PNAS COMMENTARY

The enigma of forgetting

John T. Wixted[a,1]

Ebbinghaus (1) showed, long ago, that forgetting follows a curvilinear trajectory, with information forgotten rapidly at first and more slowly with the continued passage of time. But why do we forget at all and why in just that way (instead of, for example, with the rate of forgetting remaining constant over time)? Setting aside pathological conditions like amnesia, common explanations for the ubiquitous phenomenon of forgetting include 1) pure decay (e.g., the processes required to maintain structural changes to the synapse after learning are subject to random error), 2) interference (e.g., other memorized information inhibits, competes with, or degrades to-be-remembered information), 3) context drift (the current context differs from the context at the time of learning, analogous to "state-dependent" memory), and 4) the adaptive inhibition of information that is still represented in the brain but is no longer needed (e.g., in Pavlovian conditioning, extinguished and seemingly forgotten memories can be restored with a single conditioning trial). One intriguing way to shed light on the theoretical mechanisms of forgetting is to investigate variables that "undo" forgetting, showing that what has been forgotten is not necessarily gone forever. In that regard, in PNAS, Bäuml et al. (2) report a counterintuitive pattern of results that provides compelling support for the idea that context drift plays a role in forgetting.

Before delving into the findings, consider, first, some longstanding laws and concepts that help to frame the issue: 1) the power law of forgetting, 2) Jost's law of forgetting, 3) Ribot's law of retrograde amnesia, and 4) the notion that memories consolidate over time. As has long been known, the time course of forgetting is generally well characterized by a power law (3), according to which $R = a_0(1+kt)^{-b}$, where R is the amount of information retained in long-term memory, t is the "retention interval" (i.e., how much time has passed since the information was encoded), and a_0, b, and k are parameters. According to this equation, when $t = 0$ (i.e., immediately after learning), $R = a_0$, which means that a_0 represents the "degree of learning." In other words, it represents how much retrievable information was initially encoded into long-term memory, when the learning context was still in effect. The parameter b represents the rate of forgetting (the larger its value, the faster information is lost with the passage of time), and the parameter k is simply a scaling constant which can be set to one for convenience. When the tested retention intervals all substantially exceed zero, as is usually the case in studies of forgetting, a simpler approximation of this equation can be (and usually is) used such that $R = at^{-b}$, where a represents the degree of retention at $t = 1$ (essentially still capturing the degree of learning) and b still represents the rate of forgetting.

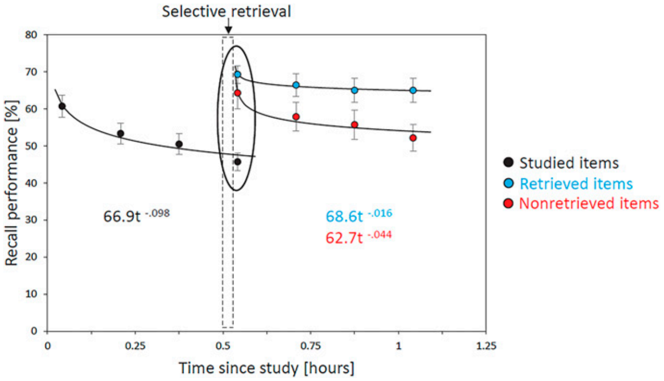

Fig. 1. Results of experiment 1. Following retrieval practice for half the items (selective retrieval), recall of both the tested and untested items was enhanced relative to recall of the studied items. However, only the tested items exhibited a reduced rate of forgetting from that point forward.

A noteworthy feature of the power law of forgetting is that the proportional rate of decay slows with the passage of time (4). For example, retrievable information might decline by 40% in the first 24 h after learning, by another 20% in the next 24 h, by another 10% in the next 24 h, and so on (unlike the constant proportional rate of decay that would be observed if forgetting were exponential in form). This property is also enshrined in Jost's (5) second law of forgetting, which states that, for two memories of the same strength but different ages, the older will decay more slowly than the younger. Ribot's (6) law of retrograde amnesia provides a possible explanation of why that might be. This law states that, as memories age, they become more resistant to the effects of disruptive forces like brain damage, electroconvulsive shock, and (one might reasonably assume) interference caused by new learning. Such increasing resistance to disruption may reflect the fact that memories consolidate and therefore stabilize over time (7, 8). As Squire and Kandel (7) put it: "A memory that has become consolidated is robust and resistant to interference. In its initial stages, even memory that would otherwise persist is highly susceptible to disruption if, for example, an attempt is made to learn some other similar material" (p. 4). Conceivably, the consolidation of long-term potentiation in the hippocampus in the hours after learning (synaptic consolidation) and the later consolidation of memories in cortex (systems consolidation) both have the effect of stabilizing memories, making them resistant to

Author affiliations: [a]Department of Psychology, University of California San Diego, La Jolla, CA 92093

Author contributions: J.T.W. wrote the paper.

The author declares no competing interest.

See companion article, "Selective memory retrieval can revive forgotten memories," 10.1073/pnas.2114377119.

[1] Email: jwixted@ucsd.edu.

disruption and helping to explain why the proportional rate of decay declines throughout the entire life of a memory.

The rate of forgetting not only slows with the passage of time, but, as has long been known, the rate at which it slows can be changed by an educationally relevant experimental manipulation known as the testing effect (9): Retrieving a previously learned item on a later memory test increases its chances of being retrieved on a future test (essentially restoring a, the original degree of learning) while also slowing the subsequent rate of forgetting (i.e., decreasing b). Why testing memory has the effects it does has long been a mystery, but the PNAS paper (2) indicates that “context reinstatement” plays a key role, and it does so by resurrecting apparently forgotten memories.

In each of three experiments, participants studied to-be-remembered information during the encoding phase (e.g., a list of words like “poet,” “moon,” “island,” and “hotel”). In the control condition, recall was cued by presenting the first letter of each word (“p___,” “m___,” “i___,” “h___”) after various retention intervals, such as 10, 20, 30, and 40 min in experiment 1. Each retention interval involved a different subgroup of participants, and the resulting forgetting function was well characterized by the power law. In the experimental condition, no testing occurred following initial learning until a “retrieval practice phase” was implemented 30 min later. The tests administered during this phase were similar to the cued recall tests for the control condition except that only half the words were tested, and there were two rounds of cued recall for the same subset of words. During the first cued recall test during retrieval practice, performance was similar to the control condition at the 30-min delay (as would be expected). However, performance improved during the second round of testing during retrieval practice, and this occurred even though no feedback was provided during the first test. The question of interest was what the subsequent pattern of forgetting would be (i.e., following two rounds of retrieval practice) for the tested and untested items.

Participants in the retrieval practice condition received a cued recall test for both the tested items and the nontested items at delays of 10, 20, 30, or 40 min following the retrieval practice phase. For the subgroup tested 10 min after retrieval practice, performance was restored to the level observed in the control condition at their 10-min mark following original learning. Remarkably, this was true for both the tested and the untested items (Fig. 1). With regard to the untested items, this is an intuitively puzzling result. Why would the untested items now be more accessible than they would have been had retrieval practice for the tested items not occurred? This surprising result can be explained by assuming that selective retrieval of half the studied items reactivated the temporal context that was in effect during study, which, in turn, facilitated the recall of both the tested and untested items to an equal degree (because they shared that same study context).

These findings may be the most direct and compelling evidence for the long-hypothesized phenomenon of context reinstatement. Moreover, in accordance with the longstanding distinction between availability and accessibility (10), they demonstrate that the memory traces of words not recalled under the cueing conditions in effect for the control group at the 40-min mark were nevertheless available in memory even though they were not accessible for retrieval (11). Conceptually similar results have been observed in animal studies in which protein synthesis inhibitors prevented the synaptic consolidation of cells that were active during memory formation. Even so, the memories could be recovered via direct optogenetic activation of those cells (12). Both context reinstatement and the direct activation of unconsolidated memory cells recovered memories that had apparently been forgotten. Whether there is a deeper connection between the two approaches to undoing forgotten memories is a mystery for future research to investigate, but the surface similarities are intriguing.

Consider next the subsequent rate of forgetting for tested and untested items following retrieval practice. For the untested items, the rate of forgetting matched the rate of forgetting for the control condition. It was as if context reinstatement simply wound back the clock to the degree of learning and state of consolidation in effect at the time of initial learning. In other words, for the untested items, reinstatement of the original context may have destabilized their context–item associations in such a way that the associations were, in essence, formed anew and had to be consolidated all over again (13). By contrast, memories of the items tested during the retrieval practice phase were not only recovered, but, going forward from that point, the rate of forgetting was reduced as well. Why? One possibility is that the explicit retrieval of a tested item may not only have reinstated the original context (rendering the tested items once again retrievable) but also selectively destabilized the tested item’s representation in a such a way that, upon reconsolidation, the strengthened trace could inherit the state of consolidation in effect at the time of retrieval practice.

> **Why testing memory has the effects it does has long been a mystery, but the PNAS paper indicates that “context reinstatement” plays a key role, and it does so by resurrecting apparently forgotten memories.**

This idea is somewhat analogous to synaptic tagging and capture, whereby a weak memory benefits from protein synthesis induced by the contemporaneous encoding of a strong memory (14). Analogously, upon being destabilized and rendered updatable by virtue of explicit retrieval (during retrieval practice), the tested item’s representation, upon reconsolidation, might benefit from the consolidation-related structural changes that had already occurred by the time retrieval occurs. This would help to explain why the rate of forgetting going forward is slower and perhaps more resistant to disruption than it would be for a memory being encoded for the first time.

These are speculative ideas, but a bit more cross-disciplinary speculation may be needed before the well-known effect of testing memory is fully understood. For the most part, the field of cognitive psychology has adamantly resisted the notion that consolidation offers anything useful for understanding forgetting, a proud tradition that dates back to the 1930s and continues to this day. At the same time, neuroscientists have paid less-than-optimal attention to what cognitive psychology has learned about the cue-dependent nature of memory and its role in forgetting. This paper by Bäuml and Trißl (2) offers findings that should be intriguing enough to warrant the attention of scientists in both camps.

1. H. Ebbinghaus, *Memory: A Contribution to Experimental Psychology* (Columbia University, New York, 1913).
2. K.-H. T. Bäuml, L. Trißl, Selective memory retrieval can revive forgotten memories. *Proc. Natl. Acad. Sci. U.S.A.* **119**, 10.1073/pnas.2114377119 (2022).
3. J. T. Wixted, S. K. Carpenter, The Wickelgren power law and the Ebbinghaus savings function. *Psychol. Sci.* **18**, 133–134 (2007).
4. J. T. Wixted, On common fround: Jost's (1897) law of forgetting and Ribot's (1881) law of retrograde amnesia. *Psychol. Rev.* **111**, 864–879 (2004).
5. A. Jost, Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen [in German]. *Z. Psychol. Physiol. Sinnesorgane* **16**, 436–472 (1897).
6. T. Ribot, *Les Maladies de la Memoire* (Germer Bailliere, Paris, 1881).
7. L. R. Squire, E. R. Kandel, *Memory: From Mind to Molecules* (Scientific American Library, New York, 1999).
8. J. L. McGaugh, Memory–A century of consolidation. *Science* **287**, 248–251 (2000).
9. H. L. Roediger III, J. D. Karpicke, Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol. Sci.* **17**, 249–255 (2006).
10. E. Tulving, Z. Pearlstone, Availability versus accessibility of information in memory for words. *J. Verbal Learn. Verbal Behav.* **5**, 381–391 (1966).
11. T. J. Ryan, P. W. Frankland, Forgetting as a form of adaptive engram cell plasticity. *Nat. Rev. Neurosci.* **23**, 173–186 (2022).
12. T. J. Ryan, D. S. Roy, M. Pignatelli, A. Arons, S. Tonegawa, Memory. Engram cells retain memory under retrograde amnesia. *Science* **348**, 1007–1013 (2015).
13. Y. Dudai, The neurobiology of consolidations, or, how stable is the engram? *Annu. Rev. Psychol.* **55**, 51–86 (2004).
14. R. L. Redondo, R. G. Morris, Making memories last: The synaptic tagging and capture hypothesis. *Nat. Rev. Neurosci.* **12**, 17–30 (2011).