Neuron
Perspective

CellPress

The Persistence and Transience of Memory

Blake A. Richards[1,2,3] and Paul W. Frankland[4,5,6,7,8,*]
[1]Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON M1C 1A4, Canada
[2]Department of Cell and Systems Biology, University of Toronto, Toronto, ON M5S 3G5, Canada
[3]Learning in Machines and Brains Program, Canadian Institute for Advanced Research, Toronto, ON M5G 1M1, Canada
[4]Hospital for Sick Children, Program in Neurosciences and Mental Health, Peter Gilgan Centre for Research and Learning, 686 Bay Street, Toronto, ON M5G 0A4, Canada
[5]Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada
[6]Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada
[7]Institute of Medical Sciences, University of Toronto, Toronto, ON M5S 1A8, Canada
[8]Child and Brain Development Program, Canadian Institute for Advanced Research, Toronto, ON M5G 1M1, Canada
*Correspondence: paul.frankland@sickkids.ca
http://dx.doi.org/10.1016/j.neuron.2017.04.037

The predominant focus in the neurobiological study of memory has been on remembering (persistence). However, recent studies have considered the neurobiology of forgetting (transience). Here we draw parallels between neurobiological and computational mechanisms underlying transience. We propose that it is the interaction between persistence and transience that allows for intelligent decision-making in dynamic, noisy environments. Specifically, we argue that transience (1) enhances flexibility, by reducing the influence of outdated information on memory-guided decision-making, and (2) prevents overfitting to specific past events, thereby promoting generalization. According to this view, the goal of memory is not the transmission of information through time, per se. Rather, the goal of memory is to optimize decision-making. As such, transience is as important as persistence in mnemonic systems.

> We do not remember days, we remember moments. The richness of life lies in memories we have forgotten.
>
> —Cesare Pavese (*This Business of Living*)

Introduction

Memory allows for the transmission of information through time. Most people, including many scientists, view the ideal mnemonic system as one of perfect *persistence*. That is, a system that transmits the greatest amount of information, with the highest possible fidelity, across the longest stretches of time. However, the few examples we have of individuals with something approximating this "perfect" mnemonic persistence suggest that remembering everything comes at a price. The Soviet clinical neuropsychologist A. R. Luria described the case of Patient S., a man with "vast memory" who could only forget something if he actively willed himself to do so (Luria, 1968). Nonetheless, according to Luria's accounts, Patient S. was handicapped by his apparent super-human memory. While on one hand, he was able to remember instances in exquisite detail, his memory was inflexible and he was unable to generalize across instances. This points to the importance of forgetting (or *transience*) as a critical component of a healthy mnemonic system.

Perhaps reflecting this preoccupation with memory as a means of making information permanent, the traditional focus in neurobiological studies of memory has been on mechanisms that promote the persistence of information (Bliss and Collingridge, 1993; Kandel, 2001; McGaugh, 2000; Poo et al., 2016). But this focus is shifting. There has been a recent increase in the number of studies concerned with the neurobiological mechanisms of memory transience (Berry and Davis, 2014; Frankland et al., 2013; Hardt et al., 2013a). Here we first briefly review the large literature concerned with neurobiological mechanisms of memory persistence. We then turn to the fledgling literature concerned with neurobiological mechanisms of memory transience. Based on principles from machine learning and computational neuroscience, we propose that it is the interaction between these two processes (i.e., persistence × transience) that optimizes memory-guided decision-making in changing and noisy environments. Specifically, we propose that only by combining persistence and transience can individuals exhibit flexible behavior and generalize past events to new experiences.

Persistence

A Neurobiological Definition of Persistence

Remembering transports us back in time, allowing us to re-experience some past event or experience, a form of mental time travel (Tulving, 2002). At the neural level, this suggests that some aspect of our present brain state reflects a past brain state corresponding to the remembered event. Perhaps most simply, remembering might involve reactivation of the patterns of neural activity that were present at encoding. This is the scenario favored by many neuroscientists (e.g., Josselyn et al., 2015; Tonegawa et al., 2015). Yet computationally there are alternate ways in which our present brain state might reflect our past brain state. As long as our current brain state is statistically dependent on our previous brain states, some information is preserved (Richards and Frankland, 2013). According to this perspective, any circuit level changes that increase (or decrease) the probability of particular brain states appearing promote persistence and, therefore, the transmission of information through time.

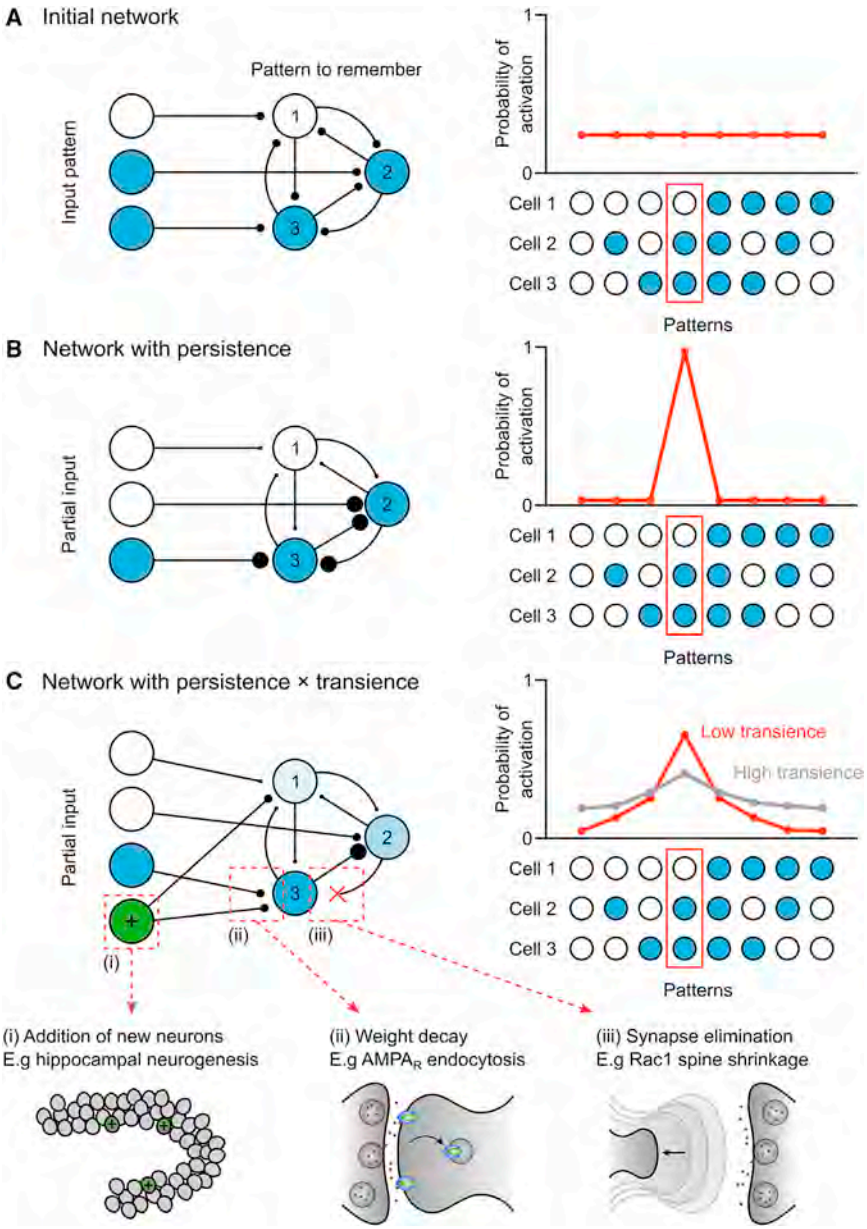

Figure 1. Persistence and Transience in Memory Networks

(A) In a naive network with uniform and/or random synaptic connections, the probability of any individual activity pattern is roughly equivalent. (Here patterns are illustrated by showing active cells in blue.) Memory storage requires that the specific pattern of activation induced by inputs to the network must be stored (illustrated here as cells 2 and 3 being active and highlighted with a red box in the diagram to the right).

(B) To store this specific pattern, the network can potentiate the synapses between the co-active cells and depontentiate the other synapses. This will increase the probability that the specific activity pattern will re-emerge later, even in response to partial inputs to the network.

(C) By employing mechanisms for mnemonic transience, such as the addition of new neurons, synaptic decay, or synaptic elimination, the network can generalize the increased probability of reactivation to other similar patterns of activity (illustrated here by adjacent patterns to the remembered one).

This framework assumes that persistence requires that changes induced during encoding are relatively *stable*. There is support for this idea, at least in the short- to intermediate term, and we review these data below. In particular, we highlight recent chemo- and optogenetic "engram" studies that show that remembering is associated with stable network changes and reactivation of patterns of activity present at encoding.

Persistence in the Short- and Intermediate Term

The reactivation of neurons that were active at the time of encoding can be achieved with fairly simple rules for forming or altering synaptic connections. Indeed, the classic articulation of memory storage, as first proposed by Hebb (Hebb, 1949), is that some form of synaptic strengthening between coactive neurons during encoding provides the basis for formation of cell assemblies that correspond to the engram (Josselyn et al., 2015; Tonegawa et al., 2015) (Figures 1A and 1B). The subsequent discovery (Bliss and Gardner-Medwin, 1973; Bliss and Lomo, 1973) that high-frequency stimulation induces long-lasting increases in synaptic strength between neurons (or long-term potentiation; LTP) provided the modern framework for understanding how cell assemblies, and therefore memories, might be formed and maintained (Stevens, 1998). Ex vivo experiments, predominantly using the hippocampal slice preparation, identified a large number of intracellular signaling mechanisms that are necessary for either the induction or maintenance phases of LTP (Bliss and Collingridge, 1993; Malenka and Bear, 2004; Sacktor, 2011; Sanes and Lichtman, 1999). Moreover, pharmacological or genetic interventions that targeted these same cascades in vivo typically produced analogous effects on memory formation (Morris et al., 1986; Pastalkova et al., 2006; Silva et al., 1992; Whitlock et al., 2006) (for an exception, see, for example, Bannerman et al., 2006). While most studies have emphasized parallels between synaptic strengthening and memory persistence, weakening of synapses (e.g., via long-term depression; LTD) can equally be used to persistently store information by promoting brain states that reflect the past (Bear, 1996; Hopfield, 1982; Kemp and Manahan-Vaughan, 2007). Indeed, interventions that eliminate LTD typically also disrupt memory formation (Kemp and Manahan-Vaughan, 2007).

Perhaps the most direct support for the idea that remembering involves reactivation of neural patterns that were present during encoding has come from recent genetic tagging experiments.

Crucially, these methods allow neural ensembles active during memory encoding to be manipulated at later time points using opto- or chemogenetics. Using a variety of different amygdala- and hippocampus-dependent tasks, three types of evidence have emerged from these experiments (Josselyn et al., 2015). First, neurons that were activated at the time of encoding are reactivated at above-chance levels when the corresponding memory is "naturally" retrieved (Denny et al., 2014; Reijmers et al., 2007; Tanaka et al., 2014). Using these methods, reactivation rates were quite modest in some regions (e.g., dentate gyrus, ~5%; Denny et al., 2014) but more robust in others (e.g., CA1, ~40%; Tanaka et al., 2014). Second, if reactivation of these "tagged" neurons is prevented in a recall test, memory retrieval is compromised (Berndt et al., 2016; Denny et al., 2014; Han et al., 2009; Hsiang et al., 2014; Park et al., 2016; Rashid et al., 2016; Tanaka et al., 2014; Zhou et al., 2009). Preventing reactivation of tagged neurons impairs expression of both aversively motivated (Berndt et al., 2016; Denny et al., 2014; Han et al., 2009; Park et al., 2016; Rashid et al., 2016; Tanaka et al., 2014; Zhou et al., 2009) as well as appetitively motivated (Hsiang et al., 2014) memories. Third, activation of populations of tagged cells is sufficient to induce "artificial" recall (Cowansage et al., 2014; Liu et al., 2013; Ohkawa et al., 2015; Ramirez et al., 2013; Rogerson et al., 2016; Yiu et al., 2014). These artificially recalled memories seem to behave similarly to natural memories. For example, inhibiting protein synthesis during retrieval blocks reconsolidation of artificially expressed fear memories, leading to reduced conditioned fear levels in subsequent tests (Kim et al., 2014). Together, these studies indicate that partial reactivation of the activity patterns present at encoding is both necessary and sufficient for hippocampus- and amygdala-dependent memory persistence.

Transience

A Neurobiological Definition of Transience

If stable changes in synaptic connectivity promote persistence, then, conversely, forgetting occurs when modified synapses are destabilized. In situations in which neural connectivity can be assumed to be reasonably static (for example, over short spans of time), transience might involve reversing potentiated or depressed synaptic connections or eliminating newly formed synaptic connections (Figure 1C). However, over longer time frames connectivity is likely less stable. In these situations, manipulations that promote circuit dynamism likely also promote transience, whereas manipulations that promote circuit stability likely promote persistence.

While the neurobiological study of forgetting is in its infancy, recent studies have found examples corresponding to both of these means of achieving transience. We review these below.

Transience on Short Timescales

Artificial Induction of Forgetting. At a cellular level, just as synaptic strengthening is associated with insertion of GluA2-containing AMPA receptors in the postsynaptic membrane, depotentiation is associated with reversal of these changes (Collingridge et al., 2010; Hardt et al., 2013b). Therefore, interventions that promote GluA2-containing AMPA receptor endocytosis might also promote forgetting. Conversely, interventions that inhibit this process might prevent forgetting. A recent series of studies has begun to address these hypotheses.

The atypical protein kinase C (PKC) isoform, PKM-ζ, plays a key role in maintaining LTP and memory (Sacktor, 2011). Following LTP induction, administration of inhibitors of PKM-ζ such as the peptide ZIP leads to depotentiation in hippocampal slice preparations (Ling et al., 2002). Similarly, following memory formation, local infusion of ZIP induces memory erasure (Pastalkova et al., 2006; Tsokas et al., 2016). These erasure effects have been observed using a variety of different behavioral paradigms (Serrano et al., 2008), as well as using genetic interventions to inhibit PKM-ζ (Tsokas et al., 2016) as an alternative to ZIP.

Several lines of evidence suggest that the amnestic effects of PKM-ζ inhibition are mediated by GluA2-containing AMPA receptor endocytosis. In hippocampal slice preparations, administration of PKM-ζ increases AMPA-mediated currents and promotes insertion of GluA2-containing AMPA receptors into the postsynaptic membrane (Ling et al., 2006; Yao et al., 2008). This suggests that PKM-ζ promotes LTP maintenance by preventing GluA2-containing AMPA receptor endocytosis. Critically, when GluA2-containing AMPA receptor endocytosis is prevented (by bath application of an interfering peptide, GluA2-3Y), ZIP no longer reverses LTP (Dong et al., 2015; Migues et al., 2010). An identical pattern of results is observed behaviorally. Following memory formation, preventing endocytosis of GluA2-containing AMPA receptors eliminates the amnestic effects of ZIP (Dong et al., 2015; Migues et al., 2010).

While these studies indicate that reversing learning-induced changes in synaptic strength produces forgetting, these PKM-ζ manipulations affect all synapses in the targeted region, and not just those associated with the memory. Two recent studies have addressed this anatomical specificity issue by targeting only potentiated synapses.

The first of these used a Pavlovian fear conditioning paradigm (Nabavi et al., 2014). When a conditioned stimulus (CS), such as a tone, is paired with an unconditioned stimulus (US), such as a footshock, animals develop conditioned fear to the tone. This form of conditioning depends on plasticity in the lateral amygdala (LA), which receives cortical and thalamic inputs conveying information about the CS (i.e., tone) and thalamic inputs conveying information about the US (i.e., shock).

During conditioning, the CS input into LA is potentiated (McKernan and Shinnick-Gallagher, 1997; Rogan et al., 1997), such that presentation of the CS alone is now sufficient to elicit fear behaviors (Blair et al., 2001). Nabavi and colleagues showed that this behaviorally induced potentiation can be mimicked by replacing the tone with optogenetic stimulation of the auditory cortex. Following pairing of optical stimulation of this pathway with a US, presentation of the optical CS alone was sufficient to elicit fear. Crucially, they then showed that depotentiating this pathway (by low-frequency optical stimulation of the CS pathway at 1 Hz) led to loss of conditioned responses to the optical CS. Remarkably, re-potentiation of the same pathway restored conditioned fear (Nabavi et al., 2014).

Synapses may weaken, but they also may be eliminated. Therefore, elimination of synapses potentiated during learning should also lead to memory loss. This technically challenging experiment was recently conducted by Hayashi-Takagi and

colleagues (2015). In these experiments, mice were generated in which a photactivatable version of the Rho GTPase Rac1 was targeted to potentiated spines via a dendritic targeting element of Arc mRNA. Subsequent photoactivation of this Rac1 construct led to spine shrinkage, and this allowed them to test whether shrinking recently potentiated spines would erase a memory. To do this, mice were trained on a rotarod. This form of motor learning induces significant synaptic remodeling in subsets of motor cortex neurons. Following learning, light-induced shrinkage of potentiated spines led to loss of acquired motor skills, and thus showed that elimination of learning-induced synaptic growth leads to loss of the corresponding memory (Hayashi-Takagi et al., 2015).

Natural Forgetting. Memory loss following spine shrinkage, ZIP administration, or experimentally induced depotentiation is rapid. However, psychological studies of natural forgetting indicate that memory loss is typically much more gradual (Ebbinghaus, 1913). To what extent do the processes identified above contribute to constitutive decay mechanisms underlying natural forgetting?

Whether AMPA receptor endocytosis plays a role in natural forgetting has been evaluated in two recent papers (Dong et al., 2015; Migues et al., 2016). In both studies, rats were trained in hippocampus-dependent tasks where natural forgetting emerges over hours or days. In one study, rats were trained in an inhibitory avoidance task (Dong et al., 2015). In this test, rats are placed in the lighted side of a two-compartment apparatus. Rats prefer darkness, so after a short time they enter the dark compartment. However, entry into the dark compartment is punished with a footshock, and when later replaced in the lighted compartment of the apparatus rats are more reticent to enter the dark compartment. The use of strong shocks can produce avoidance that lasts weeks (Inda et al., 2011). Using lower intensity shocks produces a more transient memory, though, and, using this weaker training protocol, Dong and colleagues found that rats exhibited robust inhibitory avoidance memory when tested 1 hr, but not 24 hr, after training. Remarkably, post-training infusion of an interfering peptide that prevents GluA2 endocytosis extended the lifetime of this memory. Rats that received intra-hippocampal infusions of the peptide showed intact memory 1 day following training (Dong et al., 2015).

Studies of CA1 LTP in freely behaving rats revealed a similar pattern of results (Dong et al., 2015). Whereas strong Schaffer collateral stimulation produced a stable, non-decaying form of LTP that lasted more than 24 hr, weaker stimulation produced a decaying LTP that returned to baseline within 2 hr of tetanization. However, just as blocking GluA2 endocytosis prevented forgetting, the same manipulation prevented LTP decay over this time course (Dong et al., 2015). The behavioral effects of blocking GluA2 endocytosis on forgetting were additionally observed using different strategies to prevent GluA2 endocytosis, and with different tasks including food-conditioned place preference and an object location task (Migues et al., 2016).

In addition to weakening synaptic connections, there is also evidence that synapse elimination (via shrinkage or loss of spines) underlies natural forgetting. In particular, Rac—a key regulator of actin dynamics—appears to regulate forms of natural forgetting in flies and mice. In the initial work in flies, Rac inhibition in mushroom body neurons extended the lifespan of an odor-shock memory from hours to more than a day. Conversely, overexpression of a dominant active Rac in the mushroom body accelerated forgetting of this aversive memory (Dong et al., 2016; Shuai et al., 2010).

In mice, Rac inhibition (via hippocampal expression of a dominant-negative form of Rac1) similarly slowed forgetting of an object recognition memory. In control mice, object recognition was evident 1, but not 3, days following training. In Rac1-deficient mice, object recognition was still evident 5 days following training. In contrast, hippocampal overexpression of a constitutively active form of Rac1 accelerated decay, with performance falling to chance levels within 24 hr of training (Liu et al., 2016) (see also Jiang et al., 2016 for similar results using pharmacological inhibition of Rac1). The same manipulations bidirectionally modulated the stability of LTP, measured in CA3. Decreasing Rac1 activity prevented LTP decay, whereas increasing Rac1 activity accelerated loss of LTP (Liu et al., 2016).

Racs play important roles in a number of cellular processes including cell migration, cell polarity, and cell cycle (Hodge and Ridley, 2016). However, additional experiments in flies that targeted upstream regulators and downstream targets of Rac suggest that Racs regulate forgetting via a direct interaction with cofilin, a potent promoter of actin depolymerization. Indeed, flies expressing a mutant form of cofilin exhibited reduced forgetting in the odor-shock paradigm, phenocopying the effects of Rac inhibition (Shuai et al., 2010). Furthermore, accelerated forgetting following overexpression of the dominant active Rac was prevented by genetic disruption of this pathway. These results suggest that a Rac-cofilin pathway regulates forgetting by promoting actin depolymerization and cytoskeleton remodeling. Interventions that promoted this process induced forgetting, whereas interventions that inhibited this process promoted stabilization. More recent work has provided evidence that this pathway is modulated by dopamine (via dopamine receptor activation and the scaffolding protein Scribble), and thus provides a link between different behavioral states (e.g., sleep) and transience (Berry et al., 2012, 2015; Cervantes-Sandoval et al., 2016). During sleep, dopaminergic activity in flies is lower, leading to reduced activation of this pathway and reduced forgetting (Berry et al., 2015).

Synapse elimination has also been associated with natural forgetting in the nematode *C. elegans*. In *C. elegans*, associative learning induces synapse growth at a specific neuron (AVA). However, within hours of learning, this modified synapse shrinks, and reversion to a naive state is accompanied by loss of the associative memory. Genetic and pharmacological interventions targeting regulators of the actin cytoskeleton modulate both synapse stability and natural forgetting (Hadziselimovic et al., 2014). For example, upregulation of the Arp2/3 complex promotes actin branching, stabilizes modified synapses, and prevents forgetting. Conversely, downregulation of the Arp2/3 complex reduces actin branching, accelerates synapse loss, and accelerates forgetting.

Therefore, a collection of processes that "downscale" (i.e., weaken or eliminate) potentiated synapses likely promotes memory transience. Such synaptic downscaling may occur at any moment. However, certain phases of sleep in mammals

may represent a privileged time window for these processes to occur (Tononi and Cirelli, 2006). For an overview of how sleep, and in particular different sleep stages, may modulate persistence and transience, see Stickgold and Walker (2013).

Transience on Longer Timescales

In scenarios where neural connections are reasonably static, reversal of learning-induced changes should be sufficient to induce forgetting, as described above. However, there is plenty of evidence that patterns of brain connections are not static, especially over longer time spans. Endogenous processes continuously remodel the brain, cumulatively altering patterns of connectivity. For example, spines are dynamically regulated, with significant turnover of at least subpopulations of spines throughout the brain. Moreover, in the adult brain, neurogenesis persists in at least two regions (the subgranular and subventricular zones). Neurons generated from these niches migrate and integrate into existing hippocampal and olfactory circuits, respectively, thereby altering connectivity. Finally, ongoing neural activity itself has the potential to modify synaptic weights. Such neural activity might result from new learning, or simply reflect basal brain states (or noise).

Collectively, these processes may not only alter or eliminate existing connections but additionally introduce new connections. Moreover, they may be regulated by intrinsic and extrinsic factors (e.g., hippocampal neurogenesis is upregulated by exercise [van Praag et al., 1999] and downregulated by stress [Gould et al., 1997]). Therefore, interventions that promote remodeling should induce instability, and therefore transience. Conversely, interventions that reduce remodeling should promote stability, and therefore persistence. Perhaps the most straightforward evidence for these ideas has emerged from in vivo studies of LTP.

Rodent LTP Studies

LTP is typically studied over the course of minutes to hours in hippocampal slice preparations. However, LTP may be studied over considerably longer times in vivo, with LTP persistence dependent on the induction protocol. For example, in the perforant path-dentate gyrus synapse, weaker theta-burst stimulation produces a form of LTP that decays to baseline levels within 3–5 days (Kitamura et al., 2009; Villarreal et al., 2002). However, stronger theta-burst stimulation can produce forms of LTP that last considerably longer, even up to a year (Abraham et al., 2002). Might interventions that inhibit remodeling of hippocampal circuits make decaying forms of LTP more durable? Conversely, might interventions that promote remodeling of hippocampal circuits weaken stable forms of LTP? These hypotheses have been addressed in a series of studies.

Villarreal and colleagues asked whether reducing synaptic activation following tetanization would convert a decaying LTP into a non-decaying form (Villarreal et al., 2002). Following tetanization, rats received daily systemic injections of an NMDA antagonist. Whereas LTP decayed within days in the untreated mice, NMDA receptor blockade prevented this decay (see also Sachser et al., 2016). Rats trained in the radial arm maze exhibited forgetting over a similar time course. In parallel with the LTP experiments, NMDA receptor blockade following training improved spatial memory retention. This pattern of results has been extended to other behavioral paradigms including an object location task (Sachser et al., 2016). These results suggest that LTP is a normally persistent process that is actively reversed by NMDA receptor activation. Depotentiation depends on NMDA receptor activation (Lüscher and Malenka, 2012; O'Dell and Kandel, 1994), and therefore it is plausible that spontaneous, sporadic synaptic activation associated with noise leads to the gradual weakening of potentiated synapses (and therefore memory transience).

Abraham and colleagues addressed the converse question (Abraham et al., 2002). They asked whether an intervention that promotes synaptic remodeling in the hippocampus—environmental enrichment—would be sufficient to convert the non-decaying LTP into a decaying form. Following tetanization, rats were either housed conventionally or in an enriched environment (for 3 weeks, starting 2 weeks after tetanization). In the conventional housing condition, LTP was stable, as expected. However, in the enrichment condition, LTP decay was observed. Even in basal states, it is likely that there is considerable synaptic turnover in the hippocampus, with turnover rates in the hippocampus likely higher than those in other brain regions (Attardo et al., 2015). These results therefore suggest that turnover rates (and therefore transience) may be tuned by environmental factors.

Hippocampal Neurogenesis and Transience

Environment enrichment additionally promotes neurogenesis in the adult hippocampus (Kempermann et al., 1997). The adult dentate gyrus contains neural stem cells (Reynolds and Weiss, 1992) that can generate new neurons throughout life. As these newly generated neurons mature, they establish connections with appropriate presynaptic (e.g., perforant path inputs from entorhinal cortex) and postsynaptic (e.g., mossy fiber synapses onto CA3 pyramidal cells) partners and, in doing so, continuously remodel hippocampal circuits (Gonçalves et al., 2016). This remodeling process is competitive (McAvoy et al., 2016), with newborn neurons competing with established dentate granule cells for inputs from the entorhinal cortex and outputs onto CA3 pyramidal cells. The net result is that new synaptic connections may co-exist with or, in some cases, replace established synaptic connections (Toni et al., 2007, 2008).

Experimental support for the idea that neurogenesis regulates transience in the hippocampus has emerged from several recent studies in rodents (Akers et al., 2014; Epp et al., 2016; Ishikawa et al., 2016; Kitamura et al., 2009). For example, one study examined the stability of perforant path-dentate gyrus LTP in awake, behaving rats (Kitamura et al., 2009). In these experiments, the strength of perforant path-dentate gyrus LTP was monitored over weeks in awake, behaving rats. In control rats, LTP reverted back to baseline within about a week. However, in irradiated rats (in which hippocampal neurogenesis was eliminated), LTP was prolonged with potentiation still evident 2 weeks following induction. These results complement the Abraham et al. study (Abraham et al., 2002). Whereas promoting remodeling (via environmental enrichment) induced decay of LTP, inhibiting remodeling (via blocking neurogenesis) promoted persistence of LTP. Together, they suggest that neurogenesis provides a continuous decay signal, by remodeling hippocampal circuits and "overwriting" established LTP.

More recently, the relationship between neurogenesis and transience has been assessed behaviorally (Akers et al., 2014; Epp et al., 2016; Ishikawa et al., 2016). In one study, Akers and

colleagues trained mice in contextual fear conditioning (Akers et al., 2014). During a 6 week period following training, mice either had access to a running wheel in their home cage or were housed conventionally, and contextual fear was subsequently assessed. As expected, voluntary exercise was robustly neurogenic, doubling rates of neurogenesis. However, these exercise-induced increases in hippocampal neurogenesis weakened the contextual fear memory. These forgetting effects were observed using other aversively and appetitively motivated hippocampus-dependent tasks (including water maze, Barnes maze, inhibitory avoidance, and an olfactory paired associate task; Akers et al., 2014; Epp et al., 2016; Ishikawa et al., 2016). Similar results were observed using genetic (e.g., conditional deletion of *p53* [e.g., Akers et al., 2014] and pharmacological [e.g., memantine; Akers et al., 2014; Ishikawa et al., 2016]) interventions to artificially elevate hippocampal neurogenesis in post-training periods. Conversely, suppressing hippocampal neurogenesis produced the opposite pattern of results. Following training in the water maze, forgetting emerged over a 6 week period in conventionally housed control mice. Suppressing hippocampal neurogenesis within this window attenuated this form of natural forgetting (Epp et al., 2016).

These bidirectional effects of manipulating levels of hippocampal neurogenesis support the idea that neurogenic remodeling of hippocampal circuits plays a causal role in natural forgetting. Similar to other forgetting mechanisms described above, neurogenesis-based remodeling presumably changes connectivity in such a way as to reduce the likelihood of particular activity patterns reappearing.

Persistence × Transience

> In the practical use of our intellect, forgetting is as important as remembering.—William James (*The Principles of Psychology*)

Above, we reviewed a number of neurobiological mechanisms that can promote mnemonic transience. The most intuitive explanation for why the brain possesses these mechanisms is that they help to "make room" for new memories. However, when we consider the sheer number of neurons and synapses in the brain, it would seem that there is ample capacity to store many more memories than we actually do. For example, the human brain is estimated to have roughly 80–90 billion neurons (Azevedo et al., 2009). If we were to reserve only a tenth of those for memories of specific events, then according to computational estimates of capacity in auto-associative networks, we could reliably store approximately one billion individual memories (Amit et al., 1985). Furthermore, when we consider sparsely encoded memories this number can increase by several orders of magnitude (Amari, 1989). Given that it is apparently possible to remember far more than most of us actually do, why did evolution endow most individuals with brains that work to prevent faithful transmission of information through time? In other words, is there a utility to memory transience, given the seemingly obvious benefits of memory persistence?

We propose that memory transience is required in a world that is both changing and noisy. In changing environments, forgetting is adaptive because it allows for more flexible behavior. In noisy environments, forgetting is adaptive because it prevents overfitting to peculiar occurrences. According to this perspective, memory persistence is not always useful. For example, persistence of memory for aspects of the world that are either transient or uncommon would be detrimental since it might lead to inflexible behavior and/or incorrect predictions. Rather, persistence is only useful when it maintains those aspects of experience that are either relatively stable and/or predictive of new experiences. Therefore, it is only through the *interaction* of persistence and transience (persistence × transience) that memory actually serves its true purpose: using the past to intelligently guide decision-making (for related viewpoints, see Dudai and Carruthers, 2005; Schacter et al., 2007). Below, we review the computational case for using transience to increase behavioral flexibility and promote generalization. In addition, we identify the parallels between how transience is used computationally and how it appears to be implemented in the brain.

Transience for Behavioral Flexibility

New learning represents significant challenges for neural networks that use distributed representations (French, 1999; Lewandowsky and Li, 1995; McCloskey and Cohen, 1989; Ratcliff, 1990). The challenges are 2-fold. New learning might overwrite previous memories (i.e., catastrophic interference), and in turn, new learning is impeded by existing, stored memories (i.e., proactive interference) (Burgess et al., 1991; McCloskey and Cohen, 1989; Palm, 2013; Siegle and Hasselmo, 2002). This is the "stability versus plasticity" dilemma in neural networks (Abraham and Robins, 2005; Carpenter and Grossberg, 1987). As such, according to the traditional view, memory persistence is incompatible with behavioral flexibility because a network that is good at maintaining persistent memories will be poor at learning new information, especially if it conflicts with previous experiences. However, recent neural network models that use external memory devices or synapses that change over multiple timescales challenge the universality of this dilemma (Graves et al., 2016; Kirkpatrick et al., 2017; Santoro et al., 2016a). Moreover, another strategy the brain can use to solve this dilemma is to sparsely encode experiences using orthogonal representations, which may potentially arise from pattern separation processes (see Yassa and Stark, 2011 for a review). The contextual dependence of memory is one example of this strategy: by maintaining orthogonal patterns, memories that are encoded in a particular context are more likely to be expressed in that context, but not other contexts (Maren et al., 2013). This type of strategy maximizes the number of patterns that can be stored within a neural network without interference (Amari, 1989).

However, in dynamic environments it might also be important to discard outdated information regardless of any capacity constraints (Kraemer and Golding, 1997). If the environment changes, but our memories do not, then we may perseverate to our own detriment. Therefore, transience may facilitate decision-making by eliminating outdated (and potentially misleading) information, allowing an organism to respond more efficiently to changes in its environment.

Consistent with this idea, recent studies provide evidence that forgetting is necessary for flexible behavior in dynamic environments (Dong et al., 2016; Epp et al., 2016; Shuai et al., 2010).

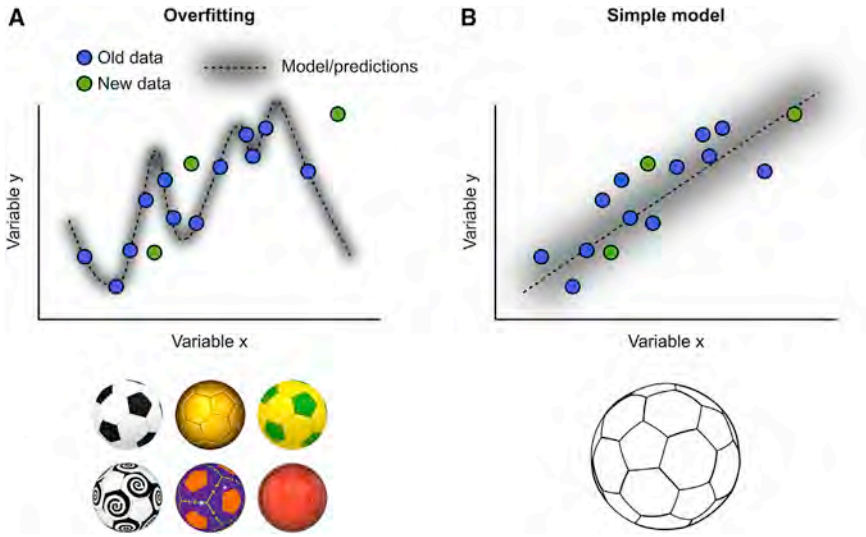

Figure 2. Avoiding Overfitting with Simple Models and Memories

(A) When performing a regression in statistics, using a function with many parameters builds a model (dotted line and shading) that fits the old data very well (blue dots), but that fails to predict new data (green dots). The mnemonic equivalent of a complex model would be to store memories for the specific patterns on every soccer ball that we have ever seen (bottom image).

(B) In contrast, using a function with few parameters builds a model that might not perfectly describe the old data, but will be better at predicting new data. The mnemonic equivalent would be to forget most details regarding soccer balls we have seen, and instead remember that they generally are made up of interlocking pentagons and hexagons (bottom image). This will lead to better prediction of the appearance of new soccer balls we encounter.

As introduced above, Shuai and colleagues trained flies to discriminate two odors (odor A, paired with shock [A+] versus odor B, not paired with shock [B−]) and found that Rac1 inhibition slowed forgetting (Shuai et al., 2010). They then asked to what extent slower forgetting would now interfere with reversal learning. Accordingly, they retrained the flies but reversed the odor-shock contingencies (i.e., A− and B+). Flies in which Rac1 was inhibited (i.e., flies displaying slower forgetting) exhibited impaired reversal learning, indicating that increased persistence of odor-shock memories interfered proactively with new learning (thereby reducing flexibility). Conversely, flies in which Rac1 was activated had the opposite phenotype. They exhibited accelerated forgetting, and this increased forgetting facilitated reversal learning (thereby increasing flexibility). This pattern of results extended to five different lines of flies engineered to express mutations linked to autism spectrum disorder that also interfere with Rac activity. All these lines of flies with disrupted Rac function exhibited impaired forgetting, and this, in turn, impaired reversal learning (Dong et al., 2016).

Further evidence for the idea that forgetting is necessary for flexible behavior in changing environments comes from rodent studies. Epp and colleagues (2016) examined reversal learning following neurogenesis-mediated forgetting. In one experiment, they trained mice in the water maze to find a platform in a fixed location. Increasing hippocampal neurogenesis after training induced forgetting of this location. Mice were subsequently retrained in the same maze, but the platform was moved to the opposite quadrant. The mice with enhanced hippocampal neurogenesis found the new platform location more efficiently (i.e., reversal learning was improved). The converse pattern was observed when hippocampal neurogenesis levels were experimentally reduced after initial water maze training. Post-training suppression of neurogenesis sustained memory for the original platform location and, in turn, interfered with learning the new reversal location.

A similar pattern of results was observed in a context-odor paired associate task (Epp et al., 2016). Increasing neurogenesis following training induced forgetting of learned paired associates (e.g., A1 and B2) but facilitated subsequent reversal learning (i.e., learning A2 and B1). Notably, this facilitation was not due to a general enhancement of learning. This was demonstrated by showing that the benefit in new learning was only observed when there was an explicit conflict with the original learning. Mice with increased neurogenesis showed no benefit when subsequently trained on novel context-odor pairs (e.g., C3 and D4). These findings suggest that adult hippocampal neurogenesis promotes forgetting, and forgetting enhances behavioral flexibility by removing or weakening outdated information. For related papers that have examined the relationship between neurogenesis and flexibility, see Burghardt et al. (2012); Garthe et al. (2009), (2016); Luu et al. (2012); Swan et al. (2014); and Winocur et al. (2012).

Transience for Regularization

In addition to promoting behavioral flexibility in changing environments, we propose that mnemonic transience is likely a means for the brain to avoid *overfitting* in noisy environments. Overfitting refers to a pernicious problem in statistics and machine learning: when overly precise models are fit to a finite dataset, it leads to inaccurate predictions due to the focus on the particulars of that dataset (Hawkins, 2004). In other words, with limited data it is easy to identify false patterns that are specific to the data, but do not generalize to new situations (Figure 2A). Tools have been developed to prevent this form of overfitting in statistics and machine learning. This usually involves restricting complexity by constraining the numbers of parameters (e.g., synapses) used to model the data (MacKay, 2003). For instance, if too many parameters are used to model the data, it is straightforward to describe all of the data. However, an "overfitted" model cannot generalize and predict new data points (MacKay, 2003). One solution is to use simple models (Figure 2B). The well-known heuristic for scientists, Occam's razor, states this in a more intuitive manner: the simpler the explanation, the broader its application.

Memories can be viewed as models of the past. By this, we mean that memories are simplified representations that capture the essence, but not necessarily the detail, of past events. This runs somewhat counter to early conceptualizations of memory as a largely reproductive process. For example, in epilepsy patients, observations that electrical stimulation of the medial temporal lobes elicited ''a record of the stream of consciousness'' (Penfield and Milner, 1958) reinforced the idea of memory as storing high-fidelity records of past events (much like a tape recorder) (Moscovitch, 2012).

Nowadays, most people recognize that memory is not a faithful reproduction of past experience. However, they might assume that this is the case because of some capacity constraint and/or flaws in processes underlying mnemonic persistence (Buonomano, 2011; Schacter, 2002). But when memories are framed as models, a memory that models the past with the simplest possible explanation is one that obeys Occam's razor. As such, simplification is not merely a side effect of constraints or flaws. Rather, it is an essential component of adaptive memory. By avoiding overfitting, simple memories will then be more successful at predicting new experiences in noisy environments. That is, simple memories that store the gist of our experiences and avoid complicated details will be better for generalizing to future events (Kumaran et al., 2016; McClelland et al., 1995; Moscovitch et al., 2016; Richards et al., 2014). One way to construct simple memories is to engage in some controlled, partial forgetting. Specifically, in order to store only the gist of an experience, statistically insignificant details must be forgotten (Sekeres et al., 2016).

Within machine learning, simplification is achieved through a process known as ''regularization.'' Regularization refers to any process used to constrain models and promote generalization (MacKay, 2003). Common regularization techniques used in artificial neural networks are ''weight decay'' (MacKay, 2003), ''sparse coding'' (Olshausen and Field, 1996, 1997, 2004), and ''drop-out'' (Srivastava et al., 2014). Here, we will focus on weight decay, and related techniques, including ''weight elimination'' (LeCun et al., 1989) and ''noise injection'' (Hinton and van Camp, 1993), as they provide the most obvious parallels to mnemonic transience in the brain. Weight decay works by reducing the strength of synaptic connections (MacKay, 2003), weight elimination works by removing synaptic connections (LeCun et al., 1989), and noise injection works by adding variability to synaptic connections (Hinton and van Camp, 1993).

Some neuroscientists might find these a surprising set of strategies for machine learning researchers to employ. Why work against the very learning that the network produces? However, if the goal is to generate simple models, then these strategies are useful. They help to constrain the complexity of the neural network, thereby restricting the total number of bits of information represented by the synaptic weights (Hinton and van Camp, 1993; MacKay, 2003). This forces neural networks to develop simple models of the data, which, in turn, improves generalization capabilities (Hinton and van Camp, 1993; LeCun et al., 1989; MacKay, 2003).

Interestingly, the addition of regularization to a learning algorithm can also be viewed as a form of Bayesian learning (MacKay, 1992). For example, in a neural network the activity patterns to be remembered may be considered as data (D), and the synaptic connections (or weights) as model parameters (W). For perfect recall, a network would select synaptic weights that maximize the probability of reactivating the stored patterns (i.e., synaptic weights that maximize the likelihood function, $P(D|W)$). In contrast, it can be shown that a network using regularization maximizes the posterior function, $P(W|D)$ (MacKay, 1992). We note that, superficially, this is a different way of viewing Bayesian learning compared to how many neuroscientists may think of it, because the parameters of the model are not things like the mean or standard deviation of a Gaussian distribution. Instead, the parameters of the model are the synaptic connections, and the prior that promotes simplification actually promotes particular synaptic arrangements (e.g., sparse connectivity). Nonetheless, the same formal rules apply, and the role of the prior is to ensure that the model does not overfit the data. In the case of memory storage, the goal is to select synaptic connections that help to recall the core features of stored patterns without focusing too heavily on the details.

These approaches for regularization resemble forms of *partial* forgetting. Although they do not lead to a complete elimination of previous learning, they can eliminate portions of previously stored information (Hinton and van Camp, 1993). Interestingly, weight decay, weight elimination, and noise injection all have recognizable analogs among the collection of neurobiological mechanisms of transience that we reviewed above (Figure 1C). Weight decay arguably corresponds to neurobiological mechanisms that weaken previously potentiated synaptic connections, including depotentiation (via endocytosis of GluA2-containing AMPARs; Hardt et al., 2013b) and synaptic downscaling during REM sleep (Tononi and Cirelli, 2006). Similarly, it is reasonable to suggest that weight elimination corresponds to neurobiological mechanisms that eliminate previously potentiated synaptic connections, including Rac-mediated spine shrinkage (Hayashi-Takagi et al., 2015) and Arp2/3 destabilization (Hadziselimovic et al., 2014). Finally, noise injection can be viewed as analogous to neurobiological mechanisms that add variability to synaptic connections, including NMDA-mediated plasticity (Villarreal et al., 2002), spinogenesis/spine turnover (Abraham et al., 2002; Attardo et al., 2015), and neurogenesis-mediated circuit remodeling (Akers et al., 2014; Kitamura et al., 2009). Therefore, both artificial neural networks and the brain appear to use comparable strategies to restrict information retention. We propose that these parallels reflect a deeper normative account of memory transience, namely, that transience is used by the brain to avoid mnemonic overfitting.

In neural networks, the outcome of minimizing overfitting is generalization. Consistent with this, a recent study showed that preventing forgetting impairs the development of generalization in rats (Migues et al., 2016). If rodents are tested shortly after contextual fear conditioning, they exhibit conditioned fear to the trained context, but not to an alternate context. However, if they are tested at longer retention delays, they exhibit conditioned fear to both the training context and the alternate context (Wiltgen and Silva, 2007; Winocur et al., 2007). This form of context generalization is a good example of how avoiding overfitting might be beneficial: when contextual fear memories generalize, an animal is no longer committed to a

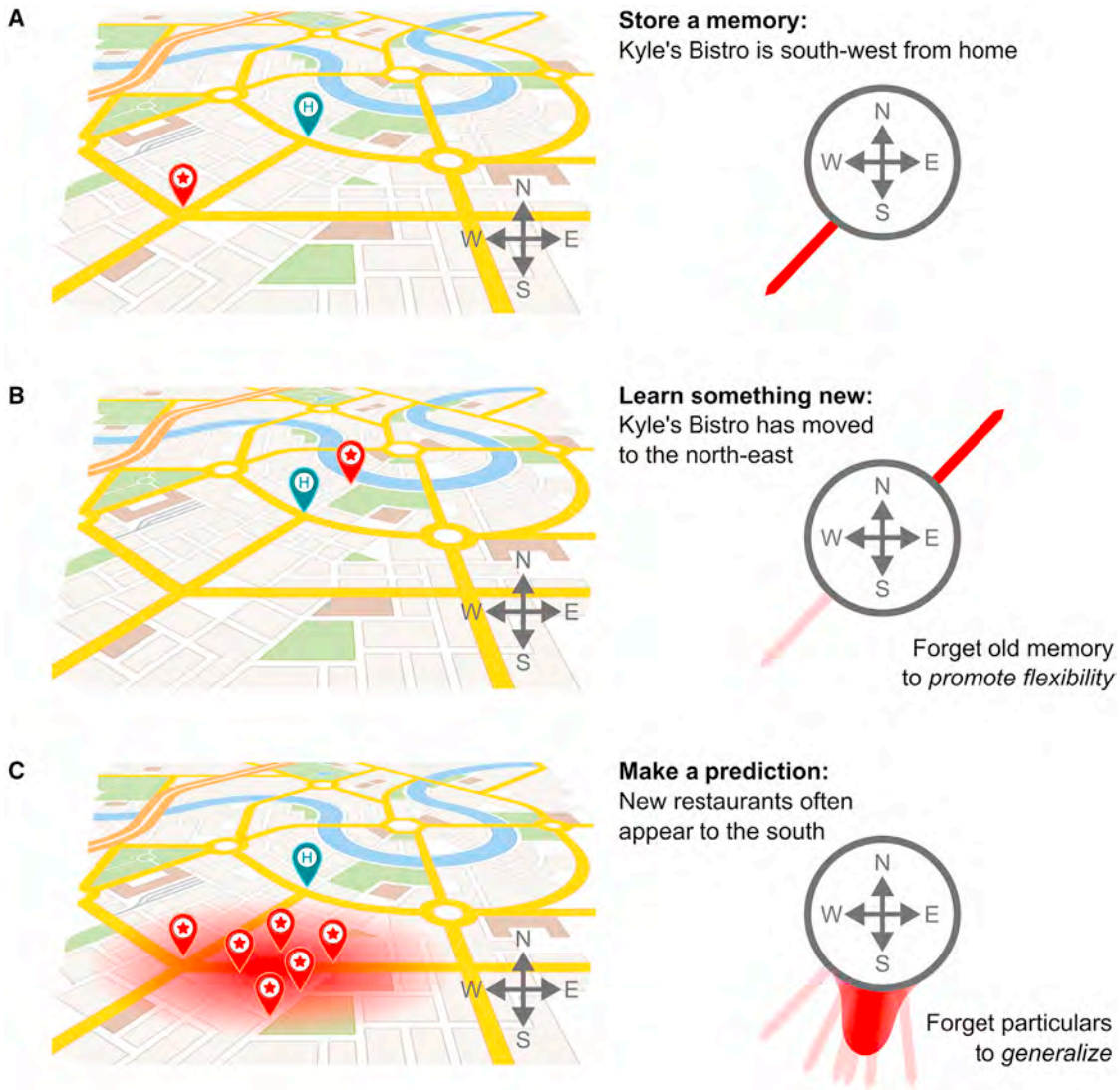

Figure 3. Mnemonic Transience Promotes Flexibility and Generalization

(A) The restaurant business is volatile. While successful restaurants may stay in the same location for many years, occasionally they move (e.g., to bigger premises). Moreover, many restaurants fail. These failed businesses are replaced by new restaurants, more often than not in the same area (e.g., neighborhoods that are high density, are walkable, and have good public transport links). Therefore, a city dweller might encode the location of their favorite eatery (e.g., Kyle's Bistro; red star), located southwest from their home (H).

(B) Transience allows flexible updating when Kyle's Bistro moves to a new location, northwest of their home.

(C) Transience also facilitates generalization, allowing the individual to predict that new restaurants will typically open up in the neighborhood south of their home.

specific set of circumstances in order to recognize danger. Migues and colleagues found that inhibiting hippocampal GluA2-containing AMPAR endocytosis following conditioning prevented the time-dependent emergence of generalization, indicating that the same mechanisms that lead to forgetting (i.e., GluA2-containing AMPAR endocytosis) also promote memory generalization (Migues et al., 2016).

Conclusion: Persistence × Transience for Optimal Decision-Making

Historically, the neurobiological study of memory has focused on how we remember rather than how we forget. However, in other traditions, most notably psychology, there has been a greater appreciation of the importance of forgetting (Rubin and Wenzel, 1996; Wimber et al., 2015). In this review, we outlined our current understanding of the neurobiological mechanisms underlying forgetting and attempted to address a broader question: what is the mnemonic benefit of transience?

Based on principles from machine learning and computational neuroscience, we proposed that in environments that change and that are noisy, transience offers two advantages for memory-guided decision-making (Figure 3). First, transience enhances behavioral flexibility by eliminating outdated information. Second, transience promotes generalization by preventing

overfitting memories to specific instances from the past that may not accurately predict the future. Other authors have made similar arguments previously (Hardt et al., 2013a; Kraemer and Golding, 1997; Nørby, 2015). However, these arguments did not explicitly discuss the computational foundations for this process, nor did they directly link these computational considerations to the neurobiology of transience.

A handful of papers have explicitly explored the advantages of transience for memory-guided decision-making (Brea et al., 2014; Fusi et al., 2007; Hunt and Chittka, 2015; Santoro et al., 2016b). Brea and colleagues modeled decision-making in flies in an associative memory paradigm (Brea et al., 2014). They found that forgetting represented the statistically optimal strategy for maximizing reward rates in dynamic environments. That is, in an environment where action-outcome contingencies change over time, it was important for an agent to engage in gradual forgetting; otherwise, behavior remained inflexible and the overall reward rate declined. They further found that tuning the rate of forgetting to the temporal dynamics of the environment maximized reward rates. When action-outcome contingencies changed frequently, faster forgetting was optimal. In contrast, when action-outcome contingencies changed infrequently, slower forgetting was best.

Fusi et al. (2007) arrived at a similar conclusion. They generated a model of a decision-making neural circuit and compared its output to psychophysical data from primates. The neural network and monkeys were engaged in a sensorimotor association task. Subjects had to learn to associate stimuli with either a leftward or rightward movement, and the associations were occasionally reversed at unexpected times. Fusi et al. (2007) found that the primates' behavior could be best described by a model that combined both a slow and fast synaptic plasticity rule. The slow learning rule led to stable long-term knowledge about the overall probability of which direction was rewarded. The fast synaptic plasticity rule allowed the model to adjust to the reversals by quickly forgetting the most recent association and returning to a stochastic baseline based on the long-term trends. Accordingly, the model both optimized its performance in the task and matched the experimental data by (1) being sensitive to the statistics of the sensorimotor associations on multiple timescales and (2) forgetting specific associations that could quickly change.

Similarly, Santoro et al. (2016b) found that a shift from precise memories to generalized memories over time enhances foraging success in noisy, changing environments (Santoro et al., 2016b). Using a neural network model with two memory systems (one for precise memories and one for general statistical patterns), they showed that the rate of reward can increase if an agent forgets precise memories and transitions to general models that have undergone regularization. This transition over time from precise memories to more general memories that average multiple instances has been observed experimentally in both mice in the water maze (Richards et al., 2014) and bees in a foraging task (Hunt and Chittka, 2015).

Interestingly, in parallel to the Brea et al. (2014) findings, Santoro et al. (2016b) found that there is an interaction between the dynamics of the environment and the ideal balance between persistence and transience. In environments where change occurs frequently, it is advantageous to rapidly shift toward generalized models, whereas in static environments where change occurs infrequently, it is advantageous to maintain specific memories for longer periods of time. Given that it is possible to encounter many different environments with different temporal dynamics, a good strategy may be to rely on multiple memory systems that have different balances between persistence and transience (Benna and Fusi, 2016; Roxin and Fusi, 2013). Indeed, there is some evidence that the emphasis on persistence versus transience varies in different mnemonic systems. For example, certain types of emotional memory stored in the amygdala may be protected from mechanisms of transience in order to enhance survival (Maren and Quirk, 2004). Moreover, there is evidence for more rapid forgetting of episodic memories (dependent on hippocampus) and slower forgetting of more general (semantic or schematized) memories (dependent on the neocortex) (Ritchey et al., 2015).Therefore, differences in the balance between persistence and transience may reflect specializations in flexible behavior versus statistical generalization (e.g., McClelland et al., 1995).

In this perspective, we have emphasized that memory should not be viewed simply as a means for high-fidelity transmission of information through time. Rather, we stressed that the goal of memory is to guide intelligent decision-making. Others have similarly discussed memory-guided decision-making within a reinforcement learning framework (Gershman and Daw, 2017). These accounts highlighted how different memory systems (e.g., model-free versus model-based versus episodic) interact in decision-making. Here we highlighted the importance of mnemonic transience. By outlining how transience can optimize memory-guided decision-making in changing and noisy environments, we emphasized how this might allow individuals to exhibit flexible behavior and generalize past events to new experiences. From this perspective, forgetting is not necessarily a failure of memory. Rather, it may represent an investment in a more optimal mnemonic strategy. With the growing literature on mnemonic transience, the time is ripe for exploring these concepts further.

ACKNOWLEDGMENTS

This review was supported by a Natural Sciences and Engineering Research Council grant (NSERC; RGPIN-2014-04947) and a 2016 Google Faculty Research Award (#2016_665) to B.A.R., and a Canadian Institute of Health Research (CIHR; FDN143227) grant to P.W.F. P.W.F. is a CIHR Canada Research Chair in Memory Research. We thank Dean Buonomano, Sheena Josselyn, and Adam Santoro for comments on earlier versions of this manuscript.

REFERENCES

Abraham, W.C., and Robins, A. (2005). Memory retention—the synaptic stability versus plasticity dilemma. Trends Neurosci. *28*, 73–78.

Abraham, W.C., Logan, B., Greenwood, J.M., and Dragunow, M. (2002). Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. J. Neurosci. *22*, 9626–9634.

Akers, K.G., Martinez-Canabal, A., Restivo, L., Yiu, A.P., De Cristofaro, A., Hsiang, H.-L.L., Wheeler, A.L., Guskjolen, A., Niibori, Y., Shoji, H., et al. (2014). Hippocampal neurogenesis regulates forgetting during adulthood and infancy. Science *344*, 598–602.

Amari, S.-I. (1989). Characteristics of sparsely encoded associative memory. Neural Netw. *2*, 451–457.

Amit, D.J., Gutfreund, H., and Sompolinsky, H. (1985). Storing infinite numbers of patterns in a spin-glass model of neural networks. Phys. Rev. Lett. *55*, 1530–1533.

Attardo, A., Fitzgerald, J.E., and Schnitzer, M.J. (2015). Impermanence of dendritic spines in live adult CA1 hippocampus. Nature *523*, 592–596.

Azevedo, F.A.C., Carvalho, L.R.B., Grinberg, L.T., Farfel, J.M., Ferretti, R.E.L., Leite, R.E.P., Jacob Filho, W., Lent, R., and Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J. Comp. Neurol. *513*, 532–541.

Bannerman, D.M., Rawlins, J.N.P., and Good, M.A. (2006). The drugs don't work—or do they? Pharmacological and transgenic studies of the contribution of NMDA and GluR-A-containing AMPA receptors to hippocampal-dependent memory. Psychopharmacology (Berl.) *188*, 552–566.

Bear, M.F. (1996). A synaptic basis for memory storage in the cerebral cortex. Proc. Natl. Acad. Sci. USA *93*, 13453–13459.

Benna, M.K., and Fusi, S. (2016). Computational principles of synaptic memory consolidation. Nat. Neurosci. *19*, 1697–1706.

Berndt, A., Lee, S.Y., Wietek, J., Ramakrishnan, C., Steinberg, E.E., Rashid, A.J., Kim, H., Park, S., Santoro, A., Frankland, P.W., et al. (2016). Structural foundations of optogenetics: Determinants of channelrhodopsin ion selectivity. Proc. Natl. Acad. Sci. USA *113*, 822–829.

Berry, J.A., and Davis, R.L. (2014). Active forgetting of olfactory memories in Drosophila. Prog. Brain Res. *208*, 39–62.

Berry, J.A., Cervantes-Sandoval, I., Nicholas, E.P., and Davis, R.L. (2012). Dopamine is required for learning and forgetting in Drosophila. Neuron *74*, 530–542.

Berry, J.A., Cervantes-Sandoval, I., Chakraborty, M., and Davis, R.L. (2015). Sleep facilitates memory by blocking dopamine neuron-mediated forgetting. Cell *161*, 1656–1667.

Blair, H.T., Schafe, G.E., Bauer, E.P., Rodrigues, S.M., and LeDoux, J.E. (2001). Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. Learn. Mem. *8*, 229–242.

Bliss, T.V., and Collingridge, G.L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. Nature *361*, 31–39.

Bliss, T.V., and Gardner-Medwin, A.R. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the unanaestetized rabbit following stimulation of the perforant path. J. Physiol. *232*, 357–374.

Bliss, T.V., and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J. Physiol. *232*, 331–356.

Brea, J., Urbanczik, R., and Senn, W. (2014). A normative theory of forgetting: lessons from the fruit fly. PLoS Comput. Biol. *10*, e1003640.

Buonomano, D. (2011). Brain Bugs: How the Brain's Flaws Shape Our Lives (W.W. Norton & Company).

Burgess, N., Shapiro, J.L., and Moore, M.A. (1991). Neural network models of list learning. Network *2*, 399–422.

Burghardt, N.S., Park, E.H., Hen, R., and Fenton, A.A. (2012). Adult-born hippocampal neurons promote cognitive flexibility in mice. Hippocampus *22*, 1795–1808.

Carpenter, G.A., and Grossberg, S. (1987). ART 2: self-organization of stable category recognition codes for analog input patterns. Appl. Opt. *26*, 4919–4930.

Cervantes-Sandoval, I., Chakraborty, M., MacMullen, C., and Davis, R.L. (2016). Scribble scaffolds a signalosome for active forgetting. Neuron *90*, 1230–1242.

Collingridge, G.L., Peineau, S., Howland, J.G., and Wang, Y.T. (2010). Long-term depression in the CNS. Nat. Rev. Neurosci. *11*, 459–473.

Cowansage, K.K., Shuman, T., Dillingham, B.C., Chang, A., Golshani, P., and Mayford, M. (2014). Direct reactivation of a coherent neocortical memory of context. Neuron *84*, 432–441.

Denny, C.A., Kheirbek, M.A., Alba, E.L., Tanaka, K.F., Brachman, R.A., Laughman, K.B., Tomm, N.K., Turi, G.F., Losonczy, A., and Hen, R. (2014). Hippocampal memory traces are differentially modulated by experience, time, and adult neurogenesis. Neuron *83*, 189–201.

Dong, Z., Han, H., Li, H., Bai, Y., Wang, W., Tu, M., Peng, Y., Zhou, L., He, W., Wu, X., et al. (2015). Long-term potentiation decay and memory loss are mediated by AMPAR endocytosis. J. Clin. Invest. *125*, 234–247.

Dong, T., He, J., Wang, S., Wang, L., Cheng, Y., and Zhong, Y. (2016). Inability to activate Rac1-dependent forgetting contributes to behavioral inflexibility in mutants of multiple autism-risk genes. Proc. Natl. Acad. Sci. USA *113*, 7644–7649.

Dudai, Y., and Carruthers, M. (2005). The Janus face of Mnemosyne. Nature *434*, 567.

Ebbinghaus, H. (1913). Memory: A Contribution to Experimental Psychology (Columbia University).

Epp, J.R., Silva Mera, R., Köhler, S., Josselyn, S.A., and Frankland, P.W. (2016). Neurogenesis-mediated forgetting minimizes proactive interference. Nat. Commun. *7*, 10838.

Frankland, P.W., Köhler, S., and Josselyn, S.A. (2013). Hippocampal neurogenesis and forgetting. Trends Neurosci. *36*, 497–503.

French, R.M. (1999). Catastrophic forgetting in connectionist networks. Trends Cogn. Sci. *3*, 128–135.

Fusi, S., Asaad, W.F., Miller, E.K., and Wang, X.-J. (2007). A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. Neuron *54*, 319–333.

Garthe, A., Behr, J., and Kempermann, G. (2009). Adult-generated hippocampal neurons allow the flexible use of spatially precise learning strategies. PLoS ONE *4*, e5464.

Garthe, A., Roeder, I., and Kempermann, G. (2016). Mice in an enriched environment learn more flexibly because of adult hippocampal neurogenesis. Hippocampus *26*, 261–271.

Gershman, S.J., and Daw, N.D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. Annu. Rev. Psychol. *68*, 101–128.

Gonçalves, J.T., Schafer, S.T., and Gage, F.H. (2016). Adult neurogenesis in the hippocampus: from stem cells to behavior. Cell *167*, 897–914.

Gould, E., McEwen, B.S., Tanapat, P., Galea, L.A., and Fuchs, E. (1997). Neurogenesis in the dentate gyrus of the adult tree shrew is regulated by psychosocial stress and NMDA receptor activation. J. Neurosci. *17*, 2492–2498.

Graves, A., Wayne, G., Reynolds, M., Harley, T., Danihelka, I., Grabska-Barwińska, A., Colmenarejo, S.G., Grefenstette, E., Ramalho, T., Agapiou, J., et al. (2016). Hybrid computing using a neural network with dynamic external memory. Nature *538*, 471–476.

Hadziselimovic, N., Vukojevic, V., Peter, F., Milnik, A., Fastenrath, M., Fenyves, B.G., Hieber, P., Demougin, P., Vogler, C., de Quervain, D.J.-F., et al. (2014). Forgetting is regulated via Musashi-mediated translational control of the Arp2/3 complex. Cell *156*, 1153–1166.

Han, J.-H., Kushner, S.A., Yiu, A.P., Hsiang, H.-L.L., Buch, T., Waisman, A., Bontempi, B., Neve, R.L., Frankland, P.W., and Josselyn, S.A. (2009). Selective erasure of a fear memory. Science *323*, 1492–1496.

Hardt, O., Nader, K., and Nadel, L. (2013a). Decay happens: the role of active forgetting in memory. Trends Cogn. Sci. *17*, 111–120.

Hardt, O., Nader, K., and Wang, Y.-T. (2013b). GluA2-dependent AMPA receptor endocytosis and the decay of early and late long-term potentiation: possible mechanisms for forgetting of short- and long-term memories. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130141.

Hawkins, D.M. (2004). The problem of overfitting. J. Chem. Inf. Comput. Sci. *44*, 1–12.

Hayashi-Takagi, A., Yagishita, S., Nakamura, M., Shirai, F., Wu, Y.I., Loshbaugh, A.L., Kuhlman, B., Hahn, K.M., and Kasai, H. (2015). Labelling and optical erasure of synaptic memory traces in the motor cortex. Nature *525*, 333–338.

Hebb, D.O. (1949). The Organization of Behavior: A Neuropsychological Theory (John Wiley & Sons).

Hinton, G.E., and van Camp, D. (1993). Keeping the neural networks simple by minimizing the description length of the weights. In Proceedings of the Sixth Annual Conference on Computational Learning Theory—COLT '93

Hodge, R.G., and Ridley, A.J. (2016). Regulating Rho GTPases and their regulators. Nat. Rev. Mol. Cell Biol. *17*, 496–510.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA *79*, 2554–2558.

Hsiang, H.-L.L., Epp, J.R., van den Oever, M.C., Yan, C., Rashid, A.J., Insel, N., Ye, L., Niibori, Y., Deisseroth, K., Frankland, P.W., and Josselyn, S.A. (2014). Manipulating a "cocaine engram" in mice. J. Neurosci. *34*, 14115–14127.

Hunt, K.L., and Chittka, L. (2015). Merging of long-term memories in an insect. Curr. Biol. *25*, 741–745.

Inda, M.C., Muravieva, E.V., and Alberini, C.M. (2011). Memory retrieval and the passage of time: from reconsolidation and strengthening to extinction. J. Neurosci. *31*, 1635–1643.

Ishikawa, R., Fukushima, H., Frankland, P.W., and Kida, S. (2016). Hippocampal neurogenesis enhancers promote forgetting of remote fear memory after hippocampal reactivation by retrieval. eLife *5*, e17464.

Jiang, L., Mao, R., Zhou, Q., Yang, Y., Cao, J., Ding, Y., Yang, Y., Zhang, X., Li, L., and Xu, L. (2016). Inhibition of Rac1 activity in the hippocampus impairs the forgetting of contextual fear memory. Mol. Neurobiol. *53*, 1247–1253.

Josselyn, S.A., Köhler, S., and Frankland, P.W. (2015). Finding the engram. Nat. Rev. Neurosci. *16*, 521–534.

Kandel, E.R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. Science *294*, 1030–1038.

Kemp, A., and Manahan-Vaughan, D. (2007). Hippocampal long-term depression: master or minion in declarative memory processes? Trends Neurosci. *30*, 111–118.

Kempermann, G., Kuhn, H.G., and Gage, F.H. (1997). More hippocampal neurons in adult mice living in an enriched environment. Nature *386*, 493–495.

Kim, J., Kwon, J.-T., Kim, H.-S., Josselyn, S.A., and Han, J.-H. (2014). Memory recall and modifications by activating neurons with elevated CREB. Nat. Neurosci. *17*, 65–72.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., et al. (2017). Overcoming catastrophic forgetting in neural networks. Proc. Natl. Acad. Sci. USA *114*, 3521–3526.

Kitamura, T., Saitoh, Y., Takashima, N., Murayama, A., Niibori, Y., Ageta, H., Sekiguchi, M., Sugiyama, H., and Inokuchi, K. (2009). Adult neurogenesis modulates the hippocampus-dependent period of associative fear memory. Cell *139*, 814–827.

Kraemer, P.J., and Golding, J.M. (1997). Adaptive forgetting in animals. Psychon. Bull. Rev. *4*, 480–491.

Kumaran, D., Hassabis, D., and McClelland, J.L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. Trends Cogn. Sci. *20*, 512–534.

LeCun, Y., Denker, J.S., and Solla, S.A. (1989). Optimal brain damage. In Advances in Neural Information Processing Systems 2, R.E. Howard and L.D. Jackel, eds., pp. 598–605.

Lewandowsky, S., and Li, S.-C. (1995). Catastrophic interference in neural networks. In Interference and Inhibition in Cognition, F.N. Dempster and C.J. Brainerd, eds. (Elsevier), pp. 329–361.

Ling, D.S.F., Benardo, L.S., Serrano, P.A., Blace, N., Kelly, M.T., Crary, J.F., and Sacktor, T.C. (2002). Protein kinase Mzeta is necessary and sufficient for LTP maintenance. Nat. Neurosci. *5*, 295–296.

Ling, D.S.F., Benardo, L.S., and Sacktor, T.C. (2006). Protein kinase Mzeta enhances excitatory synaptic transmission by increasing the number of active postsynaptic AMPA receptors. Hippocampus *16*, 443–452.

Liu, X., Ramirez, S., and Tonegawa, S. (2013). Inception of a false memory by optogenetic manipulation of a hippocampal memory engram. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130142.

Liu, Y., Du, S., Lv, L., Lei, B., Shi, W., Tang, Y., Wang, L., and Zhong, Y. (2016). Hippocampal activation of Rac1 regulates the forgetting of object recognition memory. Curr. Biol. *26*, 2351–2357.

Luria, A.R. (1968). The Mind of a Mnemonist: A Little Book about a Vast Memory (Harvard University Press).

Lüscher, C., and Malenka, R.C. (2012). NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). Cold Spring Harb. Perspect. Biol. *4*, a005710.

Luu, P., Sill, O.C., Gao, L., Becker, S., Wojtowicz, J.M., and Smith, D.M. (2012). The role of adult hippocampal neurogenesis in reducing interference. Behav. Neurosci. *126*, 381–391.

MacKay, D.J.C. (1992). Bayesian interpolation. In Maximum Entropy and Bayesian Methods, K.M. Hanson and R.N. Silver, eds. (Springer), pp. 39–66.

MacKay, D.J.C. (2003). Information Theory, Inference and Learning Algorithms (Cambridge University Press).

Malenka, R.C., and Bear, M.F. (2004). LTP and LTD: an embarrassment of riches. Neuron *44*, 5–21.

Maren, S., and Quirk, G.J. (2004). Neuronal signalling of fear memory. Nat. Rev. Neurosci. *5*, 844–852.

Maren, S., Phan, K.L., and Liberzon, I. (2013). The contextual brain: implications for fear conditioning, extinction and psychopathology. Nat. Rev. Neurosci. *14*, 417–428.

McAvoy, K.M., Scobie, K.N., Berger, S., Russo, C., Guo, N., Decharatanachart, P., Vega-Ramirez, H., Miake-Lye, S., Whalen, M., Nelson, M., et al. (2016). Modulating neuronal competition dynamics in the dentate gyrus to rejuvenate aging memory circuits. Neuron *91*, 1356–1373.

McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. *102*, 419–457.

McCloskey, M., and Cohen, N.J. (1989). Catastrophic interference in connectionist networks: the sequential learning problem. Psychol. Learn. Motiv. *24*, 109–165.

McGaugh, J.L. (2000). Memory—a century of consolidation. Science *287*, 248–251.

McKernan, M.G., and Shinnick-Gallagher, P. (1997). Fear conditioning induces a lasting potentiation of synaptic currents in vitro. Nature *390*, 607–611.

Migues, P.V., Hardt, O., Wu, D.C., Gamache, K., Sacktor, T.C., Wang, Y.T., and Nader, K. (2010). PKMzeta maintains memories by regulating GluR2-dependent AMPA receptor trafficking. Nat. Neurosci. *13*, 630–634.

Migues, P.V., Liu, L., Archbold, G.E.B., Einarsson, E.Ö., Wong, J., Bonasia, K., Ko, S.H., Wang, Y.T., and Hardt, O. (2016). Blocking synaptic removal of GluA2-containing AMPA receptors prevents the natural forgetting of long-term memories. J. Neurosci. *36*, 3481–3494.

Morris, R.G., Anderson, E., Lynch, G.S., and Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature *319*, 774–776.

Moscovitch, M. (2012). Memory before and after H.M.: an impressionistic historical perspective. In Epilepsy and Memory, A. Zeman, N. Kapur, and M. Jones-Gotman, eds. (Oxford University Press), pp. 19–50.

Moscovitch, M., Cabeza, R., Winocur, G., and Nadel, L. (2016). Episodic memory and beyond: the hippocampus and neocortex in transformation. Annu. Rev. Psychol. *67*, 105–134.

Nabavi, S., Fox, R., Proulx, C.D., Lin, J.Y., Tsien, R.Y., and Malinow, R. (2014). Engineering a memory with LTD and LTP. Nature *511*, 348–352.

Nørby, S. (2015). Why forget? On the adaptive value of memory loss. Perspect. Psychol. Sci. *10*, 551–578.

O'Dell, T.J., and Kandel, E.R. (1994). Low-frequency stimulation erases LTP through an NMDA receptor-mediated activation of protein phosphatases. Learn. Mem. *1*, 129–139.

Ohkawa, N., Saitoh, Y., Suzuki, A., Tsujimura, S., Murayama, E., Kosugi, S., Nishizono, H., Matsuo, M., Takahashi, Y., Nagase, M., et al. (2015). Artificial association of pre-stored information to generate a qualitatively new memory. Cell Rep. *11*, 261–269.

Olshausen, B.A., and Field, D.J. (1996). Natural image statistics and efficient coding. Network *7*, 333–339.

Olshausen, B.A., and Field, D.J. (1997). Sparse coding with an overcomplete basis set: a strategy employed by V1? Vision Res. *37*, 3311–3325.

Olshausen, B.A., and Field, D.J. (2004). Sparse coding of sensory inputs. Curr. Opin. Neurobiol. *14*, 481–487.

Palm, G. (2013). Neural associative memories and sparse coding. Neural Netw. *37*, 165–171.

Park, S., Kramer, E.E., Mercaldo, V., Rashid, A.J., Insel, N., Frankland, P.W., and Josselyn, S.A. (2016). Neuronal allocation to a hippocampal engram. Neuropsychopharmacology *41*, 2987–2993.

Pastalkova, E., Serrano, P., Pinkhasova, D., Wallace, E., Fenton, A.A., and Sacktor, T.C. (2006). Storage of spatial information by the maintenance mechanism of LTP. Science *313*, 1141–1144.

Penfield, W., and Milner, B. (1958). Memory deficit produced by bilateral lesions in the hippocampal zone. AMA Arch. Neurol. Psychiatry *79*, 475–497.

Poo, M.-M., Pignatelli, M., Ryan, T.J., Tonegawa, S., Bonhoeffer, T., Martin, K.C., Rudenko, A., Tsai, L.-H., Tsien, R.W., Fishell, G., et al. (2016). What is memory? The present state of the engram. BMC Biol. *14*, 40.

Ramirez, S., Liu, X., Lin, P.-A., Suh, J., Pignatelli, M., Redondo, R.L., Ryan, T.J., and Tonegawa, S. (2013). Creating a false memory in the hippocampus. Science *341*, 387–391.

Rashid, A.J., Yan, C., Mercaldo, V., Hsiang, H.-L.L., Park, S., Cole, C.J., De Cristofaro, A., Yu, J., Ramakrishnan, C., Lee, S.Y., et al. (2016). Competition between engrams influences fear memory formation and recall. Science *353*, 383–387.

Ratcliff, R. (1990). Connectionist models of recognition memory: constraints imposed by learning and forgetting functions. Psychol. Rev. *97*, 285–308.

Reijmers, L.G., Perkins, B.L., Matsuo, N., and Mayford, M. (2007). Localization of a stable neural correlate of associative memory. Science *317*, 1230–1233.

Reynolds, B.A., and Weiss, S. (1992). Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. Science *255*, 1707–1710.

Richards, B.A., and Frankland, P.W. (2013). The conjunctive trace. Hippocampus *23*, 207–212.

Richards, B.A., Xia, F., Santoro, A., Husse, J., Woodin, M.A., Josselyn, S.A., and Frankland, P.W. (2014). Patterns across multiple memories are identified over time. Nat. Neurosci. *17*, 981–986.

Ritchey, M., Montchal, M.E., Yonelinas, A.P., and Ranganath, C. (2015). Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. eLife *4*, http://dx.doi.org/10.7554/eLife.05025.

Rogan, M.T., Stäubli, U.V., and LeDoux, J.E. (1997). Fear conditioning induces associative long-term potentiation in the amygdala. Nature *390*, 604–607.

Rogerson, T., Jayaprakash, B., Cai, D.J., Sano, Y., Lee, Y.-S., Zhou, Y., Bekal, P., Deisseroth, K., and Silva, A.J. (2016). Molecular and cellular mechanisms for trapping and activating emotional memories. PLoS ONE *11*, e0161655.

Roxin, A., and Fusi, S. (2013). Efficient partitioning of memory systems and its importance for memory consolidation. PLoS Comput. Biol. *9*, e1003146.

Rubin, D.C., and Wenzel, A.E. (1996). One hundred years of forgetting: a quantitative description of retention. Psychol. Rev. *103*, 734–760.

Sachser, R.M., Santana, F., Crestani, A.P., Lunardi, P., Pedraza, L.K., Quillfeldt, J.A., Hardt, O., and Alvares, L. de O. (2016). Forgetting of long-term memory requires activation of NMDA receptors, L-type voltage-dependent Ca2+ channels, and calcineurin. Sci. Rep. *6*, 22771.

Sacktor, T.C. (2011). How does PKMζ maintain long-term memory? Nat. Rev. Neurosci. *12*, 9–15.

Sanes, J.R., and Lichtman, J.W. (1999). Can molecules explain long-term potentiation? Nat. Neurosci. *2*, 597–604.

Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., and Lillicrap, T. (2016a). Meta-learning with memory-augmented neural networks. Proceedings of the 33rd International Conference on Machine Learning, pp. 1842–1850.

Santoro, A., Frankland, P.W., and Richards, B.A. (2016b). Memory transformation enhances reinforcement learning in dynamic environments. J. Neurosci. *36*, 12228–12242.

Schacter, D.L. (2002). The Seven Sins of Memory: How the Mind Forgets and Remembers (Houghton Mifflin Harcourt).

Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. Nat. Rev. Neurosci. *8*, 657–661.

Sekeres, M.J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., and Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: differential rates of forgetting for detail types in episodic memory. Learn. Mem. *23*, 72–82.

Serrano, P., Friedman, E.L., Kenney, J., Taubenfeld, S.M., Zimmerman, J.M., Hanna, J., Alberini, C., Kelley, A.E., Maren, S., Rudy, J.W., et al. (2008). PKMzeta maintains spatial, instrumental, and classically conditioned long-term memories. PLoS Biol. *6*, 2698–2706.

Shuai, Y., Lu, B., Hu, Y., Wang, L., Sun, K., and Zhong, Y. (2010). Forgetting is regulated through Rac activity in Drosophila. Cell *140*, 579–589.

Siegle, G.J., and Hasselmo, M.E. (2002). Using connectionist models to guide assessment of psychological disorder. Psychol. Assess. *14*, 263–278.

Silva, A.J., Paylor, R., Wehner, J.M., and Tonegawa, S. (1992). Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. Science *257*, 206–211.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. *15*, 1929–1958.

Stevens, C.F. (1998). A million dollar question: does LTP = memory? Neuron *20*, 1–2.

Stickgold, R., and Walker, M.P. (2013). Sleep-dependent memory triage: evolving generalization through selective processing. Nat. Neurosci. *16*, 139–145.

Swan, A.A., Clutton, J.E., Chary, P.K., Cook, S.G., Liu, G.G., and Drew, M.R. (2014). Characterization of the role of adult neurogenesis in touch-screen discrimination learning. Hippocampus *24*, 1581–1591.

Tanaka, K.Z., Pevzner, A., Hamidi, A.B., Nakazawa, Y., Graham, J., and Wiltgen, B.J. (2014). Cortical representations are reinstated by the hippocampus during memory retrieval. Neuron *84*, 347–354.

Tonegawa, S., Liu, X., Ramirez, S., and Redondo, R. (2015). Memory engram cells have come of age. Neuron *87*, 918–931.

Toni, N., Teng, E.M., Bushong, E.A., Aimone, J.B., Zhao, C., Consiglio, A., van Praag, H., Martone, M.E., Ellisman, M.H., and Gage, F.H. (2007). Synapse formation on neurons born in the adult hippocampus. Nat. Neurosci. *10*, 727–734.

Toni, N., Laplagne, D.A., Zhao, C., Lombardi, G., Ribak, C.E., Gage, F.H., and Schinder, A.F. (2008). Neurons born in the adult dentate gyrus form functional synapses with target cells. Nat. Neurosci. *11*, 901–907.

Tononi, G., and Cirelli, C. (2006). Sleep function and synaptic homeostasis. Sleep Med. Rev. *10*, 49–62.

Tsokas, P., Hsieh, C., Yao, Y., Lesburguères, E., Wallace, E.J.C., Tcherepanov, A., Jothianandan, D., Hartley, B.R., Pan, L., Rivard, B., et al. (2016). Compensation for PKMζ in long-term potentiation and spatial long-term memory in mutant mice. eLife *5*, e14846.

Tulving, E. (2002). Episodic memory: from mind to brain. Annu. Rev. Psychol. *53*, 1–25.

van Praag, H., Kempermann, G., and Gage, F.H. (1999). Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. Nat. Neurosci. *2*, 266–270.

Villarreal, D.M., Do, V., Haddad, E., and Derrick, B.E. (2002). NMDA receptor antagonists sustain LTP and spatial memory: active processes mediate LTP decay. Nat. Neurosci. *5*, 48–52.

Whitlock, J.R., Heynen, A.J., Shuler, M.G., and Bear, M.F. (2006). Learning induces long-term potentiation in the hippocampus. Science *313*, 1093–1097.

Wiltgen, B.J., and Silva, A.J. (2007). Memory for context becomes less specific with time. Learn. Mem. *14*, 313–317.

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., and Anderson, M.C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. Nat. Neurosci. *18*, 582–589.

Winocur, G., Moscovitch, M., and Sekeres, M. (2007). Memory consolidation or transformation: context manipulation and hippocampal representations of memory. Nat. Neurosci. *10*, 555–557.

Winocur, G., Becker, S., Luu, P., Rosenzweig, S., and Wojtowicz, J.M. (2012). Adult hippocampal neurogenesis and memory interference. Behav. Brain Res. *227*, 464–469.

Yao, Y., Kelly, M.T., Sajikumar, S., Serrano, P., Tian, D., Bergold, P.J., Frey, J.U., and Sacktor, T.C. (2008). PKM zeta maintains late long-term potentiation by N-ethylmaleimide-sensitive factor/GluR2-dependent trafficking of postsynaptic AMPA receptors. J. Neurosci. *28*, 7820–7827.

Yassa, M.A., and Stark, C.E.L. (2011). Pattern separation in the hippocampus. Trends Neurosci. *34*, 515–525.

Yiu, A.P., Mercaldo, V., Yan, C., Richards, B., Rashid, A.J., Hsiang, H.-L.L., Pressey, J., Mahadevan, V., Tran, M.M., Kushner, S.A., et al. (2014). Neurons are recruited to a memory trace based on relative neuronal excitability immediately before training. Neuron *83*, 722–735.

Zhou, Y., Won, J., Karlsson, M.G., Zhou, M., Rogerson, T., Balaji, J., Neve, R., Poirazi, P., and Silva, A.J. (2009). CREB regulates excitability and the allocation of memory to subsets of neurons in the amygdala. Nat. Neurosci. *12*, 1438–1443.