theory does not follow from the LEAD Theory. It is compatible with any gradable theory of stative forgetting. It is an alternative to Williamson's strong process theory, which denies that there is any stative forgetting. If the Reductionist Process Theory is correct, then Williamson's strong process theory is incorrect, and several of his other claims from Section 3 teeter. On the Reductionist Process Theory, since there are processes of forgetting, there are states of forgetting. But if we grant Williamson that forgetting is factive, then the state of forgetting is a factive stative attitude that does not require knowledge. So, contra Williamson, knowledge is not the *most general* factive stative attitude. As a result, there is less evidence for Williamson's claim that the concept of knowledge is central in our thinking.

## 7 Conclusion

I have surveyed many dimensions of forgetting and challenged the few theories about it on offer. I have put forth a unified theory of the state of forgetting, and a theory of the process of forgetting, making the relation between the process and state clear. Together, my two main proposals remember all forgetting.<sup>21</sup>

## Notes

- 1 See especially Conee and Feldman (2011, p. 304). For discussion, see Frise (2015; 2017a).
- 2 A view commonly attributed to John Locke. Behan (1979) reads him differently.
- 3 Cf. Michaelian (2011, p. 402) and Tulving (1983, p. 47).
- 4 Halamish et al. (2011, 632) and Tulving and Pearlstone (1966, 389).
- 5 Friedman and Castel (2011).
- 6 Bernecker (2010, p. 198).
- 7 See Schacter (2001).
- 8 Cf. Koriat et al. (2004, p. 651).



- 9 Harris et al. (2010, p. 255) and Michaelian (2011, pp. 402–4). Cf. Roediger et al. (2010, p. 2) and Tulving (1974, p. 74).
- 10 See Tulving and Pearlstone (1966) and Michaelian (2011, p. 403–7). I note that elimination guarantees inaccessibility—what you no longer have, you cannot access. So, the *loss* relations appear to be just varieties of inaccessibility.
- 11 Cf. Arango-Muñoz (2013) and Halamish et al. (2011). See also Arango-Muñoz and Michaelian (2014) and Michaelian (2016).
- 12 Mere access failure explains the mere retrieval failure of information, which seems to have special epistemological significance. See Frise (2017b, section 4).
- 13 Cf. Annis (1980, p. 330) and Schacter (2001, pp. 51–60).
- 14 Cf. Conee and Feldman (2011, p. 305).
- 15 Schacter (2001, p. 27) thinks we forget even information that was only in working memory. Michaelian (2011, p. 402) disagrees. Presumably Schacter would say that information was learned, and Michaelian would not.
- 16 Cf. Martin and Deutscher (1966, pp. 182–3) on prompting.
- 17 McDaniel and Einstein (2007, p. 239) argue that the “retrieval of the prospective memory intention can occur even when the intention to retrieve has been suspended.” That is, prospective *remembering* can be spontaneous, not requiring an intention to retrieve the intention to act. This is orthogonal to my claim that prospective *forgetting* requires an intention to retrieve.
- 18 See Frise (2018) and Michaelian (2016).
- 19 Cf. Moon (2012, p. 356).
- 20 This is so on the simulation theory and various causal theories of remembering (Michaelian, 2016, Chs. 5–6). Incidentally, Bernecker (2008, p. 27) says “the notion of forgetting can be adequately defined only by way of appeal to the notion of remembering.” I see it as a further bonus that the LEAD Theory makes no such appeal.
- 21 I thank David James Barnett, Earl Conee, Dorothea Debus, Christoph Hoerl, Corey Maley, Kevin McCain, Kourken Michaelian, Andrew Moon, Denis Perrin, Sarah Robins, two anonymous referees, and an audience at the 2016 Memory and Subjectivity Conference for helpful discussion of a draft of this paper. I wrote this while supported by a grant from the Templeton Religion Trust. The views expressed here may not reflect those of the Templeton Religion Trust.

## References

- Annis, D. B. (1980). Memory and justification. *Philosophy and Phenomenological Research*, 40(3), 324–333.
- Arango-Muñoz, S. (2013). Scaffolded memory and metacognitive feelings. *Review of Philosophy and Psychology*, 4(1), 135–152.
- Arango-Muñoz, S., & Michaelian, K. (2014). Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical Inquiries*, 2(1), 97–122.
- Behan, D. P. (1979). Locke on persons and personal identity. *Canadian Journal of Philosophy*, 9(1), 53–75.
- Bernecker, S. (2008). *The metaphysics of memory*. New York: Springer.
- Bernecker, S. (2010). *Memory: A philosophical study*. Oxford: Oxford University Press.
- Conee, E., & Feldman, R. (2011). Replies. In T. Dougherty (Ed.), *Evidentialism and its discontents*. Oxford: Oxford University Press.

- Friedman, M. C., & Castel, A. D. (2011). Are we aware of our ability to forget? Metacognitive predictions of directed forgetting. *Memory & Cognition*, 39(8), 1448–1456.
- Frise, M. (2015). Epistemology of memory. In J. Fieser & B. Dowden (Eds.), *Internet encyclopedia of philosophy*. Retrieved from www.iep.utm.edu/epis-mem/
- Frise, M. (2017a). Internalism and the problem of stored beliefs. *Erkenntnis*, 82(2), 285–304.
- Frise, M. (2017b). Preservationism in the epistemology of memory. *The Philosophical Quarterly*, 67(268), 486–507.
- Frise, M. (2018). Eliminating the problem of stored beliefs. *American Philosophical Quarterly*, 55(1), 63–79.
- Halamish, V., McGillivray, S., & Castel, A. D. (2011). Monitoring one's own forgetting in younger and older adults. *Psychology and Aging*, 26(3), 631–635.
- Harris, C. B., J. Sutton, & A. Barnier. (2010). Autobiographical forgetting, social forgetting and situated forgetting. In S. D. Sala (Ed.), *Forgetting* (pp. 253–284). London: Psychology Press.
- Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, 133(4), 643–656.
- Martin, C. B., & Deutscher, M. (1966, April). Remembering. *Philosophical Review*, 75, 161–196.
- McDaniel, M. A., & Einstein, G. O. (2007). Spontaneous retrieval in prospective memory. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III*. New York: Psychology Press.
- McGrath, M. (2007). Memory and epistemic conservatism. *Synthese*, 157(1), 1–24.
- Michaelian, K. (2011). The epistemology of forgetting. *Erkenntnis*, 74(3), 399–424.
- Michaelian, K. (2016). *Mental time travel: Episodic memory and our knowledge of the personal past*. Cambridge, MA: MIT Press.
- Moon, A. (2012). Three forms of internalism and the new evil demon problem. *Episteme*, 9(4), 345–360.
- Naylor, A. (2015). Justification and forgetting. *Pacific Philosophical Quarterly*, 96(3), 372–391.
- Pappas, G. S. (1987). Suddenly he knows. In S. Luper (Ed.), *The possibility of knowledge: Nozick and his critics*. Totowa, NJ: Rowman & Littlefield.
- Roediger III, Henry L., Weinstein, Y., & Agarwal, P. K. (2010). Forgetting: preliminary considerations. In S. D. Sala (Ed.), *Forgetting* (pp. 1–22). New York: Psychology Press.
- Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Boston: Mariner Books.
- Tulving, E. (1974). Cue-dependent forgetting. *American Scientist*, 62(1), 74–82.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Oxford University Press.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 381–391.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.