

The molecular memory code and synaptic plasticity: a synthesis

Samuel J. Gershman

Department of Psychology and Center for Brain Science, Harvard University
Center for Brains, Minds and Machines, MIT
Correspondence: gershman@fas.harvard.edu

November 11, 2021

Abstract

The most widely accepted view of memory in the brain holds that synapses are the storage sites of memory, and that memories are formed through associative modification of synapses. This view has been challenged on conceptual and empirical grounds. As an alternative, it has been proposed that molecules within the cell body are the storage sites of memory, and that memories are formed through biochemical operations on these molecules. This paper proposes a synthesis of these two views, grounded in a computational theory of memory. Synapses are conceived as storage sites for the parameters of an approximate posterior probability distribution over latent causes. Intracellular molecules are conceived as storage sites for the parameters of a generative model. The theory stipulates how these two components work together as part of an integrated algorithm for learning and inference.

Keywords: memory, free energy, synaptic plasticity, learning, inference

Contents

1	Introduction	3
2	Empirical and conceptual issues	4
2.1	Synaptic plasticity and memory	4
2.2	The trouble with synapses	6
2.2.1	Timescales	6
2.2.2	Selectivity	7
2.2.3	Content	8
2.2.4	Memory persistence and synaptic instability	10
2.2.5	Savings	12
2.2.6	Dissociations between memory and synaptic integrity	13
2.2.7	Intrinsic plasticity: from excitability to coding	14
2.2.8	Memory transfer	15
2.2.9	Learning in single cells	18
3	In search of the memory molecule	19
3.1	RNA and other cytoplasmic macromolecules	19
3.2	Nuclear mechanisms: DNA methylation and histone modification	22
4	A new theoretical perspective	23
4.1	Two computational problems: inference and learning	23
4.2	An algorithmic solution	24
4.2.1	Minimizing free energy by gradient descent	25
4.2.2	Informativeness and associability	25
4.3	Biologically plausible neural implementation	26
4.3.1	Circuit model	26
4.3.2	Connectivity	27
4.3.3	Plasticity rules	28
4.3.4	Querying the generative model	28
4.4	Summary and implications	29
5	Insights from cognitive science	29
6	Conclusions	31

“The lack of interest of neurophysiologists in the macromolecular theory of memory can be accounted for by recognizing that the theory, whether true or false, is clearly premature. There is no chain of reasonable inferences by means of which our present, albeit highly imperfect, view of the functional organization of the brain can be reconciled with the possibility of its acquiring, storing and retrieving nervous information by encoding such information in molecules of nucleic acid or protein.”

Stent (1972), *Prematurity and uniqueness in scientific discovery*

1 Introduction

Understanding the biological basis of memory is one of the most profound puzzles in neuroscience. Solving this puzzle requires answers to two questions: (i) What is the *content* of memory, and (ii) what is the *structure* of memory? Content refers to the information encoded in memory. Structure refers to the code that maps information into a physical trace (or “engram”). There are standard textbook answers to these questions: The content of memory consists of associations between events; the code is a mapping from associations to synaptic strengths (Martin et al., 2000; Takeuchi et al., 2014).

There are major problems with these textbook answers, which I will summarize below. An alternative hypothesis, originally proposed in the mid 20th century (see Gaito, 1976, for a review) and recently revived (Gallistel, 2017; Abraham et al., 2019), postulates that memories are encoded in an intracellular molecular substrate. The precise nature of this substrate is speculative and a matter of debate. Gallistel has argued that the content of memory consists of “facts” (explicit representations of variables in the environment) rather than associations (Gallistel and King, 2011; Gallistel, 2017, 2021). These facts are read out by the spiking activity of neurons (presumably mediated by some biochemical process within the cell) and thereby made computationally accessible to downstream neurons. Synaptic plasticity, on this view, plays no role in memory storage.

As we will see, the molecular hypothesis has many merits. But if it turns out to be true, we are left with a residual puzzle: What is the function of synaptic plasticity? Note that Gallistel never denies the *existence* of synaptic plasticity, only its functional implications. However, it would be odd if the brain carried out functionally inert computations that may demand on the order of 10% of the total cortical metabolic rate (Karbowski, 2019). Moreover, synaptic plasticity is *not* functionally inert, since it appears to be causally related to learned behavior (reviewed by Takeuchi et al., 2014, and discussed further below).

I will develop new answers to these questions, starting with a distinction between two kinds of memory, one synaptic and one intracellular, that serve different computational functions. Following Gallistel, I propose that an intracellular molecular mechanism stores “facts”—in particular, the parameters of a *generative model* (a concept I elaborate below). To do some useful computation with these facts (e.g., perception, prediction, action), the brain needs to infer the latent causes of sensory observations by inverting the generative model. This is generally intractable for complex generative models, but can be approximated using a parametrized mapping from observations to a distribution over latent causes—the *inference model*.¹ I hypothesize that: (i) this inference model

¹The inference model is sometimes referred to as a *recognition model* (Dayan et al., 1995; Hinton and Zemel, 1994; Kingma and Welling, 2013) or *inference network* (Mnih and Gregor, 2014; Rezende and Mohamed, 2015). In the machine learning literature, the inference model is typically implemented as a deep neural network.

is implemented by spiking activity in a network of neurons; (ii) its parameters are stored at the synapse; and (iii) synaptic plasticity updates these parameters to optimize the inference model.

Importantly, the generative model and the inference model fit together as part of a single optimization problem, namely minimizing free energy. This is already a familiar idea in theoretical neuroscience (e.g., Friston et al., 2006; Friston, 2010; Gershman, 2019). The most biologically detailed implementations of free energy minimization rest upon a predictive coding scheme in which feedforward pathways (neurons in superficial cortical layers) convey prediction errors and feedback pathways (neurons in deep cortical layers) convey predictions (Bastos et al., 2012; Bogacz, 2017; Friston, 2005). Both generative and inference parameters are learned via Hebbian synaptic plasticity rules. Thus, the main departure explored here is the idea that the generative model is implemented by intracellular computations, and learning the parameters of the generative model is implemented by molecular plasticity within cells rather than at the synapse.²

To set the stage for my theory, I begin by surveying the empirical and conceptual issues with synaptic plasticity as a memory mechanism. This motivates a consideration of alternative (or complementary) mechanisms, in particular those based on intracellular molecular substrates. To understand how these synaptic and intracellular mechanisms fit together, I embark on an abstract theoretical treatment of the optimization problem facing the brain. I show how the optimization problem can be tractably (albeit approximately) solved, and how this solution can be implemented in a biologically plausible system. Finally, I discuss how this proposal reconciles the divergent views of memory in the brain.

2 Empirical and conceptual issues

How did we get to the current consensus that synaptic plasticity is the basis of memory formation and storage? While there exists a massive literature on synaptic plasticity from physiological, cellular, and molecular perspectives, here I will focus on the most compelling evidence that synaptic plasticity plays a role in the behavioral expression of memory. I will then critically evaluate this body of evidence.

2.1 Synaptic plasticity and memory

The classical way to study synaptic plasticity is to stimulate a presynaptic axon with a high-frequency tetanus and measure the size of evoked excitatory postsynaptic potentials (EPSPs). Bliss and Lømo (1973) first reported that this procedure resulted in a long-lasting increase in EPSP amplitude—a phenomenon known as *long-term potentiation* (LTP). Later it was discovered that low-frequency stimulation reduces EPSP amplitude (Dudek and Bear, 1992)—a phenomenon known as *long term depression* (LTD). These findings were remarkable because they fit conveniently into an associative conceptualization of memory that dates back to Plato and Aristotle, and which was later championed by the British empiricist philosophers. In the 20th century, associative memory was propelled to the forefront of experimental psychology. Donald Hebb is usually credited with the hypothesis that the strengthening of connections between neurons might be the basis of memory formation (“neurons that fire together, wire together”; Hebb, 1949), though William

²It is also worth noting that predictive coding can only be applied to a restricted class of models, and also entails a Gaussian approximation of the posterior (see Gershman, 2019, for further discussion).

James proposed essentially the same idea (in terms of “organic materials” rather than neurons) more than half a century earlier (James, 1890).

By now, these ideas have become so entrenched that it’s difficult for neuroscientists to talk about memory without association. It has become a matter of *definition* rather than *theory*. For the moment, I will defer a critique of this definitional assumption. Taking it at face value, what is the evidence that LTP and LTD are causally responsible for memory as measured behaviorally? Few studies have addressed this question directly, despite our wealth of knowledge about LTP and LTD.

Some early attempts to address this question interfered with various cellular components that were considered to be crucial for synaptic plasticity. For example, NMDA receptors are necessary for LTP, and antagonizing these receptors in the hippocampus impairs learning in a hippocampus-dependent place learning task (the Morris water maze; Morris et al., 1986). Unfortunately, this study, and many similar ones, suffers from a logical flaw: NMDA receptors (like many other components involved in synaptic plasticity) play multiple roles in the cell, so we do not know if the impairment is due specifically to blocked plasticity or to some other dysfunction in the cell (Shors and Matzel, 1997; Nicoll, 2017). Moreover, subsequent work showed that animals can produce intact performance in the Morris water maze despite NMDA antagonism, as long as they are given some pretraining (Bannerman et al., 1995; Saucier and Cain, 1995). Thus, NMDA receptors do not appear to be a necessary component of learning, at least under certain conditions. I will discuss other examples of dissociations between LTP and learning in a later section.

Other studies have asked whether changes in synaptic strength are accompanied by changes in behavior. Whitlock et al. (2006) showed that inhibitory avoidance learning produced LTP at hippocampal synapses, accompanied by trafficking of AMPA receptors to the postsynaptic membrane. In the same spirit, Rogan et al. (1997) showed that fear conditioning produces LTP in the lateral amygdala, paralleling changes in cue-evoked fear responses (see also McKernan and Shinnick-Gallagher, 1997). Using an auditory frequency discrimination task, Xiong et al. (2015) showed that corticostriatal synapses were potentiated specifically for cortical neurons tuned to frequencies associated with reward, paralleling changes in discrimination performance.

Later studies provided causal evidence that trafficking of AMPA receptors to the postsynaptic membrane of lateral amygdala neurons is necessary for acquiring a fear memory (Rumpel et al., 2005), and endocytosis of AMPA receptors is necessary for the extinction of fear memory (Kim et al., 2007; Clem and Haganir, 2010). Using optogenetic induction of LTP and LTD at amygdala synapses, Nabavi et al. (2014) were able to create, inactivate, and reactivate a fear memory. In motor cortex, optical erasure of dendritic spines formed during learning caused a selective disruption of the corresponding memory (Hayashi-Takagi et al., 2015).

While these studies are suggestive, they suffer from another logical flaw: they do not tell us whether synaptic plasticity is necessary for the *storage* of memory or for its *expression*. If memories were stored in an intracellular format but required synapses to be expressed behaviorally (e.g., through conditioned responding), then changes in synapses should track changes in behavioral expression, even though the synapses are not the locus of information storage. *In vitro* LTP typically decays to baseline within a few hours (longer durations have been measured *in vivo*, e.g., Abraham et al., 2002). This implies that any behavioral memory expression that outlasts this time window is not mediated by storage at the synapse in question. I will discuss the issue of long-term storage further below. For now, the important point is that the evidence for synaptic storage of information is quite weak; in contrast, it is clear that synapses play an important role in memory

expression.

2.2 The trouble with synapses

Although synapses are widely viewed as the site of memory storage, and synaptic plasticity is widely viewed as the basis of memory formation, the previous section has already indicated that decisive evidence for these views is lacking. In this section, I undertake a systematic critique of the synaptic model of memory.

2.2.1 Timescales

Hebbian rules for plasticity imply a temporal constraint: the presynaptic and postsynaptic neurons must fire close in time in order to produce a change in synaptic strength. How close? This question was answered by studies that measured LTP and LTD as a function of relative spike timing, finding that outside a window of about 40-60 milliseconds no plasticity occurs (Markram et al., 1997; Bi and Poo, 1998). This is known as *spike-timing-dependent plasticity* (STDP).

This temporal constraint presents a serious problem, since animals can easily learn associations between stimuli separated by much longer intervals. The challenge then is to explain how synaptic plasticity can produce learning at behaviorally relevant timescales. One answer to this challenge involves neurons that generate sustained responses. If a stimulus can produce an “eligibility trace” that outlasts the stimulus, then a conventional STDP rule can still produce an association (Drew and Abbott, 2006). For example, calcium plateau potentials in apical dendrites could play this role (Bittner et al., 2017). Alternatively, a neuromodulator with slower kinetics, such as a dopamine, might provide the temporal bridge between stimuli (Izhikevich, 2007; Gerstner et al., 2018).

These mechanisms can extend the timescale of synaptic plasticity from milliseconds to seconds, but they cannot explain how learning is possible over longer timescales. Animals can acquire conditioned taste aversion when sickness follows ingestion by several hours (Smith and Roll, 1967; Revusky, 1968). In this case, one might argue that a long-lasting eligibility trace exists in the digestive system—the so-called “aftertaste” hypothesis. If true, this would allow a conventional STDP mechanism to explain taste aversion learning over very long delays. However, considerable evidence opposes the aftertaste hypothesis (see Rozin and Kalat, 1971, for a review). Taste aversions can be established under physiological conditions in which a long-lasting trace is highly implausible. For example, only a single conditioning trial with a saccharin solution is needed to produce taste aversion, even when the delay to sickness is 12 hours, at which point the ingested saccharin has a negligible presence in the blood stream and digestive system. Aversion can also be learned to the high or low temperature of a drink, despite the fact that the fluid’s temperature is quickly converted to body temperature after ingestion (Nachman, 1970). Another argument against the aftertaste hypothesis rests on the assumption that whatever the trace is, it must decay over time. Indeed, the strength of conditioned taste aversion declines with delay, consistent with other Pavlovian conditioning preparations. Presenting the conditioned stimulus again should presumably “reactivate” the trace, thereby rendering taste aversion stronger. In fact, the opposite happens: taste aversion is weakened, possibly due to a “learned safety” effect (Kalat and Rozin, 1973).

This is not necessarily the end of the story, since an eligibility trace could be maintained neurally even in the absence of a peripheral trace. For delays of up to 3 hours, conditioned taste aversion depends on sustained phosphorylation of calmodulin-dependent protein kinase II in the

insular cortex (Adaikkan and Rosenblum, 2015). The existence of such a trace, however, does not imply that conditioned taste aversion is mediated by STDP. It is unclear how such a mechanism could achieve the selectivity of learning observed in conditioned taste aversion (a point I discuss more in the next section). Another issue is that a very slowly decaying trace would have to terminate relatively abruptly after the aversive outcome in order to avoid producing LTD that would partially negate the antecedent LTP. Such precise (millisecond) timing for a trace that is assumed to be governed by kinetics operating on the timescale of hours presents a biophysical conundrum.

I have not yet confronted an even more fundamental timescale problem: *there is no characteristic timescale for learning*. While it is true that associative learning is weaker when two stimuli are separated by a longer delay, the functionally relevant units of time are not absolute but relative. To understand this, consider the fact that spacing trials increases the rate of learning. Putting the delay-dependence of learning together with the spacing effect, we obtain the timescale invariance of learning: learning rate is typically constant across inter-stimulus delays as long as the inter-stimulus interval is proportionally rescaled (Gibbon et al., 1977; Ward et al., 2012; Gallistel and Gibbon, 2000). As pointed out by Gallistel and Matzel (2013), synaptic plasticity does not in general exhibit this property. On the contrary, de Johnge and Racine (1985) showed that the cumulative strength of LTP is weaker when the tetanic stimulations are separated by longer intervals. Other studies have observed a spacing effect in the same direction as behavior. Zhou et al. (2003) demonstrated stronger LTP with spaced stimulation intervals. However, unlike the spacing effect in behavioral studies of learning, the benefit of spacing diminished beyond about 5 minutes, such that the longest tested intervals (10 minutes) produced LTP comparable to that of massed presentation (0 minutes). Other studies demonstrating advantages for spaced stimulation also used the optimal (5 minute) interval without assessing longer intervals (Scharf et al., 2002; Woo et al., 2003).

2.2.2 Selectivity

One of the puzzles that any associative theory of learning must address is why associations appear to be selective. For example, in the conditioned taste aversion experiments using X-rays to induce sickness (e.g., Smith and Roll, 1967; Revusky, 1968), it takes about an hour for the animals to feel sick. Everything the animal does in the intervening hour has closer temporal proximity to the sickness event, and should (according to the temporal logic of Hebbian association) form stronger aversions. Yet the animal learns to avoid the gustatory stimulus to which it was exposed rather than to any of these other stimuli (for further discussion, see Seligman, 1970; Rozin and Kalat, 1971).

In an elegant experiment, Garcia and Koelling (1966) showed that an audiovisual stimulus occurring at exactly the same time as the gustatory stimulus did not elicit avoidance. However, when an electric shock was used as the aversive outcome, animals avoided the audiovisual stimulus rather than the gustatory stimulus. This kind of selectivity does not by itself defeat all theories of learning based on synaptic plasticity, but it does impose some strong constraints. Synapses either have to be restricted to only those stimulus-selective neurons that are eligible for association, or the synapse has to somehow “know” which presynaptic inputs to ignore. Neither of these constraints has been systematically studied in the literature on synaptic plasticity.

2.2.3 Content

We have seen above how synaptic plasticity is sensitive to time. This temporal sensitivity is critically distinct from temporal coding: synapses do not provide a representation of time that is computationally accessible to downstream neurons (Gallistel and Matzel, 2013). The distinction between sensitivity and coding matters because animals appear to use representations of time to guide their behavior, even in simple Pavlovian conditioning protocols (Savastano and Miller, 1998; Molet and Miller, 2014). A few examples will suffice to illustrate this fact.

Barnet et al. (1997) trained rats with either forward conditioning (cue1 \rightarrow outcome) or backward conditioning (outcome \rightarrow cue1). Consistent with previous studies, they found weaker responding to the cue in the backward condition.³ In addition, they trained a second-order cue that preceded the first-order cue (cue2 \rightarrow cue1), finding that responding to the second-order cue was *stronger* in the backward condition. This puzzling observation makes sense under the hypothesis that animals had learned the temporal relationship between cue1 and the outcome, which they could then reuse to drive learning about cue2. Specifically, if the animals anticipate that the outcome will occur sooner following cue2 in the backward condition, they will exhibit a stronger response (consistent with the findings, already discussed, that shorter cue-outcome intervals produce stronger conditioning when the inter-trial interval is held fixed). This nicely demonstrates how an explicit representation of temporal expectation is made computationally accessible to another learning process. It is difficult to see how a purely associative account would predict the observed differences in response strength to the cues.

In a similar vein, Cole et al. (1995) used a trace conditioning procedure in which the outcome occurred 5 seconds after the offset of cue1. They then exposed animals to cue1 \rightarrow cue2 pairings. The ensuing response to cue2 was *stronger* than the response to cue1, despite the fact that cue1 had been directly paired with the outcome and cue2 had not. Again, this is deeply puzzling from an associative point of view, but makes sense if animals had learned a temporal expectation which rendered cue2 in closer proximity to the outcome compared to cue1.

Temporal coding exemplifies a more general point about the contents of memory: the brain cannot compute with information that it does not represent in memory (Gallistel and King, 2011). If we adopt the position that the brain learns some kind of content (e.g., temporal expectations), then this information must be stored in a format that is accessible to downstream computation. This logic applies to non-temporal representations. For example, pigeons can use representations of spatial relationships to identify food locations. If two landmarks (A and B) occur in a consistent spatial relationship, and then one of them (A) is placed in a particular spatial relationship to a food location, pigeons will search in the appropriate food location when exposed to landmark B, consistent with the idea that animals have learned a spatial map that supports integration of partial information (Blaisdell and Cook, 2005; Sawa et al., 2005). Similar results have been reported with rats (Chamizo et al., 2006). A parsimonious hypothesis is that these animals store the spatial coordinates of landmarks in an allocentric map.

The challenge for synaptic models of memory is to explain how information about content like time and space is stored in a format that supports the flexible behavior described above. One influential proposal is that these variables could be encoded in the activity state of neural populations. For example, elapsed time could be encoded by the stimulus-induced dynamics of population activity, and then decoded by downstream neurons (Karmarkar and Buonomano,

³Note that excitatory, albeit weaker, responding in backward conditioning is already problematic for an STDP model, which predicts only LTD.

2007). Similarly, spatial position could be encoded by population activity in an attractor network (Burak and Fiete, 2009; Samsonovich and McNaughton, 1997). Although these models make time and space computationally accessible, they don't solve the problem of long-term storage: how can information be retrieved long after persistent activity has subsided? Even over relatively short retention intervals (on the order of cellular time constants), information stored in persistent activity is quickly degraded due to noise (Burak and Fiete, 2012). The same problem occurs for some models that encode spatial information using patterns of oscillatory interference (Zilli et al., 2009). Other models posit that very slowly varying activity states can retain long-term memories (Liu et al., 2019; Shankar and Howard, 2012). However, under some conditions memory can be retained even across periods in which stimulus coding cannot be detected in ongoing activity—so-called “activity silent” memory (Beukers et al., 2021).

A different line of theorizing employs synaptic plasticity in a central role for long-term storage of temporal and spatial information. I will focus on spatial information storage in the hippocampus, which has been extensively studied. The hippocampus contains neurons that selectively fire in particular spatial locations (“place cells”), and the spatial tuning of these cells is modifiable by synaptic plasticity (see Cobar et al., 2017, for a review). The question I seek to address here is how such a mechanism can store spatial memories in a durable format. It is known that the hippocampus is essential for the performance of memory-based navigation tasks such as the Morris water maze (Morris et al., 1982; Moser et al., 1995), including retention intervals on the order of weeks (Clark et al., 2005), making it plausible that the hippocampus is a long-term storage site for spatial information.

STDP can be used to learn a “navigation map” in which firing fields shift in the animal's traversal direction (Blum and Abbott, 1996; Redish and Touretzky, 1998). The difference between coded and actual location defines a vector that approximately points towards the goal and is therefore useful for goal-directed navigation. An important challenge for this kind of model is that the map is quickly overwritten when the animal is exposed to multiple goal locations. Gerstner and Abbott (1997) solved this problem by allowing the receptive fields of hippocampal cells to be modulated by goal location. This solution presupposes that the enduring memories of goal location is stored outside the hippocampus; the hippocampal neurons are, in effect, a read-out of these memories. Gerstner and Abbott do not provide an account of how these memories are stored.

Hasselmo (2009) has suggested that route memory can be linked to sensory cues by synaptic plasticity between sensory representations and place cells. A basic problem for this kind of model is that goal-directed navigation can operate in the absence of sensory cues: rats can navigate complex mazes even when they are deprived of sight (Carr, 1917a) and smell (Carr, 1917b), provided that they have been previously familiarized with the mazes. Place cells likewise maintain their firing fields in a dark environment when rats were previously provided with visual experience of the environment (Quirk et al., 1990). Moreover, place cells which were initially under sensory control by a distal landmark (e.g., cells with firing fields that rotate in correspondence with a cue outside the arena) will continue to fire in their preferred location even when the distal landmark has been removed completely (Muller and Kubie, 1987; O'Keefe and Speakman, 1987; Shapiro et al., 1997).

In summary, the problem of content is the problem of storing the relevant content in a computationally accessible format. Existing models of memory based on synaptic plasticity have not yet offered a comprehensive account of how spatial and temporal content could be stored and retrieved in the service of flexible navigation.

2.2.4 Memory persistence and synaptic instability

As mentioned earlier, LTP typically decays to baseline on the order of hours. Yet some childhood memories last nearly our entire lifetime. How can we resolve this discrepancy?

Before discussing possible resolutions, let us dive a bit deeper into the nature of the problem: why does LTP decay? Synapses are typically contained in small dendritic protrusions, called *spines*, which grow after induction of LTP (Engert and Bonhoeffer, 1999). Spine sizes are in constant flux (see Mongillo et al., 2017, for a review). Over the course of 3 weeks, most dendritic spines in auditory cortex will grow or shrink by a factor of 2 or more (Loewenstein et al., 2011). In barrel cortex, spine size changes are smaller but still substantial (Zuo et al., 2005). Spines are also constantly being eliminated and replaced, to the extent that most spines in auditory cortex are replaced entirely over the same period of time (Loewenstein et al., 2015). In the hippocampus, the lifetime of spines is even shorter—approximately 1-2 weeks (Attardo et al., 2015).⁴ Importantly, much of the variance in dendritic spine structure is independent of plasticity pathways and ongoing neural activity (Minerbi et al., 2009; Dvorkin and Ziv, 2016; Quinn et al., 2019; Yasumatsu et al., 2008), indicating that these fluctuations are likely not generated by the covert operation of classical plasticity mechanisms. Collectively, these observations paint a picture of profound synaptic instability.

Another set of clues comes from studies of animals whose brains undergo radical remodeling during certain phases of their life cycle (Blackiston et al., 2015). Holometabolous insects undergo complete metamorphosis from the larval to adult stages, including large-scale cell death, neurogenesis, and synaptic pruning. For example, olfactory projection neurons in the antennal lobe of the fruit fly *Drosophila* persist from the larval to the adult stage, but their synapses with mushroom body neurons are disassembled during metamorphosis (Marin et al., 2005). Furthermore, the mushroom body itself is substantially reorganized during metamorphosis (Armstrong et al., 1998). Nonetheless, conditioned olfactory avoidance acquired during the larval stage is retained in the adult stage (Tully et al., 1994). This is made more puzzling from a synaptic perspective by the fact that disrupting the inputs to the mushroom body impairs the acquisition olfactory conditioning (Heisenberg et al., 1985). Apparently these synapses are required for learning but not for long-term storage.

Some mammals undergo significant brain remodeling during hibernation. Perhaps the most well-studied example is the family of European and arctic ground squirrels, whose brains experience massive dendritic spine retraction during hibernation (Popov et al., 1992; Von der Ohe et al., 2006), resulting in a 50-65% loss of synapses (Von der Ohe et al., 2007). Remarkably, this pruning is reversed within 2 hours of arousal. While studies of memory retention across hibernation have yielded mixed results (see Blackiston et al., 2015), evidence suggests that at least some information learned prior to hibernation is retained by ground squirrels after arousal (Millesi et al., 2001). Again, the challenge raised by such results is how behaviorally expressed memories could be more persistent than their putative synaptic storage site.

Even more dramatically, memories can persist in some species across decapitation and regeneration. Planarians (a class of flatworms) are distinguished by their incredible powers of regeneration from tiny tissue fragments (in some cases less than 0.4% of their body mass; Morgan, 1901). Planarians are also capable of learning: if repeatedly shocked after presentation of a light, a planarian will learn to avoid the light (Thompson and McConnell, 1955). Now suppose you cut off

⁴Some of the discrepancies between estimates of spine turnover rates may reflect methodological choices, such as the type of cranial window (see Xu et al., 2007).

a planarian's head after it has learned to avoid light. Within a week, the head will have regrown. The critical question is: will the new head remember to avoid light? Remarkably, a number of experiments, using light-shock conditioning and other learning tasks, suggested (albeit controversially) that the answer is yes (McConnell et al., 1959; Corning and John, 1961; Shomrat and Levin, 2013). What kind of memory storage mechanism can withstand utter destruction of brain tissue?

Now that we understand the challenges posed by synaptic instability, we can discuss some possible solutions. The review that follows is not exhaustive; for present purposes, it suffices to briefly introduce the key ideas and supporting evidence.

A standard answer appeals to *systems consolidation*, the process by which memories are gradually transferred from a temporary store in the hippocampus to a more durable representation in the neocortex (Squire et al., 2015). This kind of solution still begs the question: how are the new representations durable if they too depend on transient synapses? One possibility is that distributed cortical memories are more robust to degradation compared to hippocampal memories. Another possibility is that multiple traces co-exist, perhaps even in the hippocampus itself (Nadel et al., 2000), which would also confer robustness.

A number of modeling studies have explored synaptic learning rules that attempt to overcome the problem of instability. One approach relies on dendritic spines with different lifetimes. Experiments have shown that some spines are stable on the timescale of months (Yang et al., 2009). Models with multiple timescales of plasticity could potentially harness this property to construct a form of synaptic consolidation, whereby memories are passed through a cascade of progressively more stable storage sites (Fusi et al., 2005; Benna and Fusi, 2016).

Multiple timescales can also be realized at the molecular level using bistable switches (Crick, 1984). The most well-known example of a hypothetical switch is calmodulin-dependent protein kinase II (CaMKII), which assembles into a holoenzyme structure consisting of 12 CaMKII subunits (Lisman et al., 2012). LTP induction leads to phosphorylation of several subunits, and this phosphorylation state is maintained through autophosphorylation by neighboring subunits. In this way, the switch can persistently maintain memories. In support of this hypothesis, inhibiting CaMKII can reverse LTP (Barcomb et al., 2016; Sanhueza et al., 2011). At the behavioral level, inhibiting CaMKII in the hippocampus after learning impairs conditioned place preference (Rossetti et al., 2017), and inhibiting CaMKII in the nucleus accumbens impairs amphetamine sensitization (Loweth et al., 2013).⁵

Another class of models makes use of the fact that connected neurons typically have multiple synaptic contacts (Fares and Stepanyants, 2009). With appropriate coordination, these synapses can collectively store memories that outlive the lifetime of individual synapses (Deger et al., 2012; Fauth et al., 2015). Alternatively, clusters of receptors on a single synapse can be coordinated (by correlation of receptor insertation rate between neighboring locations) to extend memory lifetime (Shouval, 2005).

Some models achieve stabilization using memory reactivation ("rehearsal"), which can effectively repair degraded synapses (Wittenberg et al., 2002; Acker et al., 2019; Shaham et al., 2021; Fauth and van Rossum, 2019). Under some conditions, the correlation structure of reactivated memories can even strengthen memories that aren't reactivated (so-called "implicit rehearsal"; Wei and Koulakov, 2014).

⁵It should be noted that, while a role for CaMKII in LTP induction is widely supported, some studies find that it is not necessary for memory maintenance/expression (Malinow et al., 1989; Murakoshi et al., 2017; Chen et al., 2001).

Finally, memories could be maintained at the network level if some properties of the network remain invariant to changes in individual synapses (Susman et al., 2019) or are protected by some form of compensatory plasticity (Raman and O’Leary, 2021). Susman et al. (2019) studied how the complex eigenvalues of the network dynamics are affected by homeostatic plasticity and synaptic noise. They showed that the complex part of the eigenvalue spectrum, which determines the structure of fluctuations around the set-points of network activity, encodes long-term memories that are robust to erasure by homeostasis. They then showed that STDP mainly affects the imaginary part of the spectrum. These memories are defined not by fixed points (as in classical models of memory, such as the Hopfield network), but by limit cycles—time-varying stable states.

In summary, “garden variety” synaptic instability (i.e., the kind that appears under typical conditions of brain function) appears to be a surmountable challenge for computational models of memory persistence. The more radical varieties of synaptic stability reviewed above (e.g., during metamorphosis, hibernation, and decapitation) are not as easily dealt with. They suggest a memory storage site that transcends synapses—and possibly even cells—entirely.

2.2.5 Savings

Many forms of learning show “savings” after the original behavioral expression of learning has disappeared. For example, a conditioned response can be extinguished by repeatedly presenting the conditioned stimulus alone; when the stimulus is then reconditioned, acquisition is faster (Napier et al., 1992; Ricker and Bouton, 1996), demonstrating that the memory has not been entirely lost due to extinction. Rapid reacquisition has been observed in many different preparations, ranging from habituation of the galvanic skin response (Davis, 1934) to instrumental learning (Bullock and Smith, 1953) and visuomotor adaptation (Krakauer et al., 2005).

Savings also appears in other ways (see Bouton, 2004, for a review). Pavlov (1927) noted that extinguished responses spontaneously recover after extinction, a finding that has been demonstrated repeatedly (Rescorla, 2004). Like rapid reacquisition, spontaneous recovery has also been observed for habituation (Prosser and Hunter, 1936), instrumental learning (Rescorla, 1997), and visuomotor adaptation (Kojima et al., 2004).

A natural synaptic model of extinction assumes that the relevant synapse is depotentiated—i.e., extinction reverses the synaptic effects of conditioning. This explanation is consistent with the aforementioned endocytosis of AMPA receptors during extinction (Kim et al., 2007; Clem and Hugarir, 2010). If true, then we would not expect any savings, because the associative memory is erased. Indeed, LTP does not exhibit savings: reinduction after depotentiation proceeds at the same pace as original induction (de Johnge and Racine, 1985). Clearly, this property of LTP is at variance with the behavioral findings, motivating alternative hypotheses about the nature of extinction, according to which extinction memories are stored separately rather than erasing the original memory (e.g., Milad and Quirk, 2012; Gershman et al., 2017). These alternative hypotheses do not, without further elaboration, explain how memory is stored at the level of cells.

Studies of experimental amnesia provide another source of evidence against a synaptic model of savings. It is commonly believed that long-lasting (“late”) LTP requires protein synthesis, since inhibiting protein synthesis prevents LTP induction (Stanton and Sarvey, 1984). Accordingly, memory is almost completely obliterated within a few hours following training under systemic protein synthesis inhibition (Squire and Barondes, 1972). However the same study demonstrated that spontaneous recovery occurs when animals are tested 3 days later. Evidently, blocking LTP was not sufficient to eliminate the memory permanently (see also Power et al., 2006; Lattal and

Abel, 2004). Strikingly, the protein synthesis inhibitor can itself act as a retrieval cue, stimulating retrieval of memories that were supposedly erased by the same treatment (Bradley and Galal, 1988; Briggs and Olson, 2013; Gisquet-Verrier et al., 2015). The next section discusses other cases of recovery from experimental amnesia.

2.2.6 Dissociations between memory and synaptic integrity

If synapses are the storage site of memory, then memories should not survive the disruption of synapses. Several reports challenge this assumption.

Sensitization of siphon withdrawal in the sea slug *Aplysia* is produced by long-term facilitation (LTF) at synapses between sensory and motor neurons (Frost et al., 1985). LTF can also be induced in dissociated cell cultures using pulses of serotonin, which mediates sensitization, and is accompanied by growth of new presynaptic varicosities (Glanzman et al., 1990). LTF can be reversed by inhibiting protein synthesis following a brief ‘reminder’ pulse of serotonin (Cai et al., 2012), which causes the reversion of presynaptic varicosities to their pretraining number (Chen et al., 2014). However, the reverted synapse differs morphologically from the pretraining synapse, suggesting that protein synthesis inhibition does not simply undo the effects of LTF. Even more remarkably, Chen and colleagues showed that sensitization memory can survive synaptic degradation: a small amount of additional training following the administration of protein synthesis inhibition was sufficient to renew sensitization in previously trained animals, but not in naive animals. These findings (echoing others in the rodent literature; Power et al., 2006; Lattal and Abel, 2004) call into question the widely held belief that the sensory-motor synapse is the storage site of sensitization memory in *Aplysia* (Kandel, 2001).

Protein synthesis inhibitors also disrupt contextual fear conditioning in rodents when administered immediately after training. When injected locally into the hippocampus, they specifically suppress LTP at postsynaptic neurons that were active during memory formation (so-called “engram cells”; Ryan et al., 2015). Surprisingly, optogenetic reactivation of these engram cells can produce conditioned fear even after protein synthesis inhibition. In other words, engram reactivation can induce recovery from amnesia. This phenomenon was not specific to the hippocampus; Ryan and colleagues also observed that conditioned fear to a tone could be optogenetically reactivated after local injection of protein synthesis inhibitors into the lateral amygdala.

It is uncontroversial that protein synthesis inhibitors disrupt synaptic plasticity, and that this disruption interferes (at least temporarily) with memory expression. What these studies demonstrate is that the link between synaptic plasticity and memory *storage* is more tenuous. A number of authors have pointed out that commonly used protein synthesis inhibitors such as cycloheximide and anisomycin have other cellular effects, including apoptosis, inhibition of catecholamine function, control of post-translational protein modification, and even the *elevation* of protein synthesis due to negative feedback regulation (Routtenberg and Rekart, 2005; Rudy et al., 2006). Furthermore, a large (and largely forgotten) literature shows that the amnesic effects of cycloheximide and anisomycin can be attenuated by a wide range of treatments (e.g., amphetamines, caffeine, nicotine) that do not reverse the inhibition of protein synthesis (Martinez Jr et al., 1981). We should therefore not accept uncritically the thesis that protein synthesis inhibitors achieve their amnesic effects through destabilization of a memory trace.

Beyond protein synthesis inhibitors, Shors and Matzel (1997) have noted that “among the dozens of compounds shown to retard the induction of LTP in the hippocampus, only a few directly influence learning, and many others have no effect on learning” (p. 606). A similar

conclusion is suggested by genetic deletion studies. For example, hippocampal LTP is debilitated in mice lacking the γ isoform of protein kinase C (Abeliovich et al., 1993a), yet these mice exhibit relatively normal learning of the Morris water maze, a hippocampus-dependent task (Abeliovich et al., 1993b). Hippocampal LTP is also debilitated in mice lacking the GluR-A subunit of the AMPA receptor, yet spatial learning in the water maze was unimpaired (Zamanillo et al., 1999). The same story has been reported for brain-derived neurotrophin factor knockout mice (Montkowski and Holsboer, 1997). Mice lacking endothelial nitric oxide synthase actually show *superior* performance in the Morris water maze (Frisch et al., 2000) despite deficient LTP (Wilson et al., 1999). Taken together, these results show that synaptic plasticity and memory performance can be dissociated in many different ways.

2.2.7 Intrinsic plasticity: from excitability to coding

It has long been known that non-synaptic (or *intrinsic*) plasticity occurs in neurons (Mozzachiodi and Byrne, 2010; Titley et al., 2017). Most studies have focused on plasticity of neuronal excitability, which can be measured in different ways (most commonly the threshold or slope of the neuron's input-output function). In one early demonstration, Woody and Black-Cleworth (1973) measured the amount of current required to elicit a spike in cortical motor neurons after eyeblink conditioning. The current threshold was reduced for cells projecting to eyeblink musculature.

Changes in intrinsic excitability are often coupled with changes in synaptic strength. For example, sensitization of siphon withdrawal in *Aplysia* is accompanied by both facilitation of sensorimotor synapses (reviewed in the previous section) and increases in sensory neuron excitability (Cleary et al., 1998). Increased excitability can augment synaptic strength by broadening presynaptic action potentials, resulting in more neurotransmitter release (Byrne and Kandel, 1996). Excitability may also control the threshold for induction of synaptic plasticity (Triesch, 2007), or serve as an eligibility trace to synaptic plasticity across long delays between presynaptic and postsynaptic firing (Janowitz and Van Rossum, 2006). On the other hand, changes in excitability and synaptic strength can sometimes be dissociated (see Mozzachiodi and Byrne, 2010, for a review), suggesting that they may play distinct functional roles.

Intrinsic plasticity extends beyond changes in excitability, as illustrated by studies of Purkinje cells in the cerebellum. These cells tonically inhibit the cerebellar output nuclei; pauses in firing produce overt movement, including eyeblinks. Animals can learn to produce eyeblinks, via pauses in Purkinje cell firing, that anticipate an aversive airpuff delivered to the eye when it is preceded by a reliable stimulus cue. According to standard models of cerebellar conditioning (Marr, 1969; Albus, 1971; Bullock et al., 1994), granule cells are thought to convey a temporal pattern of activity (e.g., neurons with tuning to different time intervals) that can be used to predict airpuff timing. The airpuff signal is conveyed to Purkinje cells by climbing fibers from the inferior olivary nucleus. Granule cell synapses that are active prior to climbing fiber spikes undergo LTD, such that subsequent granule cell activity inhibits the Purkinje cell.

For a number of reasons (see Johansson, 2019, for a recent review), this model is unlikely to be correct. One reason is that pharmacological or genetic inhibition of LTD appears to have little effect on adaptive timing of conditioned eyeblinks (Welsh et al., 2005; Schonewille et al., 2011), and in some preparations LTD does not even occur (Johansson et al., 2018). Even when LTD does occur, its properties are mismatched to the properties of Purkinje cell firing: LTD emerges on the order of minutes, whereas conditioned responding of Purkinje cells (similar to the conditioned eyeblink) emerges on the order of hours (Chen and Thompson, 1995; Jirenhed et al., 2007). Another reason is

that Purkinje cells appear to store information about the timing of input signals in a cell-intrinsic format (Johansson et al., 2014). I now explain this experiment in more detail.

Johansson and colleagues directly stimulated the axons of granule cells (the parallel fibers) and the climbing fibers while measuring Purkinje cell activity in decerebrate ferrets. The same parallel fibers were stimulated every 10 milliseconds for a 200-800 millisecond interval prior to climbing fiber stimulation. This feature of their experiment is critical, because it means that there is no temporal pattern encoded by the granule cell population.⁶ Any LTD that might be happening at parallel fiber synapses would be expected to affect all the stimulated synapses in a uniform way, rendering the learning mechanism useless for predicting the timing of the climbing fiber input. Nonetheless, Purkinje cells in this preparation acquired exquisitely timed anticipatory responses to parallel fiber input, and can even acquire information about multiple time intervals (Jirenhed et al., 2017). These findings imply that the learned timing information is intrinsic to the cell.

2.2.8 Memory transfer

Few topics in neuroscience are more bizarre and controversial than memory transfer. The history is too long and complex to fully recount here (see Travis, 1981; Setlow, 1997), but I will try to extract some lessons and connect this history to more recent developments.

As discussed earlier, McConnell and his collaborators showed that planarians can retain conditioned fear of a light stimulus across decapitation (McConnell et al., 1959). At the time, neurobiologists were not yet locked into synaptic theories (LTP had not yet been discovered). Much attention was focused on the hypothesis (discussed further below) that the storage site was biochemical, possibly RNA or some other macromolecule. This seemed appealing as an account for the planarian data, since the hypothetical memory molecule could potentially survive decapitation by being distributed throughout the nervous system (McConnell, 1968). In particular, McConnell speculated that memories were stored inside neoblasts—undifferentiated, pluripotent cells that circulate throughout the body of the worm, are rich in RNA, and provide the basis of its astounding regenerative powers. The biochemical hypothesis also seemed to offer a way of addressing Lashley's well-known failure to anatomically localize the memory trace based on lesion data (Lashley, 1929).

Inspired by the biochemical hypothesis, McConnell (1962) took advantage of naturally occurring cannibalism between planarians, feeding trained subjects to untrained subjects, and finding that the untrained subjects exhibited conditioned fear of the light stimulus on the first day of training. The untrained subjects had apparently ingested a memory!

Around the same time, Corning and John (1961) pursued the RNA hypothesis more directly by bathing decapitated planarians in ribonuclease (an enzyme that hydrolyzes RNA). Under this treatment, previously trained planarians did not retain memory after decapitation. A few years later, it was shown that directly injecting RNA extracted from trained planarians into untrained planarians could produce a transfer effect (Jacobson et al., 1966). The study also included a pseudo-conditioning control group, which received the same number of light and shock stimuli but without systematic temporal contiguity between their occurrences. Pseudo-conditioned RNA donors did not produce a transfer effect in recipients, supporting the conclusion that the memory being transferred encoded specific information about the relationship between stimuli. This con-

⁶Further evidence against a granule cell time code is provided by studies showing that stimulation of inputs to granule cells tends to produce a short-latency response with little variation across cells (Jörntell and Ekerot, 2006).

trol was important for ruling out the alternative hypothesis that transfer effects were explained entirely by sensitization to the light or shock (Walker and Milton, 1966; Hartry et al., 1964). It is also important to note that the transfer effect was observed without any additional training, which argues against a metaplasticity interpretation (acceleration of new learning without retention of old learning).

Despite these compelling demonstrations of memory transfer, McConnell's research was eventually dismissed as a failure (Rilling, 1996). In a recent review, Josselyn et al. (2017) reiterate this assessment:

Many attempted to replicate McConnell's memory transfer findings. Some replicated his findings, but many did not. McConnell's work was widely criticized for being poorly designed and poorly controlled. (p. 4653)

The scholarly record supporting this assessment is quite thin. Only a small handful of published studies attempted to replicate the planarian memory transfer effects, and of these only a single one (Walker, 1966) failed to replicate the result. As already noted, the Jacobson et al. (1966) study effectively controlled for several potential confounds, yet it is rarely mentioned in the literature on memory transfer.⁷

There are several reasons why the planarian memory transfer work was unfairly maligned. First, McConnell had a penchant for self-publishing his findings in his own satirical journal, *Worm Runner's Digest* (although the first transfer study was eventually published in a peer-reviewed journal). Second, there were disputes about whether planarians could acquire classical conditioning at all, let alone transfer this learning, but these were addressed by the introduction of appropriate control conditions (Block and McConnell, 1967; Jacobson et al., 1967). Third, it appears that McConnell and other planarian researchers were victims of guilt by association with later studies of memory transfer, particularly with rodents, that failed to replicate reliably. I now turn to these studies.

The first reports of memory transfer between rodents were reported by four separate laboratories in 1965 (Babich et al., 1965b,a; Fjeringstad et al., 1965; Reinis, 1965; Ungar and Ocegüera-Navarro, 1965). By the mid-1970s, hundreds of rodent memory transfer studies had been carried out. According to one tally (Dyal, 1971), approximately half of the studies found a positive transfer effect. This unreliability, combined with uncertainty about the underlying molecular substrate and issues with inadequate experimental controls, were enough to discredit the whole line of research in the eyes of the neuroscience community (Setlow, 1997).

There are a number of important caveats to this conclusion. One is that many of the failed replications transferred RNA extracts based on the hypothesis that RNA was the memory molecule. Ungar (1974), in contrast, hypothesized that the memory molecule mediating transfer was a peptide (a short amino acid sequence). Consistent with this hypothesis, Ungar's original report of memory transfer (Ungar and Ocegüera-Navarro, 1965) showed that incubating brain homogenates with chymotrypsin (which hydrolyzes peptide bonds) prior to transfer eliminated the transfer effect, whereas incubating with ribonuclease did not.

Another problem for the RNA hypothesis is that most RNA cannot cross the blood-brain barrier. Accordingly, intra-peritoneal injections of RNA extracts did not produce a measurable change in brain RNA levels (Luttges et al., 1966; Guttman et al., 1972). Ungar believed that his peptide

⁷Despite being published in *Nature*, the Jacobson paper was omitted from several contemporary surveys of the memory transfer literature (McGaugh, 1967; Smith, 1974), as well as more recent surveys (Morange, 2006; Colaço, 2018).

hypothesis did not suffer from this problem, arguing that peptides are typically small enough to cross the blood-brain barrier. However, it is now understood that the blood-brain barrier permeability to peptides is low, and that brain capillaries have high levels of peptidases (Pardridge, 1983).

If correct, the peptide hypothesis (or indeed any non-RNA hypothesis) would render the replication failures using RNA extracts irrelevant. Ungar (1970) suggested that the memory-encoding peptide or protein might be bound in a complex to RNA, separated only after injection, which would explain how RNA could be incidentally associated with transfer effects despite not conveying learned information.

A second caveat is that the sheer variety of transferred behaviors—including habituation, passive dark avoidance, active footshock avoidance, taste aversion, conditioned magazine approach, bar pressing, wire climbing, spatial alternation, spatial discrimination, size discrimination, and color discrimination—and the variety of species tested—including worms, hamsters, mice, rats, goldfish, and even chickens—makes it difficult to completely dismiss the transfer phenomenon as a peculiarity of some particular preparation. The variability of preparations also provides insight into the conditions for successful transfer. Smith (1974) reviewed a wide range of procedural variables and concluded that many made little difference to the outcome or were equivocally related to the outcome. The most important variables related to the brain extracts, such as extract dosage, chemical composition, heat, purification procedure, among others. Because these variables differed considerably across laboratories, it is not straightforward to discriminate between replication failures and novel experimental findings.

After a long hiatus, memory transfer work has recently made a comeback, using new tools from molecular biology and new model systems. Bédécarrats et al. (2018) showed that sensitization in *Aplysia* could be transferred to untrained subjects by injection of RNA. This study went beyond the earlier literature in several ways. One is that it identified a neurophysiological correlate of transfer: sensory neurons (but not motor neurons) of RNA recipients exhibited increased excitability. RNA injection also increased the strength of a subset of sensorimotor synapses, although this increase was not significant at the population level.

Another recent study showed that learned avoidance behavior in the nematode *C. elegans* can be transferred (Moore et al., 2021). Naive animals are initially attracted to media containing the toxic bacteria *P. aeruginosa*, but they learn to avoid it after exposure. Moore and colleagues showed that ingestion of homogenates from trained animals produced avoidance behavior in naive animals. Transfer of avoidance behavior is plausibly a naturally occurring phenomenon, because exposure to *P. aeruginosa* can result in autophagy and ultimately autolysis. The lysates may then be consumed by other animals. Another remarkable feature of this model system is that the mechanism of transfer has been identified: pathogen avoidance is encoded by a small RNA (P11) packaged into virus-like particles that enable extracellular transmission. I will discuss this mechanism further when I consider the RNA hypothesis in more detail.

Memories can be transferred not only “horizontally” between individuals of the same generation, but also “vertically” between individuals of different generations (see Miska and Rechavi, 2021, for a review). Horizontal and vertical forms of transfer appear to rely on some of the same intracellular mechanisms. For example, the P11 RNA mediating horizontal transfer of pathogen avoidance in *C. elegans* also mediates vertical transfer (Kaletsky et al., 2020); indeed, Moore et al. (2021) showed that horizontally transferred memories are then vertically transferred, and both depend on the *Cer1* retrotransposon that encodes the virus-like particles encapsulating P11. This

raises the intriguing possibility that there exists a common code for memory storage that supports both forms. Conceivably, vertical transfer is the evolutionarily more ancient form (used even by prokaryotic organisms like bacteria, as well as by plants and fungi) which was co-opted by more complex organisms to support horizontal transfer. This would explain why long-term memory storage seems to rely on epigenetic mechanisms like DNA methylation and histone modification, which play central roles in transgenerational inheritance (Heard and Martienssen, 2014). I discuss these mechanisms further below.

For our purposes, the important implication of memory transfer research is that it should not be possible if memories are stored synaptically, since synapses cannot be transferred. Even if one has reservations about some of the early transfer research, it seems difficult to argue that none of the hundreds of results (including from more technically sophisticated recent experiments) are real.

2.2.9 Learning in single cells

If single cells can learn, then they must be using a non-synaptic learning mechanism. Like memory transfer, the question of single cell learning has a long and controversial history, recently surveyed in Gershman et al. (2021). The conclusion of that survey was that single cells likely do learn.

Some of the earliest studies were carried out by the zoologist Herbert Spencer Jennings and reported in his seminal book, *Behavior of the Lower Organisms* (Jennings, 1906). Jennings examined a wide range of behaviors in various simple organisms, including unicellular ciliates like *Stentor* and *Paramecium*. Jennings (see also Dexter et al., 2019) demonstrated that when repeatedly stimulated, *Stentor* exhibited a regular sequence of distinct avoidance behaviors. This suggests that the response to stimulation was memory-based: the same external stimulus produced different behavioral output as a function of stimulation history.

In the half century that followed Jennings' work, there were sporadic attempts to study learning in *Paramecium*, with mixed results. The most systematic program of research was undertaken by Beatrice Gelber, a neglected pioneer highlighted in our survey (Gershman et al., 2021). Gelber developed an appetitive conditioning paradigm in which a platinum wire was repeatedly swabbed with a bacterial suspension (a food source) and then dipped into a culture containing *Paramecia*. Gelber found that cells would attach to the wire after training, even when it had not been swabbed (Gelber, 1952). They would not produce this behavior at the beginning of training or when only exposed to the wire by itself. She interpreted these results as indicating that the cells had acquired a conditioned response to a neutral stimulus (the wire).

In a series of subsequent experiments, Gelber mapped out the effects of several quantitative parameters, such as the spacing of trials (Gelber, 1962) and the retention interval prior to the wire-only test (Gelber, 1958). Her research was criticized for suffering from confounds, such as the possibility that the *Paramecia* or bacteria left a chemical residue during training that could influence subsequent behavior on the test trial (Jensen, 1957; Katz and Deterline, 1958). She went to great lengths to address these confounds (Gelber, 1957), but ultimately this was not enough to convince skeptics.

Hennessey et al. (1979) repeatedly paired a vibrational stimulus with an alternating current and showed that *Paramecia* could be trained to produce an avoidance response to the vibration. Importantly, they also showed that this conditioning effect could not be explained as an artifact of sensitization or pseudo-conditioning. Their preparation had the advantage over Gelber's in that the electrical field (unlike the bacteria) did not produce any chemical byproducts. Finally, they

showed that the conditioned response was retained for 24 hours, confirming that the cells had acquired a long-term memory.

So far I have been discussing data from unicellular organisms, but learning has also been investigated in single cells dissociated from multicellular organisms. The biochemist Daniel Koshland followed this approach using the rat adrenal cell line known as PCI2 (reviewed in Morimoto and Koshland, 1991). Although these cells are not neurons, they share an embryological origin with neurons (in the neural crest), and have been used as a model of neurosecretion due to the fact that they carry vesicles of catecholamines and acetylcholine. Neurosecretion habituates after repetitive depolarization (McFadden and Koshland, 1990a) or cholinergic stimulation (McFadden and Koshland, 1990b). After several hours without stimulation, the secretory response rebounds, but only partially; this indicates that the cell has maintained a long-term memory trace of the stimulation history. Amazingly, a single cell can maintain multiple memory traces, as evidenced by the fact that habituation to cholinergic stimulation has no effect on habituation to depolarization, and vice versa.

While we still do not understand fully the learning capabilities of single cells, the totality of evidence makes a strong case that learning does occur in single cells. The mechanisms underlying such learning are still poorly understood. I now turn to a discussion of possible mechanisms.

3 In search of the memory molecule

If the preceding sections convinced you that molecular mechanisms are at play in memory storage, the next task is to identify them. I will focus on a subset of the possible mechanisms that have been most extensively studied, omitting some (e.g., peptides and microtubules) that are potentially important but have received less direct support.

3.1 RNA and other cytoplasmic macromolecules

The emergence of modern molecular biology in the 1940s, culminating in Watson and Crick's discovery of the DNA double helix in 1953, led to a rejuvenation of ideas about the nature of the memory trace. Even before Watson and Crick's discovery, it was understood that cells have biochemical mechanisms for information storage. Gerard (1953) drew an analogy to immune cells, which store information about pathogens long after antibody levels have subsided. Katz and Halstead (1950) suggested that biochemical memory traces in the nervous system might consist of nuclear proteins, modified as a consequence of experience. Other authors suggested cytoplasmic macromolecules as memory traces (see Uphouse et al., 1974, for a review of these early proposals).

The impetus for macromolecular theories, and the RNA hypothesis in particular, came from several directions. One was the finding, reviewed earlier, that degrading RNA with ribonuclease impaired the retention of memory in decapitated planarians (Corning and John, 1961). Another was the rather bizarre finding that administration of yeast RNA (as well as DNA) to elderly patients seemed to improve their memory (Cameron, 1958); this finding was reinforced by subsequent studies in rodents (Cook et al., 1963; Solyom et al., 1966). While these results may have reflected a kind of general enhancement of cognitive or physical activity rather than specific enhancement of memory (see Corson and Enesco, 1966; Wagner et al., 1966), the contemporaneous memory transfer results reviewed earlier suggest a possibly more specific role for RNA in memory. Finally, changes to RNA were observed following learning (Hydén and Egyhazi, 1962, 1964),

and inhibition of RNA impaired learning (Dingman and Sporn, 1961; Chamberlain et al., 1963).

Broadly speaking, macromolecular theories fell into two classes: switchboard and direct coding theories. Switchboard theories viewed macromolecules as playing the role of a switch or connector, guiding the flow of information between neurons (Szilard, 1964; Rosenblatt, 1967; Ungar, 1968, 1970). These theories align more closely with modern views of synaptic memory, according to which macromolecules play a supportive role (e.g., in determining synaptic strength). A basic difficulty facing such theories is explaining the memory transfer results. For example, Ungar (1968) argued that the hypothetical molecular connectors provide a “trail” through genetically determined neural pathways:

The situation could be compared to that of a traveler who is given a jumble of paper slips identifying the different intersections at which he has to turn to reach his destination. With the help of a map which shows the location of the intersections designated by the identification marks, he could easily find the correct sequence in which he should take them to reach his goal. For the learning process, the map is the genetically determined neural organization. When the connectors included in the extract injected into the recipient animal fall in place in the neurons homologous to their sites of origin, the nerve impulse follows the trail created by the training of the donors. (p. 229)

Ungar’s account rested largely on the principle of chemospecificity (Sperry, 1963), according to which neurons use a chemical labeling system to recognize which pathways to form during development. When neurons are coactivated, they exchange chemical signals (what Szilard dubbed “transprinting”) that allows them to form a connector representing the memory. However, Ungar did not explain precisely how the molecular connectors “know” where to go after injection, or how they would reprogram target cells.

Direct coding theories posited that macromolecules functioned as the terminal storage site for information, akin to how DNA is the terminal storage site for genetic information. Direct coding theorists appreciated that RNA likely played some role in storage, but it was unclear whether RNA itself was the terminal storage site. At the time these theories were first proposed (the early 1960s), the understanding of RNA was restricted mainly to its role in protein synthesis (i.e., messenger RNA). The first non-coding RNA was not discovered until the mid-1960s, and it would be several decades before the diversity of non-coding RNA was appreciated. Thus, apart from early speculations about coding via RNA (e.g., Gaito, 1961, 1963), direct coding theories mainly focused on protein products as the terminal storage sites (Hydén, 1961; Briggs and Kitto, 1962; Landauer, 1964). On this view, RNA plays an intermediate role in the construction of durable macromolecular codes. This point of view is buttressed by modern data on post-translational protein modification, which suggests that the combinatorial space of protein “mod-forms” provides a massive storage capacity for cellular history (Prabakaran et al., 2012).

One of the most intriguing direct coding theories was proposed by Landauer (1964), who conceived of neurons as electronic filters tuned to particular input frequencies. If the frequency tuning of a neuron is determined by some biochemical signature (Landauer was vague on this point), then changes in RNA composition would generate (via protein synthesis) changes in the neuron’s signature and hence its frequency tuning. Landauer suggested that during periods of spiking activity, a neuron’s membrane becomes permeable to RNA molecules originating in glial cells. If the spiking activity of the neuron could somehow select RNA that produce particular signatures,

then the neuron could acquire long-term frequency tuning. Landauer's ideas, which were ahead of his time and remain neglected, appear to presage more recent theories of bioelectric memory codes (Fields and Levin, 2018).

The direct coding version of the RNA hypothesis is attractive on theoretical grounds due to the fact that polynucleotide sequences are energetically cheap to modify, can be read by intracellular processes at fast timescales, and (depending on secondary structure; Nowakowski and Tinoco Jr, 1997) can be thermodynamically stable (Gallistel, 2017). The addition of a single nucleotide to a polynucleotide sequence costs 1 ATP and can buy 2 bits of information. By contrast, synaptic modification costs ~ 4 orders of magnitude more ATP (Karbowski, 2019), but by one estimate can only buy around 4-5 bits of information per synapse (Bartol Jr et al., 2015).

Arguments against the RNA hypothesis appeared almost immediately after it was first proposed. Briggs and Kitto (1962) pointed out that the experimental evidence was largely indirect and could not rule out the plausible alternative hypothesis that RNA mediated memory effects via its conventional role in protein synthesis (see also Dingman and Sporn, 1964). They also argued that this protein synthesis pathway offered few opportunities for systematic RNA modification in the service of memory storage. Their conclusion was based mainly on the understanding of messenger RNA available at the time; as already noted, the diversity of RNA and the abundance of non-coding RNA was not yet appreciated. Other arguments against RNA as a memory molecule included its inability to cross the blood-brain barrier (raising questions about the RNA memory transfer experiments, as mentioned earlier), its susceptibility to degradation by endogenous ribonucleases, and the lack of experimental evidence for functional specificity.

These arguments were sufficient to kill off the RNA hypothesis for the next half-century. However, recent developments have brought it back to life. One development, already mentioned, was the finding that RNA is responsible for transfer of pathogen avoidance between nematodes (Moore et al., 2021). This development was significant not only because it rehabilitated the plausibility of memory transfer, but also because it indicated a mechanism by which transfer could occur under natural conditions, namely by packaging RNA into virus-like particles ("capsids").

This mechanism may extend far beyond transfer of bacterial pathogen avoidance. Around 8% of the human genome consists of endogenous retroviruses—genetic material that was originally inserted into germline DNA by a virus and then passed to offspring (Lander et al., 2001). These genes have been repurposed to serve a variety of functions, one of which is the construction of capsids for packaging of RNA and possibly other molecules. Capsids can be transported out of the cell by extracellular vesicles, endocytosed in other cells, and disassembled. Messenger RNA contained in the capsids can then be translated into protein. In principle, non-coding RNA transmitted through capsids could play a regulatory role in other cells. This pathway constitutes a phylogenetically ancient form of intercellular communication (Pastuzyn et al., 2018; Hantak et al., 2021). If memories are stored in RNA, this communication pathway provides a mechanism by which other neurons could share durably encoded messages.

It is tempting to speculate that intercellular transmission of RNA is a relic of the "RNA world" that likely preceded the "DNA world" in which we currently live (Gilbert, 1986). In the RNA world, genetic information was stored in RNA rather than DNA. Although this storage format is inherently less stable, it might still have been sufficient to support simple lifeforms (and continues to serve as the genetic material for some viruses). After the emergence of DNA, the information storage capabilities of RNA may have been utilized by cells to encode non-genetic information—possibly one function (among many) implemented by non-coding RNAs (Mattick and Makunin,

2006; Akhlaghpour, 2020).

3.2 Nuclear mechanisms: DNA methylation and histone modification

Molecular memory theorists in the 1960s did not yet know about the complex mechanisms for gene regulation that exist in the nucleus. Most importantly for our purposes, some of these mechanisms are modifiable by experience and can even be inherited transgenerationally (hence the designation “epigenetic”). While my focus is not on transgenerational inheritance, I think it is unlikely to be a coincidence that these mechanisms also play a role in memory storage within a single organism’s lifetime. As mentioned earlier, it seems plausible that if evolution discovered a mechanism for transgenerational memory storage, this mechanism could be repurposed for memory storage.

In a prescient discussion of stable molecular storage, Crick (1984) was inspired by the mechanism of DNA methylation to illustrate how a bistable switch could store information despite molecular turnover. The basic unit is a protein dimer whose monomers are modifiable in a binary manner by an enzyme (e.g., phosphorylation by a kinase, or methylation by methyltransferase). In principle, the dimer can be in one of 4 possible states, but Crick further suggested that enzymatic modification of a monomer would only occur if the other monomer was already modified. Thus, the dimer would persistently be in one of two states, “active” (+,+) and “inactive” (-,-). If one of the monomers were to lose its modification state (e.g., through protein degradation), the enzymatic process would repair it.

Crick’s arguments were very general; he did not seriously consider DNA methylation itself as potential storage site. Holliday (1999) addressed this possibility more directly. Because methylation of gene promoters typically suppresses transcription, the methylation state of DNA regulates gene expression and thereby controls cellular signaling. Holliday pointed out that the potential storage capacity of DNA methylation is stupendously large: if we think of the methylation state as a binary code, then even just 30 methylation sites could represent more than a billion possible codewords, and thus the millions of sites in the human genome could store a vast number of memories.

It is now well-known that DNA methylation is affected by the history of a cell’s inputs, that these changes can sometimes be long-lasting, and that they affect animal behavior (Day and Sweatt, 2010). For example, Miller et al. (2010) found that contextual fear conditioning resulted in methylation of a memory-associated gene (calcineurin) in the dorsomedial prefrontal cortex, which persisted 30 days later. Furthermore, memory expression was reduced by inhibition of DNA methyltransferase 30 days after training.

Another epigenetic mechanism that has received considerable attention in the cell biology literature is histone modification. DNA is packaged into a compact chromatin structure by binding to histone proteins, which protect the DNA from damage and prevent tangling. In addition, histones regulate gene expression, for example by controlling the access of transcription factors to DNA. It has been suggested that the pattern of posttranslational histone modifications functions as a code that stores information about a cell’s history (Jenuwein and Allis, 2001; Turner, 2002).

The study of histones in the brain has a long history, going back to early evidence that long-term memory is associated with histone acetylation (Schmitt and Matthies, 1979), a finding corroborated by later studies (Swank and Sweatt, 2001; Levenson et al., 2004). Disruption of histone acetylation, either by genetic knockout or by pharmacological inhibition of histone deacetylase, impairs long-term memory (Alarcón et al., 2004; Korzus et al., 2004; Levenson et al., 2004; Stefanko

et al., 2009). Remarkably, inhibition of histone deacetylase can even restore access to memories after severe brain atrophy (Fischer et al., 2007).

In summary, there exist nuclear mechanisms that can plausibly function as memory storage substrates, and which have been repeatedly implicated in behaviorally expressed memory. The major gap in our understanding is the nature of the memory code—the *codebook*, in the language of communication systems. Despite liberal use of the word “code” in the cell biology literature, it is still unclear what exactly is being encoded and how. The situation in neurobiology is even more challenging, because most researchers in that area do not even regard these nuclear mechanisms as codes at all. So deeply entrenched is the synaptic view of memory that these mechanisms are widely regarded as playing a supportive role for synaptic plasticity and maintenance (Zovkic et al., 2013). The idea that they could autonomously function as long-term storage substrates has only recently been pursued (see Gold and Glanzman, 2021).

4 A new theoretical perspective

Given the empirical and conceptual problems with the classical synaptic plasticity theory, we are left with the challenge of repairing the theory and bridging the gap with non-classical molecular mechanisms. I will approach this challenge in three parts, following the template of Marr’s levels (Marr, 1982): (1) The computational problem; (2) the algorithmic solution; and (3) the physical implementation. At present, this theory is just a sketch—a starting point for a new direction of thinking.

4.1 Two computational problems: inference and learning

Peripheral receptors in the nervous system collect sensory observations, and it is the job of the brain to “make sense” of these observations. Bayesian theories of the brain hold that this sense-making process corresponds to probabilistic inference over the latent causes of observations (Knill and Pouget, 2004). As already mentioned, Bayesian inference is not tractable in general, and different theories make different claims about how inference is approximated (Gershman and Beck, 2017). The intractability of Bayesian inference also creates a problem for learning the generative model, because computing the posterior over parameters also involves the same intractable marginalization.

To formalize these points, let \mathbf{x} denote the observations, \mathbf{z} the latent causes, and θ a set of parameters that govern the generative model $p(\mathbf{x}, \mathbf{z}, \theta)$, a joint probability distribution over all of the variables. To appreciate the generative flavor of this model, it is useful to decompose the joint distribution into the product of factors that define a causal process:

$$p(\mathbf{x}, \mathbf{z}, \theta) = p(\mathbf{x}|\mathbf{z}, \theta)p(\mathbf{z}|\theta)p(\theta). \quad (1)$$

Reading from right to left, this equation says that we can draw samples from the generative model in a sequence of steps: (i) draw parameters θ ; (ii) draw latents \mathbf{z} conditional on the parameters; (iii) draw observations \mathbf{x} conditional on the latents and parameters.

After observing \mathbf{x} , Bayes’ rule stipulates how to “invert” the generative model, producing a posterior distribution over \mathbf{z} and θ :

$$p(\mathbf{z}, \theta|\mathbf{x}) = \frac{p(\mathbf{x}, \mathbf{z}, \theta)}{p(\mathbf{x})}. \quad (2)$$

The denominator (the marginal likelihood) requires marginalizing the joint distribution over \mathbf{z} and θ , which is computationally intractable in general. Thus, we cannot expect the brain to carry out exact Bayesian inference for arbitrary generative models.

To facilitate the algorithmic solution I present in the next section, it will be useful to reformulate the computational problem in a different way. For any distribution of the form $q(\mathbf{z}, \theta|\mathbf{x})$, the following equality holds:

$$\log p(\mathbf{x}) = \mathcal{D} - \mathcal{F}, \quad (3)$$

where

$$\mathcal{D} = \mathbb{E}_q \left[\log \frac{q(\mathbf{z}, \theta|\mathbf{x})}{p(\mathbf{z}, \theta|\mathbf{x})} \right] \quad (4)$$

is the *Kullback-Leibler (KL) divergence* between $q(\mathbf{z}, \theta|\mathbf{x})$ and the posterior $p(\mathbf{z}, \theta|\mathbf{x})$, and

$$\mathcal{F} = \mathbb{E}_q \left[\log \frac{q(\mathbf{z}, \theta|\mathbf{x})}{p(\mathbf{x}, \mathbf{z}, \theta)} \right] \quad (5)$$

is the *Helmholtz free energy*. The notation $\mathbb{E}_q[\cdot]$ denotes an expectation with respect to $q(\mathbf{z}, \theta|\mathbf{x})$. Because the KL divergence is always non-negative, $-\mathcal{F}$ is a lower bound on the log marginal likelihood $\log p(\mathbf{x})$.⁸ The KL divergence is minimized to 0 when $q(\mathbf{z}, \theta|\mathbf{x}) = p(\mathbf{z}, \theta|\mathbf{x})$. Thus, exact Bayesian inference is equivalent to finding a conditional distribution that minimizes free energy. This is the key idea underlying variational inference algorithms (Wainwright and Jordan, 2008), and serves as the foundation of the free energy principle in neuroscience (Friston, 2010).

We haven't yet bought anything algorithmically by formulating Bayesian inference as an optimization problem, since minimizing KL divergence to 0 requires that we have access to the true posterior. There are two reasons to adopt this formulation. One is conceptual: it allows us to think about inference and learning as optimizing the same objective function (Neal and Hinton, 1998). The other reason, which I pursue in the next section, is that it allows us to derive tractable algorithms by restricting the conditional distribution in some way.

4.2 An algorithmic solution

One way to make free energy minimization tractable is to restrict the variational posterior to a class of distributions (inference models) that are differentiable with respect to their parameters, which enables the use of gradient-based techniques for optimization. This is the key idea in *neural variational inference* (Mnih and Gregor, 2014), which uses a neural network, parametrized by ϕ , as the inference model $q_\phi(\mathbf{z}|\mathbf{x})$.⁹ This technique effectively “amortizes” the posterior computation by replacing Bayes’ rule (which takes as input both observations and a generative model) and replaces it with a function that takes as input only the observations (Stuhlmüller et al., 2013; Gershman and Goodman, 2014). This can dramatically reduce the computational cost of inference, at the expense of flexibility.

⁸In statistical machine learning, $\log p(\mathbf{x})$ is sometimes referred to as the *evidence* and hence $-\mathcal{F}$ as the *evidence lower bound*.

⁹Variational autoencoders (Kingma and Welling, 2013) follow essentially the same idea, but use a different estimator of the gradient.

4.2.1 Minimizing free energy by gradient descent

The gradient of the free energy with respect to ϕ is given by:

$$\nabla_{\phi} \mathcal{F} = \mathbb{E}_q \left[\nabla_{\phi} q_{\phi}(\mathbf{z}|\mathbf{x}) \log \frac{q_{\phi}(\mathbf{z}|\mathbf{x})}{p(\mathbf{x}, \mathbf{z}, \theta)} \right]. \quad (6)$$

We can apply the same idea to learning the parameters θ , using an inference model $q_{\lambda}(\theta|\mathbf{x})$ parametrized by λ and taking the gradient of the free energy with respect to λ :

$$\nabla_{\lambda} \mathcal{F} = \mathbb{E}_q \left[\nabla_{\lambda} q_{\lambda}(\theta|\mathbf{x}) \log \frac{q_{\lambda}(\theta|\mathbf{x})}{p(\mathbf{x}, \mathbf{z}, \theta)} \right]. \quad (7)$$

The separation of updates for ϕ and λ implies a factorized (or “mean-field”) posterior approximation:

$$q(\mathbf{z}, \theta|\mathbf{x}) = q_{\phi}(\mathbf{z}|\mathbf{x})q_{\lambda}(\theta|\mathbf{x}). \quad (8)$$

The marginalization over \mathbf{z} can be approximated by drawing N samples from $q(\mathbf{z}, \theta|\mathbf{x})$:

$$\nabla_{\phi} \mathcal{F} \approx \frac{1}{N} \sum_n \nabla_{\phi} q_{\phi}(\mathbf{z}^n|\mathbf{x}) \log \frac{q_{\phi}(\mathbf{z}^n|\mathbf{x})}{p(\mathbf{x}, \mathbf{z}^n, \theta^n)}, \quad (9)$$

where \mathbf{z}^n and θ^n denote posterior samples. An analogous expression applies to $\nabla_{\lambda} \mathcal{F}$.

Using these gradients, approximate inference and learning can be carried out by stochastic gradient descent (the “stochastic” part refers to the fact that the gradient is being approximated using samples). The simplest form of stochastic gradient descent is given by:

$$\Delta\phi = -\alpha_{\phi} \nabla_{\phi} \mathcal{F} \quad (10)$$

$$\Delta\lambda = -\alpha_{\lambda} \nabla_{\lambda} \mathcal{F} \quad (11)$$

where α_{ϕ} and α_{λ} are learning rates.

This algorithmic solution provides high-level answers to the content and structure questions about memory. Content consists of the variational parameters ϕ and λ . From the perspective of free energy optimization, there is no substantive difference between these parameters. However, from the perspective of biology, there is a substantive difference: these parameters may be encoded by different substrates, one synaptic (ϕ) and one molecular (λ). The structure of memory consists of writing and reading operations for these two types of content. Memory is written via the variational update equations (Eqs. 10 and 11). Memory is read via sampling q_{ϕ} and q_{λ} . Because of their qualitatively different substrates, I hypothesize qualitatively different neural mechanisms for these operations.

4.2.2 Informativeness and associability

A complete algorithmic theory of learning and memory should explain why learning proceeds more quickly under some conditions than others. In animal learning theory, this has traditionally been discussed in terms of “associability”—the speed at which two stimuli enter into association. Associability is commonly formalized as a learning rate parameter in models of associative learning (e.g., Bush and Mosteller, 1955; Rescorla and Wagner, 1972; Mackintosh, 1975). In the context

of Pavlovian conditioning, we can define associability more broadly as the amount of training required to produce some criterion level of conditioned responding. Balsam, Gallistel, and their colleagues have argued that one cannot understand this broader notion of associability in associative terms (Balsam et al., 2006; Balsam and Gallistel, 2009; Ward et al., 2012, 2013). This argument goes hand-in-hand with their critique of synaptic plasticity as the vehicle for learning: if synaptic plasticity is associative and learning is not, then synaptic plasticity cannot implement learning.

As an alternative, Balsam and colleagues have formalized associability as the mutual information between the conditioned stimulus (CS) and the timing of the unconditioned stimulus (US). They refer to this quantity as the *informativeness* of the CS. Here I will briefly relate this conceptualization to the theoretical framework described above.

In the setting of Pavlovian conditioning, the latent cause is a scalar (z) representing the time at which the US will occur. On average, the gradient of the free energy is given by:

$$\mathbb{E}_p[\nabla_{\phi}\mathcal{F}] = \mathbb{E}_{q,p} \left[\nabla_{\phi} q_{\phi}(z|\mathbf{x}) \log \frac{q_{\phi}(z|\mathbf{x})}{p(z|\theta)} \right] + \text{const.} \quad (12)$$

where $\mathbb{E}_p[\cdot]$ denotes an expectation with respect to $p(\mathbf{x})$, and $\mathbb{E}_{q,p}[\cdot]$ denotes an expectation with respect to $q_{\phi}(z|\mathbf{x})p(\mathbf{x})$. If we further assume that the gradient of the variational posterior is approximately constant across \mathbf{x} and z , $\nabla_{\phi} q_{\phi}(z|\mathbf{x}) \approx c(\phi)$, we obtain:

$$\mathbb{E}_p[\nabla_{\phi}\mathcal{F}] \approx c(\phi) \cdot \mathbb{E}_{q,p} \left[\log \frac{q_{\phi}(z|\mathbf{x})}{p(z|\theta)} \right] + \text{const.} \quad (13)$$

The right-hand-side (ignoring the constants) is proportional to the mutual information between the CS configuration (\mathbf{x}) and US time (z). I have thus established conditions under which learning rate in the free energy framework approximately tracks informativeness, a probabilistic formalization of associability.

4.3 Biologically plausible neural implementation

The algorithmic solution of the preceding section suggests that we should look for two types of memories in the brain. I argue here that these types are subserved by qualitatively different cellular systems: inference parameters are stored at the synapse and updated via synaptic plasticity, while generative parameters are stored in a molecular format within the cytoplasm or nucleus and updated via biochemical processes.

4.3.1 Circuit model

I hypothesize a “direct coding” scheme in which the spiking of individual neurons reports a random sample of a variable conditional on the neuron’s input. This kind of scheme has been widely used in neural models of probabilistic inference (e.g., Pecevski et al., 2011; Haefner et al., 2016; Hoyer and Hyvärinen, 2003; Buesing et al., 2011; Orbán et al., 2016).¹⁰ For ease of exposition, I will assume that all variables are binary. Let ρ_i denote the firing rate of cell i , which I take to report the posterior probability that $z_i = 1$. These “latent cause” cells receive input from a separate

¹⁰Two notable alternatives to direct coding are predictive coding, in which a neuron reports a prediction error (Bastos et al., 2012; Friston, 2005; Deneve, 2008), and probabilistic population coding (Ma et al., 2006), in which a variable is represented by the spiking activity of a neural population.

population of “observation” cells reporting the occurrence of sensory variables as well as from a subset of other latent cause cells (see below). I use \mathbf{v}_i to denote the vector of all inputs to cell i .

The implementation of the posterior over latent causes is a neural network in which firing is driven by a Poisson process with rate ρ_i (see also Rezende and Gerstner, 2014). I assume that the firing rate is an exponential function of the membrane potential u_i ,

$$\rho_i = \exp(u_i - \psi_i), \quad (14)$$

with threshold ψ_i (Jolivet et al., 2006). The membrane potential is a linear function of synaptic inputs:

$$u_i = \sum_j w_{ij} v_{ij}, \quad (15)$$

where j indexes synaptic inputs to cell i , and w_{ij} is the synaptic strength. The firing rate ρ_i corresponds to the inference model component $q_\phi(z_i = 1 | \mathbf{v}_i)$, and the synaptic strengths \mathbf{W} and the thresholds ψ correspond to the inference parameters ϕ .

I propose that the posterior over generative parameters, $q_\lambda(\theta | \mathbf{x})$, is not expressed through spiking but rather through gene expression. Samples of θ correspond to RNA transcripts which specify the instructions for evaluating the joint probability of the random variables (see below). I refer to this process as “querying” the generative model. The RNA samples are controlled by a transcription program, parametrized by λ . For example, λ might correspond to nuclear marks such as histone acetylation or DNA methylation states.

4.3.2 Connectivity

To map the spiking neuron circuit onto the inference model, we need to specify the dependency structure of the inference model. I assume that q_ϕ factorizes into a set of conditional distributions, one for each latent cause. To ensure that this factorization defines a valid distribution, we require that the connectivity structure of the latent cause cells corresponds to a directed acyclic graph (i.e., no loops between cells). We can then represent the approximate posterior over \mathbf{z} as follows:

$$q_\phi(\mathbf{z} | \mathbf{x}) = \prod_i q_\phi(z_i | \mathbf{x}, \mathbf{z}_{\text{pa}(i)}), \quad (16)$$

where $\text{pa}(i)$ denotes the parents of latent cause z_i .

To ensure locality of the plasticity rules given below, each cell needs to receive additional inputs. Let \mathcal{C} denote a partition of the variable set $\mathbf{h} = \{\mathbf{x}, \mathbf{z}, \theta\}$ such that each subset contains at most one latent cause (z_i) and the other variables in the subset belong to the Markov blanket of z_i . The joint distribution can then be expressed as the product of component distributions:

$$p(\mathbf{x}, \mathbf{z}, \theta) = p(\mathbf{h}) = \prod_{c \in \mathcal{C}} p_c(\mathbf{h}_c | \mathbf{h}_{\text{pa}(c)}), \quad (17)$$

where $\text{pa}(c)$ denotes the parents of \mathbf{h}_c . Note that this factorization can be completely different from the inference model factorization.

Each component of the partition maps to a single latent cause (i.e., i and c are in one-to-one correspondence). This means that each cell i should receive inputs from all the variables in the

union of c and $\text{pa}(c)$. This poses a challenge, because these inputs cannot depend on the synaptic parameters; the cell needs to be able to sense the values of these inputs independently of the inference parameters. One way to finesse this issue is to utilize the concept of a postsynaptically *silent synapse*, which lacks AMPA receptors (and hence does not respond to moderate levels of depolarization) but possesses NMDA receptors (Malenka and Nicoll, 1997).¹¹ If, as conventionally assumed, synaptic plasticity is mediated by trafficking of AMPA receptors to the postsynaptic membrane, then an inference model that depends on AMPA receptors can be separated from a generative model that depends on NMDA receptors. Specifically, a cell could receive inputs about its probabilistic dependencies through NMDA receptors without eliciting AMPA currents that primarily control spiking activity (i.e., sampling from the inference model).

4.3.3 Plasticity rules

The weight parameters can be updated using a form of three-factor synaptic plasticity that is equivalent to a sample-based approximation of the gradient (Eq. 6):

$$\Delta w_{ij} = \alpha_\phi \underbrace{v_j}_{\text{pre}} \underbrace{(\hat{z}_i - \rho_i)}_{\text{post}} (L_i - \hat{z}_i \log \rho_i + \rho_i), \quad (18)$$

where \hat{z}_i denotes the activity state (0 or 1) of the latent cause cell. Similarly, the threshold can be updated according to:

$$\Delta \psi_i = -\alpha_\phi (\hat{z}_i - \rho_i) (L_i - \hat{z}_i \log \rho_i + \rho_i). \quad (19)$$

The term $L_i = \log p(\mathbf{v}, \hat{z}_i, \theta)$ is a signal that I argue is generated by querying the intracellular generative model (discussed further below).

The plasticity rule for the inference parameters takes an analogous form:

$$\Delta \lambda = \alpha_\lambda \nabla_\lambda q_\lambda(\theta | \mathbf{x}) (L_i - \hat{z}_i \log \rho_i + \rho_i). \quad (20)$$

The key difference is that this rule depends on the gradient of the variational posterior over generative parameters with respect to the inference parameters λ . Recall that λ is implemented as a set of transcriptional parameters (e.g., histone or DNA methylation marks) in the nucleus. Thus, the model requires that the transcription probability mass function is differentiable with respect to these marks. Since the marks are typically taken to be binary, it's not immediately clear how this would work, but it's possible that the transcription parameters correspond to mark *frequencies* (i.e., the proportion of marks in a particular binary state), which would then make the transcription probability mass function differentiable.

4.3.4 Querying the generative model

Querying the generative model consists of evaluating the probability of \mathbf{h} under the joint probability distribution parametrized by θ . If RNA encodes θ , then one possibility is that signals encoding \mathbf{x} and \mathbf{z} bind to RNA, and then a specialized molecule reads the bound complex to report the scalar log joint probability L_i , making it accessible to the plasticity rules described above.

¹¹Alternatively, the synapse may possess both AMPA and NMDA receptors, but a sufficiently small glutamate concentration would bind preferentially to NMDA receptors, due to their substantially higher affinity.

4.4 Summary and implications

The theoretical framework presented here attempts to unify synaptic and molecular memory mechanisms under the authority of a single optimization principle. Synaptic learning optimizes an approximate posterior over latent causes; synapses are the storage sites of the parameters governing this posterior. Intracellular/molecular learning optimizes an approximate posterior over generative parameters; nuclei are the storage sites of the parameters governing this posterior.

This synthesis of mechanisms can potentially explain the disjunction between properties of synaptic plasticity and properties of behavioral learning reviewed above. Learning “facts” about the environment does not require synaptic plasticity according to my account (facts are stored inside the cell), but synapses are necessary for the expression of memory for these facts, and synaptic plasticity exists to optimize this expression. This is why synaptic plasticity accompanies learning, and why disruption of synaptic plasticity interferes with the expression of memory, even though synapses may not be the sites of memory storage.

My account also explains why memories can be transferred (both horizontally and vertically), and why they can survive dramatic brain remodeling (e.g., during metamorphosis or hibernation). When synapses are destroyed, the generative parameters are preserved in the RNA codes, which can circulate between cells in virus-like capsids (Hantak et al., 2021). As long as the cells contain generic programs for reading the information stored in RNA, they are not dependent on a particular synaptic circuit. On the other hand, these memories may become less accessible after synapse elimination or remodeling. The temporary loss of accessibility, which may be restored under certain circumstances, is broadly consistent with studies of experimental amnesia reviewed earlier, which showed that memory loss due to amnesic agents (e.g., protein synthesis inhibitors) can be temporary.

5 Insights from cognitive science

While I have focused on biological considerations, the idea of amortized inference is also relevant for understanding a number of puzzling observations in the cognitive science literature. Here I briefly review some of the evidence and its theoretical implications.

I begin by noting that Bayes’ rule is purely syntactic, in the sense that it can be applied to *any* joint distribution over observations and latent causes. Interestingly, however, human inference is not purely syntactic: probability judgments are sensitive to semantic properties of the joint distribution. In particular, probability judgments are closer to the Bayesian ideal when the joint distribution is “believable” (i.e., similar to frequently encountered distributions). Cohen et al. (2017) asked people to make inferences about medical conditions after being given information about the results of a diagnostic test. They found that people diverged substantially from Bayes’ rule when the probabilities were unbelievable. For example, a diagnostic test with a false positive rate of 80% would be considered unbelievable (no such test would ever be used in the real world). Likewise, a base rate of 50% for pneumonia would be considered unbelievable (it’s not the case that every other person you meet has had pneumonia). The believability effect documented by Cohen and colleagues is reminiscent of a similar semantic bias in syllogistic reasoning, where beliefs about the plausibility of statements influence truth judgments, violating the purely syntactic principles of logic (Revlin et al., 1980; Wilkins, 1929; Evans et al., 1983).

Belief bias is a natural consequence of amortized inference (Dasgupta et al., 2020). If the func-

tion (e.g., a neural network) used to approximate Bayesian inference has limited capacity, then it inevitably has to concentrate this capacity on distributions it encounters frequently. Dasgupta et al. (2020) tested this hypothesis directly by giving people inference problems drawn from particular distributions. When presented repeatedly with highly diagnostic evidence and low diagnostic base rates, people learned to effectively ignore the base rates, leading to errors when subsequently given test problems where the base rates are highly diagnostic. When presented repeatedly with low diagnostic evidence and high diagnostic base rates, the opposite pattern occurred, with errors on test problems where the evidence was highly diagnostic. Importantly, people behaved differently on the *same* test trials depending on what distribution of inference problems they were given. Other studies have shown that people exhibit serial dependencies between their inferences on problems presented in sequence, even though the problems were strictly independent (Dasgupta et al., 2018). These serial dependencies may be related to widely observed “carryover” effects in public opinion surveys, where answers depend on the order in which questions are asked (Tourangeau et al., 1989; Moore, 2002). Taken together, these findings suggest that people do not apply a fixed, purely syntactic inference algorithm; rather, they *learn to infer*.¹² I conjecture that this learning process is implemented by synaptic plasticity.

Learning to infer is part of a broader pattern across cognitive science. In many domains, memory is used not only to store information about the world but also to store information about how to think (Dasgupta and Gershman, 2021). A person may know the rules of chess or the axioms of mathematics, but may still not be a particularly good chess player or mathematician. These are skills that are acquired from experience, but the nature of this experience is not about “fact learning” in the traditional sense of observing the world, since the relevant knowledge is acquired by *thinking more*. We are, so to speak, *learning to think*.

A few examples will illustrate how ubiquitous this kind of learning is (see Dasgupta and Gershman, 2021, for a more extensive review). When children first learn arithmetic, they rely heavily on counting, such that the amount of time it takes to add two numbers is proportional to the smaller of the two numbers (Groen and Parkman, 1972). By the age of 10, children can answer some addition problems by retrieving the solution from memory, resulting in a much flatter response time function (Ashcraft and Fierman, 1982). The transition from counting to memory retrieval is best understood as a kind of cognitive strategy learning rather than fact learning. The facts in this case are constructed by computation, and the results of these computations are stored in memory. A similar flattening of response time functions has been observed in mental rotation (Jolicoeur, 1985; Tarr and Pinker, 1989). Initially, the time it takes to recognize an object depends on the angular distance between the object and a canonical orientation. With repeated practice, people are eventually able to retrieve the outputs of mental rotation directly from memory, without mentally rotating the object through intermediate orientations.

In summary, learning to think is conceptually distinct from, and complementary to, fact learning. Gallistel has stressed the importance of fact learning for understanding the biological basis of memory (Gallistel, 2017, 2021). I wholeheartedly agree with his argument. At the same time, learning to think may be just as important. It may provide an answer to the puzzle of synaptic plasticity posed in the beginning of this paper.

¹²See Wang and Busemeyer (2013) and Wang et al. (2014) for an alternative explanation of some question order effects in terms of quantum probability theory.

6 Conclusions

To recapitulate the central arguments:

1. The available evidence makes it extremely unlikely that synapses are the site of long-term memory storage for representational content (i.e., memory for “facts” about quantities like space, time, and number).
2. Fact memories, or more generally probabilistic beliefs about facts, are plausibly stored in an intracellular molecular format.
3. Synapses may be the site of long-term memory storage for computational parameters that facilitate fast belief updating and communication.
4. These two forms of memory (representational and computational) work synergistically to optimize a common objective function (free energy). At the biological level, this synergy is realized by interactions between synaptic and molecular processes within a cell.

Much of this theory remains speculative. Why can so little can be firmly asserted, despite decades of intensive research? One reason is that the associative-synaptic conceptualization of memory is so deeply embedded in the thinking of neurobiologists that it is routinely viewed as an axiom rather than as a hypothesis. Consequently, the most decisive experiments have yet to be carried out.

Looking at the history of research on nuclear mechanisms of memory is instructive because it has long been clear to neurobiologists that DNA methylation and histone modification play an important role. However, the prevailing view has been that these mechanisms support long-term storage at synapses, mainly by controlling gene transcription necessary for the production of plasticity-related proteins (Zovkic et al., 2013). The notion that these nuclear mechanisms might themselves be storage sites was effectively invisible because such a notion was incompatible with the associative-synaptic conceptualization. It wasn't the case that a non-associative hypothesis was considered and then rejected; it was never considered at all, presumably because no one could imagine what that would look like. This story testifies to the power of theory, even when implicit, to determine how we interpret experimental data and ultimately what experiments we do (Gershman, 2021).

We need new beginnings. Are we prepared to take a step off the terra firma of synaptic plasticity and venture into the terra incognita of a molecular memory code?

Acknowledgments

The ideas presented here have benefited enormously from conversations with Randy Gallistel, Jeremy Gunawardena, David Glanzman, Johannes Bill, Anatole Gershman, Michael Levin, and Andrew Bolton. I am also grateful to Tony Zador, Aaron Milstein, Sadegh Nabavi, and Honi Sanders for comments on an earlier draft. This material is based upon work supported by the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216.

References

- Abeliovich, A., Chen, C., Goda, Y., Silva, A. J., Stevens, C. F., and Tonegawa, S. (1993a). Modified hippocampal long-term potentiation in $pkc\gamma$ -mutant mice. *Cell*, 75:1253–1262.
- Abeliovich, A., Paylor, R., Chen, C., Kim, J. J., Wehner, J. M., and Tonegawa, S. (1993b). $Pkc\gamma$ mutant mice exhibit mild deficits in spatial and contextual learning. *Cell*, 75:1263–1271.
- Abraham, W. C., Jones, O. D., and Glanzman, D. L. (2019). Is plasticity of synapses the mechanism of long-term memory storage? *NPJ Science of Learning*, 4:1–10.
- 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. *Journal of Neuroscience*, 22:9626–9634.
- Acker, D., Paradis, S., and Miller, P. (2019). Stable memory and computation in randomly rewiring neural networks. *Journal of Neurophysiology*, 122:66–80.
- Adaikkan, C. and Rosenblum, K. (2015). A molecular mechanism underlying gustatory memory trace for an association in the insular cortex. *Elife*, 4:e07582.
- Akhlaghpour, H. (2020). An RNA-based theory of natural universal computation. *arXiv preprint arXiv:2008.08814*.
- Alarcón, J. M., Malleret, G., Touzani, K., Vronskaya, S., Ishii, S., Kandel, E. R., and Barco, A. (2004). Chromatin acetylation, memory, and LTP are impaired in CBP \pm mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. *Neuron*, 42:947–959.
- Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, 10:25–61.
- Armstrong, J. D., de Belle, J. S., Wang, Z., and Kaiser, K. (1998). Metamorphosis of the mushroom bodies; large-scale rearrangements of the neural substrates for associative learning and memory in drosophila. *Learning & Memory*, 5:102–114.
- Ashcraft, M. H. and Fierman, B. A. (1982). Mental addition in third, fourth, and sixth graders. *Journal of Experimental Child Psychology*, 33:216–234.
- Attardo, A., Fitzgerald, J. E., and Schnitzer, M. J. (2015). Impermanence of dendritic spines in live adult ca1 hippocampus. *Nature*, 523:592–596.
- Babich, F. R., Jacobson, A. L., and Bubash, S. (1965a). Cross-species transfer of learning: effect of ribonucleic acid from hamsters on rat behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 54:1299.
- Babich, F. R., Jacobson, A. L., Bubash, S., and Jacobson, A. (1965b). Transfer of a response to naive rats by injection of ribonucleic acid extracted from trained rats. *Science*, 149:656–657.
- Balsam, P., Fairhurst, S., and Gallistel, C. (2006). Pavlovian contingencies and temporal information. *Journal of Experimental psychology. Animal Behavior Processes*, 32:284–294.

- Balsam, P. D. and Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neurosciences*, 32:73–78.
- Bannerman, D. M., Good, M., Butcher, S., Ramsay, M., and Morris, R. (1995). Distinct components of spatial learning revealed by prior training and NMDA receptor blockade. *Nature*, 378:182–186.
- Barcomb, K., Hell, J. W., Benke, T. A., and Bayer, K. U. (2016). The CaMKII/GluN2B protein interaction maintains synaptic strength. *Journal of Biological Chemistry*, 291:16082–16089.
- Barnet, R. C., Cole, R. P., and Miller, R. R. (1997). Temporal integration in second-order conditioning and sensory preconditioning. *Animal Learning & Behavior*, 25:221–233.
- Bartol Jr, T. M., Bromer, C., Kinney, J., Chirillo, M. A., Bourne, J. N., Harris, K. M., and Sejnowski, T. J. (2015). Nanoconnectomic upper bound on the variability of synaptic plasticity. *Elife*, 4:e10778.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76:695–711.
- Bédécarrats, A., Chen, S., Pearce, K., Cai, D., and Glanzman, D. L. (2018). RNA from trained aplysia can induce an epigenetic engram for long-term sensitization in untrained aplysia. *eNeuro*, 5.
- Benna, M. K. and Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nature Neuroscience*, 19:1697–1706.
- Beukers, A. O., Buschman, T. J., Cohen, J. D., and Norman, K. A. (2021). Is activity silent working memory simply episodic memory? *Trends in Cognitive Sciences*, 25:284–293.
- Bi, G.-q. and Poo, M.-m. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience*, 18:10464–10472.
- Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S., and Magee, J. C. (2017). Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357:1033–1036.
- Blackiston, D. J., Shomrat, T., and Levin, M. (2015). The stability of memories during brain remodeling: a perspective. *Communicative & Integrative Biology*, 8:e1073424.
- Blaisdell, A. P. and Cook, R. G. (2005). Integration of spatial maps in pigeons. *Animal Cognition*, 8:7–16.
- Bliss, T. V. and Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology*, 232:331–356.
- Block, R. A. and McConnell, J. V. (1967). Classically conditioned discrimination in the planarian, *Dugesia dorotocephala*. *Nature*, 215:1465–1466.

- Blum, K. I. and Abbott, L. F. (1996). A model of spatial map formation in the hippocampus of the rat. *Neural Computation*, 8:85–93.
- Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, 76:198–211.
- Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, 11:485–494.
- Bradley, P. and Galal, K. (1988). State-dependent recall can be induced by protein synthesis inhibition: behavioural and morphological observations. *Developmental Brain Research*, 40:243–251.
- Briggs, J. F. and Olson, B. P. (2013). Reexposure to the amnestic agent alleviates cycloheximide-induced retrograde amnesia for reactivated and extinction memories. *Learning & Memory*, 20:285–288.
- Briggs, M. and Kitto, G. (1962). The molecular basis of memory and learning. *Psychological Review*, 69:537–541.
- Buesing, L., Bill, J., Nessler, B., and Maass, W. (2011). Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS computational Biology*, 7:e1002211.
- Bullock, D., Fiala, J. C., and Grossberg, S. (1994). A neural model of timed response learning in the cerebellum. *Neural Networks*, 7:1101–1114.
- Bullock, D. H. and Smith, W. C. (1953). An effect of repeated conditioning-extinction upon operant strength. *Journal of Experimental Psychology*, 46:349–352.
- Burak, Y. and Fiete, I. R. (2009). Accurate path integration in continuous attractor network models of grid cells. *PLoS Computational Biology*, 5:e1000291.
- Burak, Y. and Fiete, I. R. (2012). Fundamental limits on persistent activity in networks of noisy neurons. *Proceedings of the National Academy of Sciences*, 109:17645–17650.
- Bush, R. R. and Mosteller, F. (1955). *Stochastic Models for Learning*. John Wiley & Sons Inc.
- Byrne, J. H. and Kandel, E. R. (1996). Presynaptic facilitation revisited: state and time dependence. *Journal of Neuroscience*, 16:425–435.
- Cai, D., Pearce, K., Chen, S., and Glanzman, D. L. (2012). Reconsolidation of long-term memory in Aplysia. *Current Biology*, 22:1783–1788.
- Cameron, D. E. (1958). The use of nucleic acid in aged patients with memory impairment. *American Journal of Psychiatry*, 114:943–943.
- Carr, H. (1917a). Maze learning with the white rat. II. Blind animals. *Journal of Animal Behavior*, 7:277–294.
- Carr, H. (1917b). Maze studies with the white rat. III. Anosmic animals. *Journal of Animal Behavior*, 7:295–306.

- Chamberlain, T., Rothschild, G., and Gerard, R. (1963). Drugs affecting RNA and learning. *Proceedings of the National Academy of Sciences of the United States of America*, 49:918–924.
- Chamizo, V., Rodrigo, T., and Mackintosh, N. (2006). Spatial integration with rats. *Learning & Behavior*, 34:348–354.
- Chen, C. and Thompson, R. F. (1995). Temporal specificity of long-term depression in parallel fiber–Purkinje synapses in rat cerebellar slice. *Learning & Memory*, 2:185–198.
- Chen, H.-X., Otmakhov, N., Strack, S., Colbran, R. J., and Lisman, J. E. (2001). Is persistent activity of calcium/calmodulin-dependent kinase required for the maintenance of LTP? *Journal of Neurophysiology*, 85:1368–1376.
- Chen, S., Cai, D., Pearce, K., Sun, P. Y., Roberts, A. C., and Glanzman, D. L. (2014). Reinstatement of long-term memory following erasure of its behavioral and synaptic expression in Aplysia. *Elife*, 3:e03896.
- Clark, R. E., Broadbent, N. J., and Squire, L. R. (2005). Hippocampus and remote spatial memory in rats. *Hippocampus*, 15:260–272.
- Cleary, L. J., Lee, W. L., and Byrne, J. H. (1998). Cellular correlates of long-term sensitization in Aplysia. *Journal of Neuroscience*, 18:5988–5998.
- Clem, R. L. and Huganir, R. L. (2010). Calcium-permeable AMPA receptor dynamics mediate fear memory erasure. *Science*, 330:1108–1112.
- Cobar, L. F., Yuan, L., and Tashiro, A. (2017). Place cells and long-term potentiation in the hippocampus. *Neurobiology of Learning and Memory*, 138:206–214.
- Cohen, A. L., Sidlowski, S., and Staub, A. (2017). Beliefs and Bayesian reasoning. *Psychonomic Bulletin & Review*, 24:972–978.
- Colaço, D. (2018). Rip it up and start again: The rejection of a characterization of a phenomenon. *Studies in History and Philosophy of Science Part A*, 72:32–40.
- Cole, R. P., Barnet, R. C., and Miller, R. R. (1995). Temporal encoding in trace conditioning. *Animal Learning & Behavior*, 23:144–153.
- Cook, L., Davidson, A. B., Davis, D. J., Green, H., and Fellows, E. J. (1963). Ribonucleic acid: effect on conditioned behavior in rats. *Science*, 141:268–269.
- Corning, W. and John, E. (1961). Effect of ribonuclease on retention of conditioned response in regenerated planarians. *Science*, 134:1363–1365.
- Corson, J. A. and Enesco, H. E. (1966). Some effects of injections of ribonucleic acid. *Psychonomic Science*, 5:217–218.
- Crick, F. (1984). Neurobiology: Memory and molecular turnover. *Nature*, 312:101–101.
- Dasgupta, I. and Gershman, S. J. (2021). Memory as a computational resource. *Trends in Cognitive Sciences*, 25:240–251.

- Dasgupta, I., Schulz, E., Goodman, N. D., and Gershman, S. J. (2018). Remembrance of inferences past: Amortization in human hypothesis generation. *Cognition*, 178:67–81.
- Dasgupta, I., Schulz, E., Tenenbaum, J., and Gershman, S. (2020). A theory of learning to infer. *Psychological Review*, 127:412–441.
- Davis, R. (1934). Modification of the galvanic reflex by daily repetition of a stimulus. *Journal of Experimental Psychology*, 17:504–535.
- Day, J. J. and Sweatt, J. D. (2010). DNA methylation and memory formation. *Nature Neuroscience*, 13:1319–1323.
- Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, 7:889–904.
- de Johnge, M. and Racine, R. J. (1985). The effects of repeated induction of long-term potentiation in the dentate gyrus. *Brain Research*, 328:181–185.
- Deger, M., Helias, M., Rotter, S., and Diesmann, M. (2012). Spike-timing dependence of structural plasticity explains cooperative synapse formation in the neocortex. *PLoS Computational Biology*, 8.
- Deneve, S. (2008). Bayesian spiking neurons I: inference. *Neural Computation*, 20:91–117.
- Dexter, J. P., Prabakaran, S., and Gunawardena, J. (2019). A complex hierarchy of avoidance behaviors in a single-cell eukaryote. *Current Biology*, 29:4323–4329.
- Dingman, W. and Sporn, M. B. (1961). The incorporation of 8-azaguanine into rat brain RNA and its effect on maze-learning by the rat: An inquiry into the biochemical basis of memory. *Journal of Psychiatric Research*, 1:1–11.
- Dingman, W. and Sporn, M. B. (1964). Molecular theories of memory. *Science*, 144:26–29.
- Drew, P. J. and Abbott, L. (2006). Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proceedings of the National Academy of Sciences*, 103:8876–8881.
- Dudek, S. M. and Bear, M. F. (1992). Homosynaptic long-term depression in area ca1 of hippocampus and effects of n-methyl-d-aspartate receptor blockade. In *Proceedings of the National Academy of Science*, volume 89, pages 4363–4367.
- Dvorkin, R. and Ziv, N. E. (2016). Relative contributions of specific activity histories and spontaneous processes to size remodeling of glutamatergic synapses. *PLoS Biology*, 14:e1002572.
- Dyal, J. (1971). Transfer of behavioural bias: reality and specificity. In Fjerdingstad, E., editor, *Chemical Transfer of Learned Information*, pages 219–264. North-Holland.
- Engert, F. and Bonhoeffer, T. (1999). Dendritic spine changes associated with hippocampal long-term synaptic plasticity. *Nature*, 399:66–70.
- Evans, J. S. B., Barston, J. L., and Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11:295–306.

- Fares, T. and Stepanyants, A. (2009). Cooperative synapse formation in the neocortex. *Proceedings of the National Academy of Sciences*, 106:16463–16468.
- Fauth, M., Wörgötter, F., and Tetzlaff, C. (2015). The formation of multi-synaptic connections by the interaction of synaptic and structural plasticity and their functional consequences. *PLoS Computational Biology*, 11:e1004031.
- Fauth, M. J. and van Rossum, M. C. (2019). Self-organized reactivation maintains and reinforces memories despite synaptic turnover. *ELife*, 8:e43717.
- Fields, C. and Levin, M. (2018). Multiscale memory and bioelectric error correction in the cytoplasm–cytoskeleton-membrane system. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 10:e1410.
- Fischer, A., Sananbenesi, F., Wang, X., Dobbin, M., and Tsai, L.-H. (2007). Recovery of learning and memory is associated with chromatin remodelling. *Nature*, 447:178–182.
- Fjerdingstad, E., Nissen, T., and Røigaard-Petersen, H. (1965). Effect of ribonucleic acid (RNA) extracted from the brain of trained animals on learning in rats. *Scandinavian Journal of Psychology*, 6:1–6.
- Frisch, C., Dere, E., Silva, M. A. D. S., Gödecke, A., Schrader, J., and Huston, J. P. (2000). Superior water maze performance and increase in fear-related behavior in the endothelial nitric oxide synthase-deficient mouse together with monoamine changes in cerebellum and ventral striatum. *Journal of Neuroscience*, 20:6694–6700.
- Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society B: Biological sciences*, 360:815–836.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11:127–138.
- Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100:70–87.
- Frost, W. N., Castellucci, V. F., Hawkins, R. D., and Kandel, E. R. (1985). Monosynaptic connections made by the sensory neurons of the gill-and siphon-withdrawal reflex in aplysia participate in the storage of long-term memory for sensitization. *Proceedings of the National Academy of Sciences*, 82:8266–8269.
- Fusi, S., Drew, P. J., and Abbott, L. F. (2005). Cascade models of synaptically stored memories. *Neuron*, 45:599–611.
- Gaito, J. (1961). A biochemical approach to learning and memory. *Psychological Review*, 68:288–292.
- Gaito, J. (1963). DNA and RNA as memory molecules. *Psychological Review*, 70:471–480.
- Gaito, J. (1976). Molecular psychobiology of memory: Its appearance, contributions, and decline. *Physiological Psychology*, 4:476–484.
- Gallistel, C. (2021). The physical basis of memory. *Cognition*, 213:104533.

- Gallistel, C. and Gibbon, J. (2000). Time, rate, and conditioning. *Psychological review*, 107(2):289–344.
- Gallistel, C. R. (2017). The coding question. *Trends in Cognitive Sciences*, 21:498–508.
- Gallistel, C. R. and King, A. P. (2011). *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience*. John Wiley & Sons.
- Gallistel, C. R. and Matzel, L. D. (2013). The neuroscience of learning: beyond the Hebbian synapse. *Annual Review of Psychology*, 64:169–200.
- Garcia, J. and Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 4:123–124.
- Gelber, B. (1952). Investigations of the behavior of *Paramecium aurelia*: I. Modification of behavior after training with reinforcement. *Journal of Comparative and Physiological Psychology*, 45:58–65.
- Gelber, B. (1957). Food or training in *Paramecium*? *Science*, 126:1340–1341.
- Gelber, B. (1958). Retention in *Paramecium aurelia*. *Journal of Comparative and Physiological Psychology*, 51:110–115.
- Gelber, B. (1962). Acquisition in *Paramecium aurelia* during spaced training. *The Psychological Record*, 12:165–177.
- Gerard, R. W. (1953). What is memory? *Scientific American*, 189:118–127.
- Gershman, S. and Goodman, N. (2014). Amortized inference in probabilistic reasoning. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 36.
- Gershman, S. J. (2019). What does the free energy principle tell us about the brain? *Neurons, Behavior, Data analysis, and Theory*, 2:1–10.
- Gershman, S. J. (2021). Just looking: The innocent eye in neuroscience. *Neuron*, 109:2220–2223.
- Gershman, S. J., Balbi, P. E., Gallistel, C. R., and Gunawardena, J. (2021). Reconsidering the evidence for learning in single cells. *Elife*, 10:e61907.
- Gershman, S. J. and Beck, J. M. (2017). Complex probabilistic inference. In Moustafa, A., editor, *Computational Models of Brain and Behavior*, pages 453–464. Wiley Hoboken, NJ.
- Gershman, S. J., Monfils, M.-H., Norman, K. A., and Niv, Y. (2017). The computational nature of memory modification. *Elife*, 6:e23763.
- Gerstner, W. and Abbott, L. (1997). Learning navigational maps through potentiation and modulation of hippocampal place cells. *Journal of Computational Neuroscience*, 4:79–94.
- Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D., and Brea, J. (2018). Eligibility traces and plasticity on behavioral time scales: experimental support of neohebbian three-factor learning rules. *Frontiers in nNural Circuits*, 12:53.

- Gibbon, J., Baldock, M., Locurto, C., Gold, L., and Terrace, H. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, 3:264–284.
- Gilbert, W. (1986). Origin of life: The RNA world. *nature*, 319:618–618.
- Gisquet-Verrier, P., Lynch, J. F., Cutolo, P., Toledano, D., Ulmen, A., Jasnow, A. M., and Riccio, D. C. (2015). Integration of new information with active memory accounts for retrograde amnesia: a challenge to the consolidation/reconsolidation hypothesis? *Journal of Neuroscience*, 35:11623–11633.
- Glanzman, D. L., Kandel, E. R., and Schacher, S. (1990). Target-dependent structural changes accompanying long-term synaptic facilitation in aplysia neurons. *Science*, 249:799–802.
- Gold, A. R. and Glanzman, D. L. (2021). The central importance of nuclear mechanisms in the storage of memory. *Biochemical and Biophysical Research Communications*, 564:103–113.
- Groen, G. J. and Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, 79:329–43.
- Guttman, H., Matwyshyn, G., and Warriner, G. (1972). Synthetic scotophobin-mediated passive transfer of dark avoidance. *Nature New Biology*, 235:26–27.
- Haefner, R. M., Berkes, P., and Fiser, J. (2016). Perceptual decision-making as probabilistic inference by neural sampling. *Neuron*, 90:649–660.
- Hantak, M. P., Einstein, J., Kearns, R. B., and Shepherd, J. D. (2021). Intercellular communication in the nervous system goes viral. *Trends in Neurosciences*, 44:248–259.
- Hartry, A. L., Keith-Lee, P., and Morton, W. D. (1964). Planaria: memory transfer through cannibalism reexamined. *Science*, 146:274–275.
- Hasselmo, M. E. (2009). A model of episodic memory: mental time travel along encoded trajectories using grid cells. *Neurobiology of Learning and Memory*, 92:559–573.
- 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.
- Heard, E. and Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, 157:95–109.
- Hebb, D. (1949). *The Organization of Behavior*. Wiley.
- Heisenberg, M., Borst, A., Wagner, S., and Byers, D. (1985). Drosophila mushroom body mutants are deficient in olfactory learning. *Journal of Neurogenetics*, 2:1–30.
- Hennessey, T. M., Rucker, W. B., and McDiarmid, C. G. (1979). Classical conditioning in Paramecia. *Animal Learning & Behavior*, 7:417–423.
- Hinton, G. E. and Zemel, R. S. (1994). Autoencoders, minimum description length, and helmholtz free energy. *Advances in Neural Information Processing Systems*, 6:3–10.

- Holliday, R. (1999). Is there an epigenetic component in long-term memory? *Journal of Theoretical Biology*, 200:339–341.
- Hoyer, P. O. and Hyvärinen, A. (2003). Interpreting neural response variability as Monte Carlo sampling of the posterior. In *Advances in Neural Information Processing Systems*, pages 293–300.
- Hydén, H. (1961). Satellite cells in the nervous system. *Scientific American*, 205:62–71.
- Hydén, H. and Egyhazi, E. (1962). Nuclear RNA changes of nerve cells during a learning experiment in rats. *Proceedings of the National Academy of Sciences of the United States of America*, 48:1366.
- Hydén, H. and Egyhazi, E. (1964). Changes in RNA content and base composition in cortical neurons of rats in a learning experiment involving transfer of handedness. *Proceedings of the National Academy of Sciences of the United States of America*, 52:1030.
- Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17:2443–2452.
- Jacobson, A., Horowitz, S., and Fried, C. (1967). Classical conditioning, pseudoconditioning, or sensitization in the planarian. *Journal of Comparative and Physiological Psychology*, 64:73–79.
- Jacobson, A. L., Fried, C., and Horowitz, S. D. (1966). Planarians and memory: I. transfer of learning by injection of ribonucleic acid. *Nature*, 209:599–601.
- James, W. (1890). *Principles of Psychology*. Macmillan.
- Janowitz, M. K. and Van Rossum, M. C. (2006). Excitability changes that complement hebbian learning. *Network: Computation in Neural Systems*, 17:31–41.
- Jennings, H. S. (1906). *Behavior of the Lower Organisms*. Columbia University Press, The Macmillan Company, agents.
- Jensen, D. D. (1957). Experiments on “learning” in Paramecia. *Science*, 125:191–192.
- Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science*, 293:1074–1080.
- Jirenhed, D.-A., Bengtsson, F., and Hesslow, G. (2007). Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. *Journal of Neuroscience*, 27:2493–2502.
- Jirenhed, D.-A., Rasmussen, A., Johansson, F., and Hesslow, G. (2017). Learned response sequences in cerebellar Purkinje cells. *Proceedings of the National Academy of Sciences*, 114:6127–6132.
- Johansson, F. (2019). Intrinsic memory of temporal intervals in cerebellar Purkinje cells. *Neurobiology of Learning and Memory*, 166:107103.
- Johansson, F., Jirenhed, D.-A., Rasmussen, A., Zucca, R., and Hesslow, G. (2014). Memory trace and timing mechanism localized to cerebellar Purkinje cells. *Proceedings of the National Academy of Sciences*, 111:14930–14934.

- Johansson, F., Jirenhed, D.-A., Rasmussen, A., Zucca, R., and Hesslow, G. (2018). Absence of parallel fibre to Purkinje cell LTD during eyeblink conditioning. *Scientific Reports*, 8:1–5.
- Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition*, 13:289–303.
- Jolivet, R., Rauch, A., Lüscher, H.-R., and Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience*, 21:35–49.
- Jörntell, H. and Ekerot, C.-F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *Journal of Neuroscience*, 26:11786–11797.
- Josselyn, S. A., Köhler, S., and Frankland, P. W. (2017). Heroes of the engram. *Journal of Neuroscience*, 37:4647–4657.
- Kalat, J. W. and Rozin, P. (1973). “learned safety” as a mechanism in long-delay taste-aversion learning in rats. *Journal of Comparative and Physiological Psychology*, 83:198–207.
- Kaletsky, R., Moore, R. S., Vrla, G. D., Parsons, L. R., Gitai, Z., and Murphy, C. T. (2020). *C. elegans* interprets bacterial non-coding RNAs to learn pathogenic avoidance. *Nature*, 586:445–451.
- Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science*, 294:1030–1038.
- Karbowski, J. (2019). Metabolic constraints on synaptic learning and memory. *Journal of Neurophysiology*, 122:1473–1490.
- Karmarkar, U. R. and Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron*, 53:427–438.
- Katz, J. and Halstead, W. (1950). Protein organization and mental function. *Comparative Psychology Monographs*, 20:1–38.
- Katz, M. S. and Deterline, W. A. (1958). Apparent learning in the Paramecium. *Journal of Comparative and Physiological Psychology*, 51:243–247.
- Kim, J., Lee, S., Park, K., Hong, I., Song, B., Son, G., Park, H., Kim, W. R., Park, E., Choe, H. K., et al. (2007). Amygdala depotentiation and fear extinction. *Proceedings of the National Academy of Sciences*, 104:20955–20960.
- Kingma, D. P. and Welling, M. (2013). Auto-encoding variational Bayes. *2nd International Conference on Learning Representations*.
- Knill, D. C. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27:712–719.
- Kojima, Y., Iwamoto, Y., and Yoshida, K. (2004). Memory of learning facilitates saccadic adaptation in the monkey. *Journal of Neuroscience*, 24:7531–7539.
- Korzus, E., Rosenfeld, M. G., and Mayford, M. (2004). CBP histone acetyltransferase activity is a critical component of memory consolidation. *Neuron*, 42:961–972.

- Krakauer, J. W., Ghez, C., and Ghilardi, M. F. (2005). Adaptation to visuomotor transformations: consolidation, interference, and forgetting. *Journal of Neuroscience*, 25:473–478.
- Landauer, T. (1964). Two hypotheses concerning the biochemical basis of memory. *Psychological Review*, 71:167–179.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, 470:187–197.
- Lashley, K. S. (1929). *Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain*. University of Chicago Press.
- Lattal, K. M. and Abel, T. (2004). Behavioral impairments caused by injections of the protein synthesis inhibitor anisomycin after contextual retrieval reverse with time. *Proceedings of the National Academy of Sciences*, 101:4667–4672.
- Levenson, J. M., O’Riordan, K. J., Brown, K. D., Trinh, M. A., Molfese, D. L., and Sweatt, J. D. (2004). Regulation of histone acetylation during memory formation in the hippocampus. *Journal of Biological Chemistry*, 279:40545–40559.
- Lisman, J., Yasuda, R., and Raghavachari, S. (2012). Mechanisms of CaMKII action in long-term potentiation. *Nature Reviews Neuroscience*, 13:169–182.
- Liu, Y., Tiganj, Z., Hasselmo, M. E., and Howard, M. W. (2019). A neural microcircuit model for a scalable scale-invariant representation of time. *Hippocampus*, 29:260–274.
- Loewenstein, Y., Kuras, A., and Rumpel, S. (2011). Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. *Journal of Neuroscience*, 31:9481–9488.
- Loewenstein, Y., Yanover, U., and Rumpel, S. (2015). Predicting the dynamics of network connectivity in the neocortex. *Journal of Neuroscience*, 35:12535–12544.
- Loweth, J. A., Li, D., Cortright, J. J., Wilke, G., Jeyifous, O., Neve, R. L., Bayer, K. U., and Vezina, P. (2013). Persistent reversal of enhanced amphetamine intake by transient CaMKII inhibition. *Journal of Neuroscience*, 33(4):1411–1416.
- Luttges, M., Johnson, T., Buck, C., Holland, J., and McGaugh, J. (1966). An examination of “transfer of learning” by nucleic acid. *Science*, 151:834–837.
- Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9:1432–1438.
- Mackintosh, N. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82:276–98.
- Malenka, R. C. and Nicoll, R. A. (1997). Silent synapses speak up. *Neuron*, 19:473–476.
- Malinow, R., Schulman, H., and Tsien, R. W. (1989). Inhibition of postsynaptic PKC or CaMKII blocks induction but not expression of LTP. *Science*, 245:862–866.

- Marin, E. C., Watts, R. J., Tanaka, N. K., Ito, K., and Luo, L. (2005). Developmentally programmed remodeling of the drosophila olfactory circuit. *Development*, 132:725–737.
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic apss and epsps. *Science*, 275:213–215.
- Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, 202:437–470.
- Marr, D. (1982). *Vision*. MIT Press.
- Martin, S. J., Grimwood, P. D., and Morris, R. G. (2000). Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual Review of Neuroscience*, 23:649–711.
- Martinez Jr, J. L., Jensen, R. A., and McGaugh, J. L. (1981). Attenuation of experimentally-induced amnesia. *Progress in Neurobiology*, 16:155–186.
- Mattick, J. S. and Makunin, I. V. (2006). Non-coding RNA. *Human Molecular Genetics*, 15:R17–R29.
- McConnell, J. V. (1962). Memory transfer through cannibalism in planarians. *Journal of Neuropsychiatry*, 3:542–548.
- McConnell, J. V. (1968). The modern search for the engram. In Corning, W. and Balaban, M., editors, *The mind: Biological Approaches to its Functions*. Wiley.
- McConnell, J. V., Jacobson, A., and Kimble, D. (1959). The effects of regeneration upon retention of a conditioned response in the planarian. *Journal of Comparative and Physiological Psychology*, 52:1–5.
- McFadden, P. N. and Koshland, D. E. (1990a). Habituation in the single cell: diminished secretion of norepinephrine with repetitive depolarization of PC12 cells. *Proceedings of the National Academy of Sciences*, 87:2031–2035.
- McFadden, P. N. and Koshland, D. E. (1990b). Parallel pathways for habituation in repetitively stimulated PC12 cells. *Neuron*, 4:615–621.
- McGaugh, J. L. (1967). Analysis of memory transfer and enhancement. *Proceedings of the American Philosophical Society*, 111:347–351.
- McKernan, M. and Shinnick-Gallagher, P. (1997). Fear conditioning induces a lasting potentiation of synaptic currents in vitro. *Nature*, 390:607–611.
- Milad, M. R. and Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: ten years of progress. *Annual Review of Psychology*, 63:129–151.
- Miller, C. A., Gavin, C. F., White, J. A., Parrish, R. R., Honasoge, A., Yancey, C. R., Rivera, I. M., Rubio, M. D., Rumbaugh, G., and Sweatt, J. D. (2010). Cortical DNA methylation maintains remote memory. *Nature Neuroscience*, 13:664–666.
- Millesi, E., Prossinger, H., Dittami, J. P., and Fieder, M. (2001). Hibernation effects on memory in european ground squirrels (*spermophilus citellus*). *Journal of Biological Rhythms*, 16:264–271.

- Minerbi, A., Kahana, R., Goldfeld, L., Kaufman, M., Marom, S., and Ziv, N. E. (2009). Long-term relationships between synaptic tenacity, synaptic remodeling, and network activity. *PLoS Biology*, 7:e1000136.
- Miska, E. A. and Rechavi, O. (2021). Can brain activity transmit transgenerationally? *Current Topics in Developmental Biology*, 144:1–18.
- Mnih, A. and Gregor, K. (2014). Neural variational inference and learning in belief networks. In *International Conference on Machine Learning*, pages 1791–1799. PMLR.
- Molet, M. and Miller, R. R. (2014). Timing: An attribute of associative learning. *Behavioural Processes*, 101:4–14.
- Mongillo, G., Rumpel, S., and Loewenstein, Y. (2017). Intrinsic volatility of synaptic connections—a challenge to the synaptic trace theory of memory. *Current Opinion in Neurobiology*, 46:7–13.
- Montkowski, A. and Holsboer, F. (1997). Intact spatial learning and memory in transgenic mice with reduced BDNF. *Neuroreport*, 8:779–782.
- Moore, D. W. (2002). Measuring new types of question-order effects: Additive and subtractive. *The Public Opinion Quarterly*, 66:80–91.
- Moore, R. S., Kaletsky, R., Lesnik, C., Cota, V., Blackman, E., Parsons, L. R., Gitai, Z., and Murphy, C. T. (2021). The role of the Cer1 transposon in horizontal transfer of transgenerational memory. *Cell*, 184:4697–4712.
- Morange, M. (2006). What history tells us VI. the transfer of behaviours by macromolecules. *Journal of Biosciences*, 31:323–327.
- Morgan, T. H. (1901). *Regeneration*. Macmillan.
- Morimoto, B. H. and Koshland, D. E. (1991). Short-term and long-term memory in single cells. *The FASEB journal*, 5:2061–2067.
- Morris, R., Anderson, E., Lynch, G. a., 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.
- Morris, R. G., Garrud, P., Rawlins, J. a., and O’Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, 297:681–683.
- Moser, M.-B., Moser, E. I., Forrest, E., Andersen, P., and Morris, R. (1995). Spatial learning with a minislab in the dorsal hippocampus. *Proceedings of the National Academy of Sciences*, 92:9697–9701.
- Mozzachiodi, R. and Byrne, J. H. (2010). More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory. *Trends in Neurosciences*, 33:17–26.
- Muller, R. U. and Kubie, J. L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience*, 7:1951–1968.

- Murakoshi, H., Shin, M. E., Parra-Bueno, P., Szatmari, E. M., Shibata, A. C., and Yasuda, R. (2017). Kinetics of endogenous CaMKII required for synaptic plasticity revealed by optogenetic kinase inhibitor. *Neuron*, 94:37–47.
- 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.
- Nachman, M. (1970). Learned taste and temperature aversions due to lithium chloride sickness after temporal delays¹. *Journal of Comparative and Physiological Psyt*, 73:22–30.
- Nadel, L., Samsonovich, A., Ryan, L., and Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus*, 10:352–368.
- Napier, R., Macrae, M., and Kehoe, E. (1992). Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology. Animal Behavior Processes*, 18:182–192.
- Neal, R. M. and Hinton, G. E. (1998). A view of the EM algorithm that justifies incremental, sparse, and other variants. In Jordan, M. I., editor, *Learning in Graphical Models*, pages 355–368. Springer.
- Nicoll, R. A. (2017). A brief history of long-term potentiation. *Neuron*, 93:281–290.
- Nowakowski, J. and Tinoco Jr, I. (1997). Rna structure and stability. In *Seminars in Virology*, volume 8, pages 153–165. Elsevier.
- O'Keefe, J. and Speakman, A. . (1987). Single unit activity in the rat hippocampus during a spatial memory task. *Experimental Brain Research*, 68:1–27.
- Orbán, G., Berkes, P., Fiser, J., and Lengyel, M. (2016). Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron*, 92:530–543.
- Pardridge, W. M. (1983). Neuropeptides and the blood-brain barrier. *Annual Review of Physiology*, 45:73–82.
- Pastuzyn, E. D., Day, C. E., Kearns, R. B., Kyrke-Smith, M., Taibi, A. V., McCormick, J., Yoder, N., Belnap, D. M., Erlendsson, S., Morado, D. R., et al. (2018). The neuronal gene arc encodes a repurposed retrotransposon gag protein that mediates intercellular RNA transfer. *Cell*, 172:275–288.
- Pavlov, I. P. (1927). *Conditioned Reflexes*. Oxford University Press.
- Pecevski, D., Buesing, L., and Maass, W. (2011). Probabilistic inference in general graphical models through sampling in stochastic networks of spiking neurons. *PLoS Computational Biology*, 7:e1002294.
- Popov, V., Bocharova, L., and Bragin, A. (1992). Repeated changes of dendritic morphology in the hippocampus of ground squirrels in the course of hibernation. *Neuroscience*, 48:45–51.
- Power, A. E., Berlau, D. J., McGaugh, J. L., and Steward, O. (2006). Anisomycin infused into the hippocampus fails to block “reconsolidation” but impairs extinction: the role of re-exposure duration. *Learning & Memory*, 13:27–34.

- Prabakaran, S., Lippens, G., Steen, H., and Gunawardena, J. (2012). Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 4:565–583.
- Prosser, C. L. and Hunter, W. S. (1936). The extinction of startle responses and spinal reflexes in the white rat. *American Journal of Physiology*, 117:609–618.
- Quinn, D. P., Kolar, A., Harris, S. A., Wigerius, M., Fawcett, J. P., and Krueger, S. R. (2019). The stability of glutamatergic synapses is independent of activity level, but predicted by synapse size. *Frontiers in Cellular Neuroscience*, 13:291.
- Quirk, G. J., Muller, R. U., and Kubie, J. L. (1990). The firing of hippocampal place cells in the dark depends on the rat's recent experience. *Journal of Neuroscience*, 10:2008–2017.
- Raman, D. and O'Leary, T. (2021). Optimal plasticity for memory maintenance during ongoing synaptic change. *Elife*, 10:e62912.
- Redish, A. D. and Touretzky, D. S. (1998). The role of the hippocampus in solving the morris water maze. *Neural Computation*, 10:73–111.
- Reinis, S. (1965). The formation of conditioned reflexes in rats after the parenteral administration of brain homogenate. *Activitas Nervosa Superior*, 7:167–168.
- Rescorla, R. A. (1997). Spontaneous recovery of instrumental discriminative responding. *Animal Learning & Behavior*, 25:485–497.
- Rescorla, R. A. (2004). Spontaneous recovery. *Learning & Memory*, 11:501–509.
- Rescorla, R. A. and Wagner, A. R. (1972). A theory of of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. and Prokasy, W., editors, *Classical Conditioning II: Current Research and theory*, pages 64–99. Appleton-Century-Crofts, New York, NY.
- Revlin, R., Leirer, V., Yopp, H., and Yopp, R. (1980). The belief-bias effect in formal reasoning: The influence of knowledge on logic. *Memory & Cognition*, 8:584–592.
- Revusky, S. H. (1968). Aversion to sucrose produced by contingent X-irradiation: Temporal and dosage parameters. *Journal of Comparative and Physiological Psychology*, 65:17–22.
- Rezende, D. and Gerstner, W. (2014). Stochastic variational learning in recurrent spiking networks. *Frontiers in computational neuroscience*, 8:38.
- Rezende, D. and Mohamed, S. (2015). Variational inference with normalizing flows. In *International Conference on Machine Learning*, pages 1530–1538. PMLR.
- Ricker, S. T. and Bouton, M. E. (1996). Reacquisition following extinction in appetitive conditioning. *Animal Learning & Behavior*, 24:423–436.
- Rilling, M. (1996). The mystery of the vanished citations: James mcconnell's forgotten 1960s quest for planarian learning, a biochemical engram, and celebrity. *American Psychologist*, 51:589–598.

- 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.
- Rosenblatt, F. (1967). Recent work on theoretical models of biological memory. In Tou, J., editor, *Computer and Information Sciences*, pages 33–56. Spartan Books.
- Rossetti, T., Banerjee, S., Kim, C., Leubner, M., Lamar, C., Gupta, P., Lee, B., Neve, R., and Lisman, J. (2017). Memory erasure experiments indicate a critical role of CaMKII in memory storage. *Neuron*, 96:207–216.
- Routtenberg, A. and Rekart, J. L. (2005). Post-translational protein modification as the substrate for long-lasting memory. *Trends in Neurosciences*, 28:12–19.
- Rozin, P. and Kalat, J. (1971). Specific hungers and poison avoidance as adaptive specializations of learning. *Psychological Review*, 78:459–486.
- Rudy, J. W., Biedenkapp, J. C., Moineau, J., and Bolding, K. (2006). Anisomycin and the reconsolidation hypothesis. *Learning & Memory*, 13:1–3.
- Rumpel, S., LeDoux, J., Zador, A., and Malinow, R. (2005). Postsynaptic receptor trafficking underlying a form of associative learning. *Science*, 308:83–88.
- Ryan, T. J., Roy, D. S., Pignatelli, M., Arons, A., and Tonegawa, S. (2015). Engram cells retain memory under retrograde amnesia. *Science*, 348:1007–1013.
- Samsonovich, A. and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, 17:5900–5920.
- Sanhueza, M., Fernandez-Villalobos, G., Stein, I. S., Kasumova, G., Zhang, P., Bayer, K. U., Otmakhov, N., Hell, J. W., and Lisman, J. (2011). Role of the CaMKII/NMDA receptor complex in the maintenance of synaptic strength. *Journal of Neuroscience*, 31:9170–9178.
- Saucier, D. and Cain, D. P. (1995). Spatial learning without NMDA receptor-dependent long-term potentiation. *Nature*, 378:186–189.
- Savastano, H. I. and Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes*, 44:147–162.
- Sawa, K., Leising, K., and Blaisdell, A. (2005). Sensory preconditioning in spatial learning using a touch screen task in pigeons. *Journal of Experimental psychology. Animal Behavior Processes*, 31:368–375.
- Scharf, M. T., Woo, N. H., Lattal, K. M., Young, J. Z., Nguyen, P. V., and Abel, T. (2002). Protein synthesis is required for the enhancement of long-term potentiation and long-term memory by spaced training. *Journal of Neurophysiology*, 87:2770–2777.
- Schmitt, M. and Matthies, H. (1979). [biochemical studies on histones of the central nervous system. iii. incorporation of [¹⁴C]-acetate into the histones of different rat brain regions during a learning experiment]. *Acta biologica et medica Germanica*, 38:683–689.

- Schonewille, M., Gao, Z., Boele, H.-J., Veloz, M. F. V., Amerika, W. E., Šimek, A. A., De Jeu, M. T., Steinberg, J. P., Takamiya, K., Hoebeek, F. E., et al. (2011). Reevaluating the role of ltd in cerebellar motor learning. *Neuron*, 70:43–50.
- Seligman, M. E. (1970). On the generality of the laws of learning. *Psychological Review*, 77:406–18.
- Setlow, B. (1997). Georges ungar and memory transfer. *Journal of the History of the Neurosciences*, 6:181–192.
- Shaham, N., Chandra, J., Kreiman, G., and Sompolinsky, H. (2021). Stochastic consolidation of lifelong memory. *bioRxiv*.
- Shankar, K. H. and Howard, M. W. (2012). A scale-invariant internal representation of time. *Neural Computation*, 24:134–193.
- Shapiro, M. L., Tanila, H., and Eichenbaum, H. (1997). Cues that hippocampal place cells encode: dynamic and hierarchical representation of local and distal stimuli. *Hippocampus*, 7:624–642.
- Shomrat, T. and Levin, M. (2013). An automated training paradigm reveals long-term memory in planarians and its persistence through head regeneration. *Journal of Experimental Biology*, 216:3799–3810.
- Shors, T. J. and Matzel, L. D. (1997). Long-term potentiation: what's learning got to do with it? *Behavioral and Brain Sciences*, 20:597–614.
- Shouval, H. Z. (2005). Clusters of interacting receptors can stabilize synaptic efficacies. *Proceedings of the National Academy of Sciences*, 102:14440–14445.
- Smith, J. C. and Roll, D. L. (1967). Trace conditioning with X-rays as an aversive stimulus. *Psychonomic Science*, 9:11–12.
- Smith, L. (1974). The interanimal transfer phenomenon: a review. *Psychological Bulletin*, 81:1078–1095.
- Solyom, L., Beaulieu, C., and Enesco, H. E. (1966). The effect of RNA on the operant conditioned behavior of white rats. *Psychonomic Science*, 6:341–342.
- Sperry, R. W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Sciences of the United States of America*, 50:703.
- Squire, L. R. and Barondes, S. H. (1972). Variable decay of memory and its recovery in cycloheximide-treated mice. *Proceedings of the National Academy of Sciences*, 69:1416–1420.
- Squire, L. R., Genzel, L., Wixted, J. T., and Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor perspectives in biology*, 7:a021766.
- Stanton, P. and Sarvey, J. (1984). Blockade of long-term potentiation in rat hippocampal CA1 region by inhibitors of protein synthesis. *Journal of Neuroscience*, 4:3080–3088.
- Stefanko, D. P., Barrett, R. M., Ly, A. R., Reolon, G. K., and Wood, M. A. (2009). Modulation of long-term memory for object recognition via HDAC inhibition. *Proceedings of the National Academy of Sciences*, 106:9447–9452.

- Stent, G. S. (1972). Prematurity and uniqueness in scientific discovery. *Scientific American*, 227:84–93.
- Stuhlmüller, A., Taylor, J., and Goodman, N. (2013). Learning stochastic inverses. *Advances in Neural Information Processing systems*, 26:3048–3056.
- Susman, L., Brenner, N., and Barak, O. (2019). Stable memory with unstable synapses. *Nature Communications*, 10:1–9.
- Swank, M. W. and Sweatt, J. D. (2001). Increased histone acetyltransferase and lysine acetyltransferase activity and biphasic activation of the ERK/RSK cascade in insular cortex during novel taste learning. *Journal of Neuroscience*, 21:3383–3391.
- Szilard, L. (1964). On memory and recall. *Proceedings of the National Academy of Sciences of the United States of America*, 51:1092–1099.
- Takeuchi, T., Duzskiewicz, A. J., and Morris, R. G. (2014). The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369:20130288.
- Tarr, M. J. and Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, 21:233–282.
- Thompson, R. and McConnell, J. (1955). Classical conditioning in the planarian, *Dugesia dorotocephala*. *Journal of Comparative and Physiological Psychology*, 48:65–68.
- Titley, H. K., Brunel, N., and Hansel, C. (2017). Toward a neurocentric view of learning. *Neuron*, 95:19–32.
- Tourangeau, R., Rasinski, K. A., Bradburn, N., and D’ANDRADE, R. (1989). Carryover effects in attitude surveys. *Public Opinion Quarterly*, 53:495–524.
- Travis, D. (1981). On the construction of creativity: The ‘memory transfer’ phenomenon and the importance of being earnest (1). In *The Social Process of Scientific Investigation*, pages 165–193. Springer.
- Triesch, J. (2007). Synergies between intrinsic and synaptic plasticity mechanisms. *Neural Computation*, 19:885–909.
- Tully, T., Cambiazo, V., and Kruse, L. (1994). Memory through metamorphosis in normal and mutant drosophila. *Journal of Neuroscience*, 14:68–74.
- Turner, B. M. (2002). Cellular memory and the histone code. *Cell*, 111:285–291.
- Ungar, G. (1968). Molecular mechanisms in learning. *Perspectives in biology and medicine*, 11(2):217–232.
- Ungar, G. (1970). Molecular mechanisms in information processing. In *International Review of Neurobiology*, volume 13, pages 223–253. Elsevier.
- Ungar, G. (1974). Peptides and memory. *Biochemical Pharmacology*, 23:1553–1558.

- Ungar, G. and Oceguera-Navarro, C. (1965). Transfer of habituation by material extracted from brain. *Nature*, 207:301–302.
- Uphouse, L. L., MacInnes, J. W., and Schlesinger, K. (1974). Role of RNA and protein in memory storage: A review. *Behavior Genetics*, 4:29–81.
- Von der Ohe, C. G., Darian-Smith, C., Garner, C. C., and Heller, H. C. (2006). Ubiquitous and temperature-dependent neural plasticity in hibernators. *Journal of Neuroscience*, 26:10590–10598.
- Von der Ohe, C. G., Garner, C. C., Darian-Smith, C., and Heller, H. C. (2007). Synaptic protein dynamics in hibernation. *Journal of Neuroscience*, 27:84–92.
- Wagner, A. R., Carder, J. B., and Beatty, W. W. (1966). Yeast ribonucleic acid: Effects on learned behavior in the rat. *Psychonomic Science*, 4:33–34.
- Wainwright, M. J. and Jordan, M. I. (2008). *Graphical Models, Exponential Families, and Variational Inference*. Now Publishers Inc.
- Walker, D. (1966). Memory transfer in planarians: an artifact of the experimental variables. *Psychonomic Science*, 5:357–358.
- Walker, D. and Milton, G. (1966). Memory transfer vs. sensitization in cannibal planarians. *Psychonomic Science*, 5:293–294.
- Wang, Z. and Busemeyer, J. R. (2013). A quantum question order model supported by empirical tests of an a priori and precise prediction. *Topics in Cognitive Science*, 5:689–710.
- Wang, Z., Solloway, T., Shiffrin, R. M., and Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proceedings of the National Academy of Sciences*, 111:9431–9436.
- Ward, R., Gallistel, C., Jensen, G., Richards, V., Fairhurst, S., and Balsam, P. (2012). Conditioned stimulus informativeness governs conditioned stimulus-unconditioned stimulus associability. *Journal of Experimental psychology. Animal Behavior Processes*, 38:217–232.
- Ward, R. D., Gallistel, C., and Balsam, P. D. (2013). It's the information! *Behavioural Processes*, 95:3–7.
- Wei, Y. and Koulakov, A. A. (2014). Long-term memory stabilized by noise-induced rehearsal. *Journal of Neuroscience*, 34:15804–15815.
- Welsh, J. P., Yamaguchi, H., Zeng, X.-H., Kojo, M., Nakada, Y., Takagi, A., Sugimori, M., and Llinás, R. R. (2005). Normal motor learning during pharmacological prevention of Purkinje cell long-term depression. *Proceedings of the National Academy of Sciences*, 102:17166–17171.
- 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.
- Wilkins, M. C. (1929). The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*.

- Wilson, R., Gödecke, A., Brown, R., Schrader, J., and Haas, H. (1999). Mice deficient in endothelial nitric oxide synthase exhibit a selective deficit in hippocampal long-term potentiation. *Neuroscience*, 90:1157–1165.
- Wittenberg, G. M., Sullivan, M. R., and Tsien, J. Z. (2002). Synaptic reentry reinforcement based network model for long-term memory consolidation. *Hippocampus*, 12:637–647.
- Woo, N. H., Duffy, S. N., Abel, T., and Nguyen, P. V. (2003). Temporal spacing of synaptic stimulation critically modulates the dependence of ltp on cyclic amp-dependent protein kinase. *Hippocampus*, 13:293–300.
- Woody, C. D. and Black-Cleworth, P. (1973). Differences in excitability of cortical neurons as a function of motor projection in conditioned cats. *Journal of Neurophysiology*, 36:1104–1116.
- Xiong, Q., Znamenskiy, P., and Zador, A. M. (2015). Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature*, 521:348–351.
- Xu, H.-T., Pan, F., Yang, G., and Gan, W.-B. (2007). Choice of cranial window type for in vivo imaging affects dendritic spine turnover in the cortex. *Nature Neuroscience*, 10:549–551.
- Yang, G., Pan, F., and Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature*, 462:920–924.
- Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., and Kasai, H. (2008). Principles of long-term dynamics of dendritic spines. *Journal of Neuroscience*, 28:13592–13608.
- Zamanillo, D., Sprengel, R., Hvalby, Ø., Jensen, V., Burnashev, N., Rozov, A., Kaiser, K. M., Köster, H. J., Borchardt, T., Worley, P., et al. (1999). Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. *Science*, 284:1805–1811.
- Zhou, Q., Tao, H. W., and Poo, M.-m. (2003). Reversal and stabilization of synaptic modifications in a developing visual system. *Science*, 300:1953–1957.
- Zilli, E. A., Yoshida, M., Tahvildari, B., Giocomo, L. M., and Hasselmo, M. E. (2009). Evaluation of the oscillatory interference model of grid cell firing through analysis and measured period variance of some biological oscillators. *PLoS Computational Biology*, 5:e1000573.
- Zovkic, I. B., Guzman-Karlsson, M. C., and Sweatt, J. D. (2013). Epigenetic regulation of memory formation and maintenance. *Learning & Memory*, 20:61–74.
- Zuo, Y., Lin, A., Chang, P., and Gan, W.-B. (2005). Development of long-term dendritic spine stability in diverse regions of cerebral cortex. *Neuron*, 46:181–189.