Time scales of memory, learning, and plasticity

Christian Tetzlaff, Christoph Kolodziejski, Irene Markelic & Florentin Wörgötter

Springer

Biological
Cybernetics

Time scales of memory, learning, and plasticity

**Christian Tetzlaff · Christoph Kolodziejski ·
Irene Markelic · Florentin Wörgötter**

Abstract If we stored every bit of input, the storage capacity of our nervous system would be reached after only about 10 days. The nervous system relies on at least two mechanisms that counteract this capacity limit: compression and forgetting. But the latter mechanism needs to know how long an entity should be stored: some memories are relevant only for the next few minutes, some are important even after the passage of several years. Psychology and physiology have found and described many different memory mechanisms, and these mechanisms indeed use different time scales. In this prospect we review these mechanisms with respect to their time scale and propose relations between mechanisms in learning and memory and their underlying physiological basis.

C. Tetzlaff (✉) · C. Kolodziejski · I. Markelic · F. Wörgötter
Bernstein Centre for Computational Neuroscience,
III. Institute of Physics–Biophysics, Georg-August-Universität,
Göttingen, Germany
e-mail: tetzlaff@physik3.gwdg.de

C. Kolodziejski
e-mail: kolo@physik3.gwdg.de

I. Markelic
e-mail: irene@physik3.gwdg.de

F. Wörgötter
e-mail: worgott@physik3.gwdg.de

1 Introduction

The most prominent method for survival that living beings have is to sense certain features of their environment and (re)act accordingly. The simplest behavior relies on direct sensorimotor coupling (reaction); more advanced behavior entails making simple sensing predictions (proaction) or even base the method on sophisticated and memorized environmental knowledge (planning) built by multisensory experience.

For the most basic mechanism—reaction—memory is not necessary because a direct connection from sensors to motors suffices (cf. Braitenberg 1984). The second—proaction—involves reaction to stimuli that are *expected*. Thus, for such a behavior the underlying mechanisms require the collection of stimuli from the past on which predictions for the future are based. The third—planning—relies on remembered stimuli, combines predictions into a plan, and must memorize the state of the plan while executing it.

The last two behaviors discussed above rely mainly on two concepts:

(1) The first concept consists of an extraction mechanism that condenses information from past experiences and performs refinement processing of the condensed information so that behaviorally useful predictions of the future can be made. This concept is often referred to as "learning," and we find mainly three mechanisms for it: (a) correlation learning is based on evaluating statistically significant relations between events in the environment, (b) reinforcement learning or reward-based learning is based on unspecific good–bad signals, and (c) supervised learning relies on specific error signals.

(2) The second concept consists of a bridging mechanism that retains the condensed past and present information

for future use. This concept is often referred to as "memory," and three subcategories of memory have been identified. Working memory processes information and usually requires attention, and short-term and long-term memory performs long-term storage of information where the difference between the two concepts lies in duration and capacity.

In higher animals the substrate responsible for learning and memory consists of specialized cells, mainly neurons and their connections, known as synapses [but see also, for instance, Perea et al. (2009) for astrocytes and Stocker and Durham (2009) for protozoa]. How do we know that neurons and their connections are really related to learning and memory? In 2000, Martin et al. in a review suggested four different criteria supporting the hypothesis that plasticity is the basic mechanism for learning and memory: *detectability, mimicry, anterograde, and retrograde alteration* (Martin et al., 2000; Martin and Morris, 2002).

Detectability means that there are changes at the synaptic level after an animal has learned or memorized something (e.g., Rioult-Pedotti et al., 1998; Whitlock et al., 2006).

If we took the changes in one animal and implemented them in another, then we would invoke the mechanism of *mimicry*. So far, mimicry cannot be experimentally induced.

By the mechanism of an *anterograde alteration* synaptic plasticity is prevented, which should then also prevent the animal from memorizing or learning something. This is the most prominent method to prove the relation between plasticity and learning (Morris, 1989; Martin et al., 2000).

Retrograde alteration implies varying the synaptic strength in such a way that a learned memory item can be "reprogrammed" to become another memory (Pastalkova et al. 2006).

Information in the environment is available usually only for a limited duration. Additionally, information that is important now might not be of relevance anymore sometime later—other information, on the other hand, might have to be considered lifelong. If we want to find correlations or even causalities—and knowledge of causalities is often live saving—in the environment, unknown time intervals must be bridged. Some correlations are immediate, like the disgusting taste of rotten berries, some are hours apart, like diarrhea after eating such berries, and some correlations are even years apart, like the effects of lethal parasites in the berries, hence the mechanisms for memorizing this have to operate on different time scales.

In the first part of this article we list and briefly review behavioral concepts as well as some physiological mechanisms (cf. Sects. 2–4) and discuss their relations (cf. Sect. 5). In the second part we discuss the advantages of different time scales versus complete storage of all information also with respect to capacity issues and processing time (cf. Sect. 6).

In the end, the reader should have a detailed impression of time scales of memory, learning, and plasticity and of the links that exist between these mechanisms. Finally, building upon this discussion, we raise open research questions (cf. Sect. 7).

2 Time scales of memory

2.1 Working memory

Working memory, a term coined by Miller et al. (1960), is a theoretical concept believed to underlie human thought processes. According to Miller et al., it is a mechanism that allows us to remember the current state of a plan that we are executing, for example, while doing mental arithmetic we need to remember intermediate results or while fleeing from a threat we must not suddenly forget why we are running. An important feature of working memory is that it has an individually limited capacity between four and seven storable items (Baddeley and Hitch 1974; Daneman and Carpenter 1980; Miller 1956; Cowan 2001); in addition, it has been shown that there is a high correlation between this working memory capacity and human intelligence (Kyllonen and Christal 1990). Note that this capacity might be increased by a process called chunking, whereby atomic information is combined into larger compounds, with the chunks being stored somewhere else. By "loading" links or references to the chunks into working memory instead of the information encoded by them, an increase in capacity can be obtained (Ericsson and Kintsch 1995). A further feature is that information in working memory is quickly lost if not constantly rehearsed (Brown 1958; Baddeley and Hitch 1974) or if it is not the focus of attention (Cowan 1995, 2005). Thus, the time scale of working memory is from milliseconds to minutes. Lesion and brain imaging studies mostly relate working memory to the prefrontal cortex (Owen 1997; Fuster 1973; Cohen et al. 1997).

2.2 Short- and long-term memory

Short-term memory Whereas the time scale of working memory is from milliseconds to minutes, short-term memory stores items on a time scale of minutes to days. However, right after a new memory is formed, it is still susceptible to perturbations and must be consolidated [e.g. during sleep (Walker et al. 2003)] to become stabilized for the next few days (synaptic consolidation; Dudai 2004). The corresponding brain location that is mostly associated with short-term memory is the hippocampus (Kumaran 2008), where changes in morphology due to synaptic plasticity are found in short-term memory tasks (Shimizu 2000). Another brain area, the amygdala, has, through the emotions, a strong

influence on the memorization process (McGaugh 2000). On the other hand, some theories propose that this link between short-term memory and specific brain areas depends on the memory context (Nadel and Moscovitch 1997; Henke 2010).

Long-term memory The memory type with the longest time scale is long-term memory, which operates on a scale of days to years—perhaps even for an entire lifetime. Changes during long-term-memory experiments are found in a distributed cortical network (Frankland et al. 2001, 2004), however, mainly in the prefrontal cortex (Frankland and Bontempi 2005). Those changes happen through synaptic and structural plasticity (Frankland and Bontempi 2005; Xu et al. 2009; Yang et al. 2009; Ziv and Ahissar 2009). Although the time scale of long-term memory could be up to "infinity," it is still under debate how items can be stored for such a long time (Frankland and Bontempi 2005).

Interaction between short- and long-term memory The interaction between short- and long-term memory remains unclear (McGaugh 2000). Does a memory item have to be stored in short-term memory before it is transferred to long-term memory (Marr et al. 1970; McClelland et al. 1995; Squire and Alvarez 1995)? Or does the learning paradigm influence the type of memory an item is stored in (Izquierdo et al. 1998; Henke 2010), which would mean that short-term memory and long-term memory act independently of each other (i.e., in parallel)? Another possibility is that short-term memory supports the consolidation of long-term memory (Dudai 2002). Such a transformation of memory items from short-term to long-term memory is speculated to happen between the hippocampus and the neocortex [system consolidation; (Dudai 2004)] and probably during sleep (Diekelmann and Born 2010). However, it is difficult to investigate consolidation in memory recall because for memory deficits it cannot be distinguished between a loss in consolidation or a loss in retrieval (Dudai 2004).

The models explained above attribute short-term memories to the hippocampus and long-term memories to the cortex. In the multiple trace theory, Nadel and Moscovitch (1997) hypothesize that the hippocampus is responsible for recent episodic memories and the cortex for semantic memories.

3 Time scales of memory that are relevant for learning processes

Novices in the field often confuse learning and memory. This is possibly due to the fact that many *processes of learning* require access to "something stored," hence, to some form of memory. This process—the learning—will then normally lead to the formation of new memories (which may or may not have anything to do with those items that had to be stored and retrieved during the learning process).

One of the most prominent learning paradigms is correlation-based learning in which two stimuli/events are correlated over time and this correlation is then remembered. Classical conditioning (Pavlov 1927) is the traditional example of this. The time difference between both events (e.g., conditioned and unconditioned stimuli) should be within seconds to evoke significant behavioral changes. With higher-order conditioning it is possible to lengthen the time duration up to minutes. However, to correlate both stimuli with each other the organism must memorize the first one. The short behaviorally relevant time scale suggests that this is stored in the working memory. In exceptional cases the first event can be emotionally biased and thus stored in the short- or even long-term memory (McGaugh 2000).

A more complex learning paradigm is reward-based learning [reinforcement learning, (Sutton and Barto 1998)]. In this approach an organism has several choices of behavior, and each choice can lead to a different outcome [reward/punishment; cf. (Sutton and Barto 1981)], which can also be probabilistic. To figure out if one choice is better than another, the organism must remember the outcomes and compare both. Thus, reward-based learning needs at least information coming from working memory (last choice). Of course, the result—which choice is better—can then be stored in the short- or long-term memory.

All of the learning paradigms discussed above are unsupervised. However, humans and several animals use for their offspring supervised learning methods [cf. Duda et al. (2000); Bishop (2007)]. This can be done, for instance, by imitation learning or by presenting an error (teacher) signal. Often there are no hard temporal constraints on when the error signal should occur, and it could even be presented much later. Thus, the organism must have memorized what it has done and what the error signal was. Thus, for the process of supervised learning much information must be retrieved from short- and long-term memory to update the behavior.

4 Time scales of physiology

In what follows, we present physiological models relevant for memory and learning. Many of these physiological mechanisms have been extensively investigated mathematically and computationally. To give the reader an overview of this broad field, we provide in Table 1 the most prominent mathematical descriptions of these physiological models.

4.1 Activity

The shortest time scale with a "memorizing" character in physiology can be found right after a spike arrives at the

Table 1 Mathematical overview of different physiological mechanisms where the temporal duration increases from 1 to 8

Physiological mechanism	Mathematical description	Remarks
(1) Persistent activity (Rall 1967)	$m_i = \frac{t}{\mu} e^{1-t/\mu} \cdot \Theta(t)$	
(2) Intrinsic plasticity (Triesch 2007)	$v_i = \frac{1}{1+\exp(-(a_i m_i + b_i))}$	a_i: slope parameter;
	$\Delta a_i = \mu \left(\frac{1}{a_i} + m_i - \left(2 + \frac{1}{v_T}\right) m_i v_i + \frac{m_i v_i^2}{v_T} \right)$	b_i: offset parameter
	$\Delta b_i = \mu \left(1 - \left(2 + \frac{1}{v_T} v_i\right) + \frac{v_i^2}{v_T} \right)$	
(3) Attractor network (Hopfield 1982)	$\Delta w_{ij} = \mu \xi_i \xi_j \Rightarrow w_{ij} = w_{ji}$	ξ: pattern; $\xi_i = \{-1, 1\}$;
	$v_i = \Theta(\sum_j w_{ij} v_j)$	network phenomenon
(4) Short-term plasticity (Mongillo et al. 2008)	$\dot{p}_{ij} = \mu \left(1 - p_{ij}\right) - r_{ij} p_{ij} \delta(t - t_{sp})$	p_{ij}: depression; r_{ij}: facilitation; R: baseline level
	$\dot{r}_{ij} = \lambda \left(R - r_{ij}\right) + R(1 - r_{ij}) \delta(t - t_{sp}) w_{ij} = p_{ij} r_{ij}$	
(5) Long-term plasticity	(a) $\dot{w}_{ij} = \mu v_i v_j$ (Hebb) (b) $\dot{w}_{ij} = \mu F(v_i, v_j, w_{ij})$	F: arbitrary polynomial
(6) Spike-timing-dep. plast. (Song et al. 2000)	$\dot{w}_{ij} = \begin{cases} \mu e^{-\Delta t/\tau_+} & \text{for } \Delta t \geq 0 \\ -\lambda e^{\Delta t/\tau_-} & \text{for } \Delta t < 0 \end{cases}$	$\tau_\pm$: decay times; Δt: spike-time difference
(7) Synaptic scaling (Tetzlaff et al. 2011)	$\dot{w}_{ij} = \mu (v_T - v_i) w_{ij}^n$	n: positive integer
(8) Structural plasticity (Butz et al. 2008)	$\Delta A_i = \mu (v_i - v_T) A_i$	A_i: # presyn. elements;
	$\Delta D_i = -\mu (v_i - v_T) D_i$	D_i: # postsyn. elements

The variables used throughout are as follows: w_{ij}, weight from neuron j to i; v_i, activity of neuron i; m_i, membrane potential; v_T, target activity; μ, λ, time scales; t_{sp}, time of spike; $\Theta(t)$, Heaviside function ($\Theta(t) = 1$ for $t > 0$ and $\Theta(t) = 0$ otherwise); $\delta(t)$, Dirac δ function; $\dot{x} = dx/dt$, time derivative of x; Δx, discrete change in x

synapse because then a postsynaptic potential (PSP) arises from each spike. The time scale of PSPs is on the order of tens of milliseconds. Thus, this is the simplest mechanism with which a certain time gap can be bridged. Although PSPs work on this very short time scale, the idea of using a slowly decaying signal for storage can be easily transferred to longer time scales of seconds to minutes. These so-called eligibility traces (Hull 1943; Rall 1967) are used in many approaches [e.g., Sutton (1988); Sutton and Barto (1998); Suri and Schultz (1999); Porr and Wörgötter (2003)] to bridge time gaps in learning algorithms, but the physiological basis of longer traces is unclear. A simple mathematical description for short traces is the alpha function (cf. Table 1—1), among others. Here, the time scale parameter μ directly relates to the temporal duration of the filter (e.g., $\mu = 0.01$ corresponds to 100 ms).

Single neurons themselves use temporal activity changes to encode memory. Intrinsic plasticity (Triesch 2007; Mozzachiodi and Byrne 2009; Gründemann and Häusser 2010) describes a mechanism that changes, based on the neuron's activity history, a neuron's input–output activity relation, thereby leading to distinct neuronal responses for different "memorized" activities. In Table 1—2 we state the original description by Triesch (2007). The physiological basis of the update rules for slope a and offset b is, however, unclear.

Many neurons combined in a network with recurrent connections exhibit so-called attractor dynamics (Hopfield 1982; Barbieri and Brunel 2008) attributed mainly to the hippocampus (Rolls and Kesner 2006; Lansner 2009; Cutsuridis and Wennekers 2009). This means that different activity patterns (e.g., incomplete input patterns) lead to the same (complete) activity pattern or cell assembly that can be interpreted as a memory item. Thus, to recall a memory item, it must be stored already in the synaptic weight matrix (see below). The corresponding update rule presented in Table 1—3 is Hebb-like (see below), and many approaches exist to overcome the original very simplistic assumptions [e.g., binary units, asymmetric weights; cf. Hertz et al. (1991)].

Another possible concept that relies merely on activities to represent memory and learning is reservoir computing (Vandoorne et al. 2008; Büsing et al. 2010) of which liquid state machines (Maass et al. 2002) and echo state networks (Jaeger and Haas 2004) are the most prominent. The abstract idea of this concept is that a complex network of calculating identities (e.g., neurons) is so diverse that each task is solved somewhere within the network (Maass et al. 2002; Buonomano and Maass 2009; Maass 2010). However, one problem with this approach is the capacity, which depends sublinearly on the number of neurons (Ganguli et al. 2008); another problem is the read-out of the task-specific information from the network (Maass et al. 2007; Legenstein et al. 2008).

4.2 Short-term plasticity

An experimentally well-established physiological theory is short-term plasticity (Markram et al. 1998), which has a time scale of milliseconds to seconds. In Tsodyks et al. (1998) the authors describe the short-term behavior of synapses with two coupled differential equations: the first describes the availability of vesicles and the second the probability of releasing the vesicles into the synaptic cleft. The arrival of spikes at the synapse leads to a calcium influx into the presynaptic terminal and, thus, to an increase in the release probability (synaptic facilitation). On the other hand, through activation of the synapse the number of available vesicles is reduced (synaptic depression). Two differential equations (cf. Table 1—4) that describe the connectivity changes from their baseline values (here R) can be used to model short-term plasticity (Mongillo et al. 2008).

The dynamics of short-term plasticity (also called dynamic synapses) in model neuronal networks are well reviewed (Barak and Tsodyks 2007; Marinazzo et al. 2007; Mejias and Torres 2009) and, in contrast to many other mechanisms, mathematically straightforward to analyze (Tsodyks et al. 1998; Bressloff 1999; Mejias and Torres 2009).

Short-term plasticity is thus a good candidate for temporally restricted (i.e., not long-lasting) memory mechanisms.

4.3 Long-term plasticity

When talking about plasticity and learning, most scientists think of Hebbian plasticity. In 1949 Hebb postulated, among other things, that (briefly) "neurons that fire together wire together" (Carla Shatz 1992). This statement is interpreted mainly in two different ways. First, a change in the wiring is possible only if both of the connected neurons are active and, thus, correlated (Oja 1982; Bienenstock et al. 1982). The second interpretation is that a change should depend only on information that is locally available, that is, the activity of the two neurons and the weight itself (Gerstner and Kistler 2002; Tetzlaff et al. 2011). For instance, for weight normalization (Rochester et al. 1956; von der Malsburg 1973) a process that takes into account weights of neighboring neurons is thus not possible for the latter. Both basic definitions of long-term plasticity are given in their simplest form in Table 1—5 a, b.

Independently of the interpretation of Hebbian plasticity this mechanism results in divergent dynamics: if an input drives the firing of a postsynaptic neuron, then plasticity potentiates the synaptic weight and, in this way, causes a stronger input drive, which in turn generates more potentiation, and so on. Neural systems have several different mechanisms (Bi and Poo 1998; Turrigiano et al. 1998; Royer and Paré 2003; Zhou et al. 2003), which are also used in theoretical studies (Rochester et al. 1956; Bienenstock et al. 1982;

Riedel and Schild 1992; Van Rossum et al. 2000) to avoid these runaway dynamics.

Hebb formulated his postulate in the middle of the last century, and it has long remained unknown whether such a mechanism really exists in the nerve cells of our brain. In 1973 Bliss and Lømo were the first to report a mechanism called long-term potentiation (LTP), which shows an increase in synaptic efficiency directly related to Hebbian plasticity. After the discovery of LTP, Lynch et al. in 1977 also found a reduction in synaptic efficiency, called long-term depression (LTD). Later in 1992 Dudek and Bear demonstrated both LTP and LTD at the same synapse. The fact that these synaptic changes due to LTP/LTD become permanent for several hours depends on a mechanism called synaptic tagging (Frey and Morris 1997; Clopath et al. 2008; Redondo and Morris 2011). Here, highly active synapses are tagged for long-lasting potentiation.

An influence of the temporal signal order on plasticity was proposed by Gerstner et al. (1996) and experimentally confirmed by Markram et al. (1997), Magee and Johnston (1997), and Bi and Poo (1998) [although Levy and Steward (1983) had already discovered this phenomenon]. These authors found that for the expression of LTP or LTD not only the activity as such matters but also the timing. Whenever there is a spike at a postsynaptic site *after* there was a spike at the presynaptic site, the strength of the synapse increases (LTP). However, if the timing is acausal, which means there is a postsynaptic spike *before* a presynaptic spike, the efficiency of the synapse decreases (LTD). To attain plasticity, spikes must be tenths of milliseconds apart. This is known as spike-timing-dependent plasticity (STDP). Synapses in different brain areas under different conditions exhibit experimentally different weight change curves, which are then fitted mostly as exponential kernels (Van Rossum et al. 2000; Froemke et al. 2005). A possible mathematical description is given in Table 1—6. The plasticity processes as such are related to LTP and LTD as discussed previously. However, the generality of STDP is heavily debated (Lisman and Spruston 2005, 2010; Shouval et al. 2010). Experiments show that the temporal order is only important in a small regime of presynaptic activation (Sjöström et al. 2001) and, furthermore, that synaptic modifications seem to be independent of the spiking of the postsynaptic cell (Golding et al. 2002). Thus, alternative models have been developed, for instance, calcium-based plasticity (Lisman 1989; Shouval et al. 2002; Yeung et al. 2004; Graupner and Brunel 2012).

A synaptic plasticity mechanism that seems to work on a longer time scale than minutes to hours is synaptic scaling (Turrigiano et al. 1998; Turrigiano and Nelson 2004; Rabinowitch and Segev 2006; Tetzlaff et al. 2011). Here, the synaptic efficiency of all inputs to a neuron increases/decreases if the activity of this neuron is lower/higher than a target activity (Davis 2006). These changes happen on a time

scale of days. Mathematically this can be described as shown in Table 1—7 with a weight dependence added for generality. It can be shown mathematically that adding synaptic scaling with quadratic weight dependence to generic plasticity rules stabilizes weight development (Tetzlaff et al. 2011).

4.4 Structural plasticity

Beside synaptic plasticity, which adapts the synaptic strength between neurons, structural plasticity decides which synapse is created and which is deleted (Holtmaat and Svoboda 2009; Caroni et al. 2012). Thus, by this mechanism the topology changes and the dynamics of the neuronal network can be modulated (Butz et al. 2009). The time scale of these changes is from hours to several days. In Table 1—8 we exemplify a model that describes the dynamics of the number (not the strength) of pre- and postsynaptic elements. The more elements exist, the higher is the synaptic efficiency.

To form new and abolish old synapses, the input and output compartments of a neuron, dendrites and axons, grow or retract depending on its average calcium or voltage level (Kater et al. 1989; Mattson and Kater 1989; Helias et al. 2008). Furthermore, synaptic plasticity influences these processes, and an important factor for rewiring is the synaptic efficiency (Levy and Desmond 1985; Adelsberger-Mangan and Levy 1994; Holtmaat and Svoboda 2009).

5 Links between time scales of memory and physiology

In this section we present experimental and theoretical links between memory concepts and their possibly underlying physiological mechanisms (Fig. 1). As there are many contributions for each aspect existing, this section could become exceedingly long. To avoid this, only key findings are discussed.

(1) *Working memory–activity (experimental)* In an experimental task, a specific set of neurons had ongoing activity during a delay period (Durstewitz et al. 2000; Funahashi et al. 1989) similar to eligibility traces. This ongoing activity can be perturbed by an external stimulus that also hinders the execution of the task (Fuster 1973).

(2) *Working memory–activity (theoretical)* Eligibility traces can be used as a very simple memory system (Suri and Schultz 1999; Porr and Wörgötter 2003); however, the physiological basis is not verified. Another model for working memory is reservoir computing, which uses the high complexity of recurrent neural networks (Maass et al. 2002). If these neural networks are stimulated by an input, then the complexity "stores" the input in the neural network's activity pattern for a certain duration. A similar approach are attractor networks (Barbieri and Brunel 2008), which, however, require a specific, biologically unconfirmed, signal to turn off working memory activity and previously stored knowledge.

(3) *Working memory–short-term plasticity (theoretical)* Short-term plasticity is a physiological candidate for working memory in neuronal networks (Barak and Tsodyks 2007; Mongillo et al. 2008). First, several neurons (a previously learned cell assembly) representing an item are preactivated by an external signal (e.g., think of the last digit of a phone number). By this activation, the synapses of these neurons have different release probabilities and vesicle numbers compared to the rest of the network. Afterward, within seconds, by a short activation of the whole network (e.g., thinking of digits) the item-representing neurons have significantly higher response properties and the item is thus recalled. However, there is only indirect experimental evidence for this link (Barak et al. 2010).

(4) *Working memory–short/long-term plasticity (theoretical)* An STDP rule for long- and short-term synaptic changes is used to create polychronous neural groups, i.e., groups of neurons that exhibit stereotypical time-locked spatiotemporal spike-timing patterns. These polychronous neural groups are then interpreted as working-memory items (Szatmáry and Izhikevich 2010). In attractor networks, the activation of previously learned activity patterns are also interpreted as working-memory items (Barbieri and Brunel 2008; Rolls 2010).

(5) *Long-/short-term memory–long-term plasticity (experimental)* For this link there exist much evidence [e.g., (Agranoff et al. 1966; Morris 1989; Rioult-Pedotti et al. 1998)], of which some, however, are contradictory (Martin et al. 2000). For instance, Mayford et al. (1996) showed in a learning task that the blockage of LTP in the hippocampus leads to a worse task performance compared to control. However, in a quite similar experimental setup the blockage of LTP leads to no significant differences (Zamanillo 1999). Regarding the neural correlate of memory, potential memory items were found, for example, in songbirds during song and recapitulation as stereotypical sequences of spike burst (Hahnloser et al. 2002) or during a spatial task, as spatiotemporal patterns (Harris et al. 2003) in the hippocampus. Additionally, typical connectivity structures, so-called motifs, were measured in the cortex (Song et al. 2005; Perin et al. 2011). However, it is not clear whether spatiotemporal patterns or motifs are memory items [or cell assemblies, see (6)] and whether they are subject to long-term plasticity as suggested by

Fig. 1 Overview of time scales of learning and memory on the one hand and of physiological mechanisms on the other hand and their links. The *number* next to a link indicates the corresponding section in the main text and *colors* indicate whether the link is theoretical (*blue*) or experimental (*red*). The black arrows depict the main memory dependence of learning. Color figure online
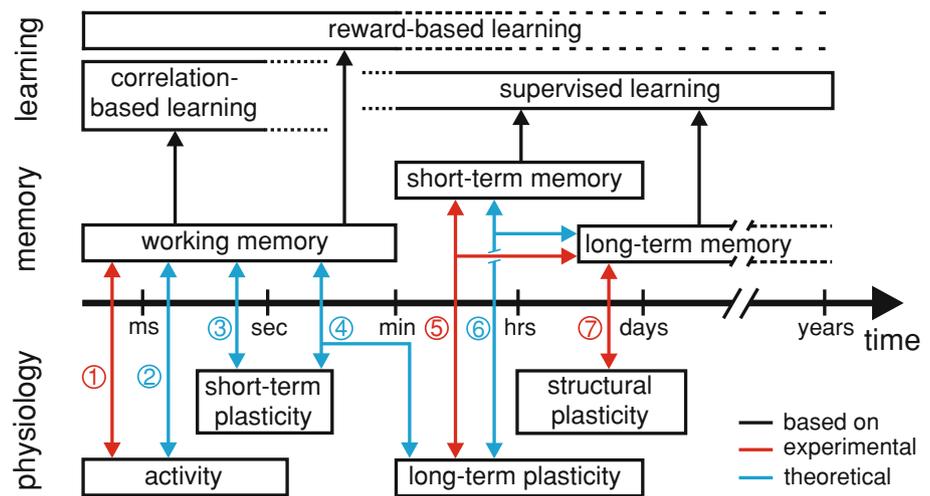
Hebb. Nevertheless, recent experimental work shows that the activation of specific groups of neurons elicits memory-related behavior (Garner et al. 2012; Liu et al. 2012; Morris and Takeuchi 2012).

(6) *Long-/short-term memory–long-term plasticity (theoretical):* As early as 1949 Hebb first speculated (Hebb 1949) that long-term plasticity should lead to memory. Hebb called these memory entities cell assemblies; highly connected cells should then represent a memory that could be elicited by stimulating only a part of the network. Attractor neural networks [e.g., Hopfield networks; (Hopfield 1982)] comprise the most prominent model for building cell assemblies with long-term plasticity. However, in these networks, after a memory item is activated, the activity of the corresponding neurons persists and the synaptic efficiencies that represent a memory item are in a "solid state," which is biologically unrealistic (Dudai 2004). In polychronization (Izhikevich 2006) spatiotemporal patterns can be interpreted as cell assemblies; however, selective storage has not yet been shown. As proposed by (Päpper et al. 2011), long-term plasticity, together with synaptic tagging, could form cell assemblies that pass through synaptic consolidation from a fragile to a more stable state.

(7) *Long-term memory–structural plasticity (experimental)* When rats experience enriched environments, their dendrites start to build significantly more spines than in a control environment (Xu et al. 2009; Yang et al. 2009). As spines are a direct indication of structural plasticity, these experiments support a link between learning and structural plasticity and—because the effect is not visible if rats had experienced the same enriched environment previously—between long-term memory and structural plasticity.

6 Why time scales?

The nervous system has by construction two very basic limitations: capacity and speed. If it stored all inputs, a very high memory capacity would be required that grew linearly with time. This would then lead to increased access time for memory retrieval [see, e.g., for Hopfield networks (Frolov and Husek 2000)]. The strategy of the nervous system to overcome these limitations apparently is to use different time scales in the memorization process. Hence, the nervous system is implicitly performing an algorithm by means of its finite storage time (i.e., by forgetting meaningless items) to *more efficiently* store and access data. But why does it use more than one finite time scale?

To answer this question, one should first consider why and how the nervous system forgets. Huge data sets exist on human forgetting [for a review see, for example, (Rubin and Wenzel 1996)] that show, as expected from everyday life, that we forget with higher probability as time passes. The exact shape of this forgetting curve is still under debate [exponential or power law; (Wixted and Ebbesen 1991; Anderson and Tweney 1997)]. However, forgetting may occur by several mechanisms, for instance, by noise and ongoing plasticity at the synaptic level (Fusi et al. 2005). Another possibility is the permanent overwriting or interference with old memories, which in turn reduces the capacity requirements of the network [see, e.g., (Mézard et al. 1986; Sikström 2002) for Hopfield models] but also deletes important memories. Processes on different time scales help to avoid such undesired interplays [e.g., the interaction between the hippocampus and the neocortex during synaptic consolidation; (Marr 1970; McClelland et al. 1995)]. Thus, different time scales could avoid the deletion of important information.

Furthermore, if different time scales interact, then one observes in robot experiments that more complex behaviors can be obtained. There exist many such examples of

interacting time scales in robotics of which we will mention briefly the following ones. In Floreano and Mondada (1996)) as well as in Ishiguro et al. (2003) evolutionary algorithms (longer time scale) are combined with plasticity mechanisms (shorter time scale) to cope with changing environments. In Steingrube et al. (2010) the adaptation of a chaotic controller (shorter time scale) is combined with an energy-reducing learning algorithm (longer time scale). In Manoonpong et al. (2012) a reactive learning algorithm (shorter time scale) is combined with a reinforcement learning algorithm (longer time scale) to increase total learning speed.

Thus, multiple time scales do indeed appear to be advantageous. On the other hand, it seems that organisms do not necessarily always make use of the potentially large time spans bridged by multiple scales. For example, no experience-based structural plasticity [(Chen et al. 2006; Varshney et al. 2011), long time scale] was found in the "simple" worm *Caenorhabditis elegans*, while it can still make use of synaptic plasticity [(Dabbish and Raizen 2011), shorter time scale].

This, and similar restrictions on time scales in other organisms, may well be due to the responses of the phyla to the evolutionary pressure they have experienced. For example, an organism that lives only for a few weeks will never "feel" the evolutionary pressure to develop long-term memories for planning food harvesting for the winter season. In other words, there exist many different short- or long-term correlations that may or may not be important for the organism in the physical world (Pitti et al. 2005). Thus, time scales of learning and memory probably only evolved if they were meaningful for the organism. The foregoing arguments also entail that, if the (relevant) environment for an organism has only a few time scales or the organism has only a limited behavioral repertoire, the underlying physiological mechanisms should also rely on only a few time scales.

7 Open questions

Up to this point we have only discussed links for which there is already experimental support or theoretical modeling. Nevertheless, to guarantee the correctness of these links, all of them must be tested according to the four criteria proposed by Martin et al. (2000) (detectability, mimicry, and anterograde and retrograde alteration). Furthermore, the following missing links should also be investigated in more detail.

First, all plasticity mechanisms we reviewed are investigated mainly for excitatory synapses. Only a few experimental studies have been done on inhibitory plasticity (Kilman et al. 2002; Woodin et al. 2003; Haas et al. 2006; Wenner 2011), and theoretical modeling is almost nonexistent (Vogels et al. 2011). It would be important to study the advantages and possibilities of inhibition so that more links to memory and learning could be made.

The theoretical modeling of activity and short- and long-term plasticity is more advanced than modeling of structural plasticity. One reason is that analytical models of topological changes in networks have not yet been developed. As a consequence, only computational models exist [cf. Table 1—8 and (Van Ooyen 1994; Poirazi and Mel 2001; Butz et al. 2009; Tetzlaff et al. 2010; Knoblauch et al. 2010)] and many of them originated from developmental neuroscience (van Ooyen 2011).

Although the separate study of structural plasticity is still difficult, the interplay of structural and long-term plasticity could give additional hints as to the discussion of the consolidation of short-term memory into long-term memory. To date, only a few studies have started to tackle this interplay (Adelsberger-Mangan and Levy 1994; Chklovskii et al. 2004).

In summary, it is interesting to note that there exist experimental support for links between (1) working memory and activity and (2) long-term memory and structural plasticity, but the very early proposed link between long-term memory and long-term plasticity (Hebb 1949) is still under debate [see, e.g., Martin et al. (2000)]. A detailed understanding of controversial experiments and their biological differences could lead to better understanding of the link between long-term plasticity and memory.

8 Summary

In this prospect we tried to argue that time scales are an inevitable property of our nervous system. In this context, we reviewed different memory and learning concepts and their underlying physiological mechanisms, always in relation to their time scales. Because it is not clear which mechanism is responsible for which behavior, we discussed possible links. This leads to many new questions, not only with respect to those links but also in research fields that use those mechanisms (e.g., robotics, artificial intelligence, machine learning).

References

Adelsberger-Mangan DM, Levy WB (1994) The influence of limited presynaptic growth and synapse removal on adaptive synaptogenesis. Biol Cybern 71:461–468

Agranoff BW, Davis RE, Brink JJ (1966) Chemical studies on memory fixation in goldfish. Brain Res 1(3):303–309

Anderson RB, Tweney RD (1997) Artifactual power curves in forgetting. Mem Cognit 25(6):867–872

Baddeley AD, Hitch G (1974) Working memory. In: Bower G (ed) The psychology of learning and motivation: advances in research and theory, vol 8. Academic, New York, pp 47–89

Barak O, Tsodyks M (2007) Persistent activity in neural networks with dynamic synapses. PLoS Comput Biol 3(2):e35

Barak O, Tsodyks M, Romo R (2010) Neuronal population coding of parametric working memory. J Neurosci 30(28):9424–9430

Barbieri F, Brunel N (2008) Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex?. Front Comput Neurosci 2(1):114–122

Bi G, Poo M (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and post-synaptic cell type. J Neurosci 18(24):10,464–10,472

Bienenstock E, Cooper LN, Munro P (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci 2(2):23–48

Bishop CM (2007) Pattern recognition and machine learning (information science and statistics). Springer, New York

Bliss TVP, Lømo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol 232(2):331–356

Braitenberg V (1984) Vehicles

Bressloff PC (1999) Mean-field theory of globally coupled integrate-and-fire neural oscillators with dynamic synapses. Phys Rev E 60(2):2160–2170

Brown J (1958) Some tests of the decay theory of immediate memory. Q J Exp Psychol 10(1):12–21

Buonomano DV, Maass W (2009) State-dependent computations: spatiotemporal processing in cortical networks. Nat Rev Neurosci 10(2):113–125

Büsing L, Schrauwen B, Legenstein R (2010) Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. Neural Comput 22:1272–1311

Butz M, Teuchert-Noodt G, Grafen K, van Ooyen A (2008) Inverse relationship between adult hippocampal cell proliferation and synaptic rewiring in the dentate gyrus. Hippocampus 18(9):879–898

Butz M, Wörgötter F, van Ooyen A (2009) Activity-dependent structural plasticity. Brain Res Rev 60(2):287–305

Carla Shatz J (1992) The developing brain. Sci Am 267: 60–67

Caroni P, Donato F, Muller D (2012) Structural plasticity upon learning: regulation and functions. Nat Rev Neurosci 13:478–490

Chen BL, Hall DH, Chklovskii DB (2006) Wiring optimization can relate neuronal structure and function. Proc Natl Acad Sci USA 103(12):4723–4728

Chklovskii D, Mel BW, Svoboda K (2004) Cortical rewiring and information storage. Nature 431:782–788

Clopath C, Ziegler L, Vasilaki E, Büsing, Gerstner W (2008) Tag-trigger-consolidation: a model of early and late long-term potentiation and depression. PLoS Comput Biol 4:e10000248

Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE (1997) Temporal dynamics of brain activation during a working memory task. Nature 386(6625):604–608

Cowan N (1995) Attention and memory: an integrated framework. Oxford University Press, New York

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 4:87–185

Cowan N (2005) Working memory capacity. Psychology Press, New York

Cutsuridis V, Wennekers T (2009) Hippocampus, microcircuits and associative memory. Neural Netw 22(8):1120–1128

Dabbish NS, Raizen DM (2011) GABAergic synaptic plasticity during a developmentally regulated sleep-like state in *C. elegans*. J Neurosci 31(44):15,932–15,943

Daneman M, Carpenter PA (1980) Individual differences in working memory and reading. J Verbal Learn Verbal Behav 466: 450–466

Davis GW (2006) Homeostatic control of neural activity: from phenomenology to molecular design. Annu Rev Neurosci 29:307–323

Diekelmann S, Born J (2010) The memory function of sleep. Nat Rev Neurosci 11(2):114–126

Duda RO, Hart PE, Stork DG (2000) Pattern classification. Wiley-Interscience, New York

Dudai Y (2002) Memory from A to Z: keywords, concepts, and beyond. Oxford University Press, Oxford

Dudai Y (2004) The neurobiology of consolidations, or, how stable is the engram? Annu Rev Psychol 55:51–86

Dudek S, Bear M (1992) Homosynaptic long-term depression in area ca1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. Proc Natl Acad Sci USA 89(10):4363–4367

Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. Nat Neurosci 3:1184–1191

Ericsson KA, Kintsch W (1995) Long-term working memory. Psychol Rev 102:211–245

Floreano D, Mondada F (1996) Evolution of plastic neurocontrollers for situated agents. In: Maes P, Mataric J, Meyer J-A, Pollack J, Wilson S (eds) From Animals to Animats 4: Proceedings of the 4th international conference on simulation of adaptive behavior. MIT Press, Cambridge, pp 401–411

Frankland PW, Bontempi B (2005) The organization of recent and remote memories. Nat Rev Neurosci 6(2):119–130

Frankland PW, Brien CO, Ohno M, Kirkwood A, Silva AJ (2001) α-CaMKII-dependent plasticity in the cortex is required for permanent memory. Nature 411:309–313

Frankland PW, Bontempi B, Talton LE, Kaczmarek L, Silva AJ (2004) The involvement of the anterior cingulate cortex in remote contextual fear memory. Science 304(5672):881–883

Frey U, Morris R (1997) Synaptic tagging and long-term potentiation. Nature 385:533–536

Froemke RC, Poo Mm, Dan Y (2005) Spike-timing-dependent synaptic plasticity depends on dendritic location. Nature 2033:2032–2033

Frolov AA, Husek D (2000) The convergence time in Hopfield network. In: Proceedings of the IEEE-INNS-ENNS international joint conference on neural networks, vol 5, Como, pp 622–626

Funahashi S, Bruce JC, Goldman-Rakic SP (1989) Mnemonic coding of visual space in the monkeys dorsolateral prefrontal cortex. J Neurophysiol 6(2):331–349

Fusi S, Drew PJ, Abbott LF (2005) Cascade models of synaptically stored memories. Neuron 45:599–611

Fuster JM (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. J Neurophysiol 36(1):61–78

Ganguli S, Huh D, Sompolinsky H (2008) Memory traces in dynamical systems. Proc Natl Acad Sci USA 105(48):18,970–18,975

Garner AR, Rowland DC, Hwang SY, Baumgaertel K, Roth BL, Kentros C, Mayford M (2012) Generation of a synthetic memory trace. Science 335:1513–1516

Gerstner W, Kistler WM (2002) Mathematical formulations of Hebbian learning. Biol Cybern 87:404–415

Gerstner W, Kempter R, Hemmen L , van Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. Nature 383:76–78

Golding NL, Staff NP, Spruston N (2002) Dendritic spikes as a mechanism for cooperative long-term potentiation. Nature 418:326–331

Graupner M, Brunel N (2012) Calcium-based plasticity model explains sensitivity of synaptic changes tp spike pattern, rate, and dendritc location. Proc Natl Acad Sci USA 109(10):3991–3996

Gründemann J, Häusser M (2010) A plastic axonal hotspot. Nature 465:1022–1023

Haas JS, Nowotny T, Abarbanel HDI (2006) Spike-timing-dependent plasticity of inhibitory synapses in the entorhinal cortex. J Neurophysiol 96(6):3305–3313

Hahnloser RHR, Kozhevnikov Aa, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419(6902):65–70

Harris K, Csicsvari J, Hirase H, Dragoi G, Buzsáki G (2003) Organization of cell assemblies in the hippocampus. Nature 424(6948):552–556

Hebb DO (1949) The organization of behavior: a neuropsychological theory. Wiley, Oxford

Helias M, Rotter S, Gewaltig MO, Diesmann M (2008) Structural plasticity controlled by calcium based correlation detection. Front Comput Neurosci 2(7):1–21

Henke K (2010) A model for memory systems based processing modes rather than consciousness. Nat Rev Neurosci 11:523–532

Hertz JA, Krogh AS, Palmer RG (1991) Introduction to the theory of neural computation. Westview, New York

Holtmaat A, Svoboda K (2009) Experience-dependent structural synaptic plasticity in the mammalian brain. Nat Rev Neurosci 10(9):647–658

Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational properties. Proc Natl Acad Sci USA 79:2554–2558

Hull CL (1943) Principles of behavior. Appleton Century Crofts, New York

Ishiguro A, Fujii A, Hotz PE (2003) Neuromodulated control of bipedal locomotion using a polymorphic CPG circuit. Adapt Behav 11(1):7–17

Izhikevich EM (2006) Polychronization: computation with spikes. Neural Comput 18(2):245–282

Izquierdo I, Barros DM, Melloe Souza T, de Souza MM, Izquierdo La, Medina JH (1998) Mechanisms for memory types differ. Nature 393(6686):635–636

Jaeger H, Haas H (2004) Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science 304(5667):78–80

Kater SB, Mattson MP, Guthrie PB (1989) Calcium-induced neuronal degeneration: a normal growth cone regulating signal gone awry (?). Ann NY Acad Sci 568:252–261

Kilman V, van Rossum MCW, Turrigiano GG (2002) Activity deprivation reduces miniature IPSC amplitude by decreasing the number of postsynaptic GABA(A) receptors clustered at neocortical synapses. J Neurosci 22(4):1328–1337

Knoblauch A, Palm G, Sommer FT (2010) Memory capacities for synaptic and structural plasticity. Neural Comput 22(2):289–341

Kumaran D (2008) Short-term memory and the human hippocampus. J Neurosci 28(15):3837–3838

Kyllonen PC, Christal RE (1990) Reasoning ability is (little more than) working-memory capacity?. Intelligence 14:389–433

Lansner A (2009) Associative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. Trends Neurosci 32(3):178–186

Legenstein R, Pecevski D, Maass W (2008) A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. PLoS Comput Biol 4(10):e1000,180

Levy WB, Desmond NL (1985) The rules of elemental synaptic plasticity. In: Levy WB, Anderson JA, Lehmkuhle S (eds) Synaptic modification, neuron selectivity, and nervous system organization, chapt 6. Lawrence Erlbaum, Hillsdale, pp 105–121

Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. Neuroscience 8(4):791–797

Lisman J (1989) A mechanism for the hebb and the anti-H ebb processes underlying learning and memory. Proc Natl Acad Sci USA 86:9574–9578

Lisman J, Spruston N (2005) Postsynaptic depolarization requirements for LTP and LTD: a critique pf spike-timing-dependent plasticity. Nat Neurosci 8(7):839–841

Lisman J, Spruston N (2010) Questions about STDP as a general model of synaptic plasticity. Front Synapt Neurosci 2:140

Liu X, Ramirez S, PangCB P T Puryear, Govindarajan A, Deisseroth K, Tonegawa S (2012) Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature 484:381–385

Lynch GS, Dunwiddie T, Gribkoff V (1977) Heterosynaptic depression: a postsynaptic correlate of long-term potentiation. Nature 266:737–739

Maass W (2010) Liquid state machines: motivation, theory, and applications. In: Cooper B, Sorbi A (eds) Computability in context: computation and logic in the real world. Imperial College Press, London, pp 275–296

Maass W, Natschläger T, Markram H (2002) Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Comput 14(11):2531–2560

Maass W, Joshi P, Sontag ED (2007) Computational aspects of feedback in neural circuits. PLoS Comput Biol 3(1):e165

Magee JC, Johnston D (1997) A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. Science 275(5297):209–213

Manoonpong P, Kolodziejski C, Wörgötter F, Morimoto J (2012) Combining correlation-based and reward-based learning in neural control for policy improvement. Adv Complex Syst (submitted)

Marinazzo D, Kappen HJ, Gielen SCaM (2007) Input-driven oscillations in networks with excitatory and inhibitory neurons with dynamic synapses. Neural Comput 19(7):1739–1765

Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275:213–215

Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical pyramidal neurons. Proc Natl Acad Sci USA 95(9):5323–5328

Marr D (1970) A theory of neocortex. Proc Roy Soc Lond B 176(1043):161–234

Marr D, Series L, Sciences B (1970) A theory for cerebral neocortex. Proc R Soc Lond B 176(1043):161–234

Martin SJ, Morris R (2002) New life in an old idea: the synaptic plasticity and memory hypothesis revisited. Hippocampus 12(5):609–636

Martin SJ, Grimwood PD, Morris RGM (2000) Synaptic plasticity and memory: an evaluation of the hypothesis. Annu Rev Neurosci 23:649–711

Mattson MP, Kater SB (1989) Excitatory and inhibitory neurotransmitters in the generation and degeneration of hippocampal neuroarchitecture. Brain Res 478(2):337–348

Mayford M, Bach ME, Huang YY, Wang L, Hawkins RD, Kandel ER (1996) Control of memory formation through regulated expression of a CaMKII transgene. Science 274(5293):1678–1683

McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampius and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102:419–457

McGaugh JL (2000) Memory—a century of consolidation. Science 287(5451):248–251

Mejias JF, Torres J (2009) Maximum memory capacity on neural networks with short-term depression and facilitation. Neural Comput 21(3):851–871, arXiv:0809.2010v1

Mézard M, Nadal JP, Toulouse G (1986) Solvable models of working memories. J Phys 47:1457–1462

Miller GA (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 63(2):81–97

Miller GA, Galanter E, Pibram KH (1960) Plans and the structure of behavior. Holt, Rinehart, Winston, New York

Mongillo G, Barak O, Tsodyks M (2008) Synaptic theory of working memory. Science 319(5869):1543–1546

Morris RG (1989) Synaptic plasticity and learning: selective impairment of learning rats and blockade of long-term potentiation in vivo by the n-methyl-D-aspartate receptor antagonist ap5. J Neurosci 9:3040–3057

Morris RGM, Takeuchi T (2012) The imaginary mind of mouse. Science 335:1455–1456

Mozzachiodi R, Byrne JH (2009) More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory. Trends Neurosci 33(1):17–26

Nadel L, Moscovitch M (1997) Memory consolidation, retrograde amnesia and the hippocampal complex. Curr Opin Neurobiol 7(2):217–227

Oja E (1982) A simplified neuron model as a principal component analyzer. J Math Biol 15(3):267–273

Owen AM (1997) The functional organization of working memory processes within human lateral frontal cortex : the contribution of functional neuroimaging. Neuroscience 9(7):1329–1339

Päpper M, Kempter R, Leibold C (2011) Synaptic tagging, evaluation of memories, and the distal reward problem. Learn Mem 18(1):58–70

Pastalkova E, Serrano P, Pinkhasova D, Wallace E, Fenton AA, Sacktor TC (2006) Storage of spatial information by the maintenance mechanism of LTP. Science 313:1141–1144

Pavlov PI (1927) Conditioned reflexes. Oxford University Press, London

Perea G, Navarrete M, Araque A (2009) Tripartite synapses: astrocytes process and control synaptic information. Trends Neurosci 32(8):421–431

Perin R, Berger TK, Markram H (2011) A synaptic organizing principle for cortical neuronal groups. Proc Natl Acad Sci USA 108(13):5419–5424

Pitti A, Lungarella M, Kuniyoshi Y (2005) Quantification of emergent behaviors induced by feedback resonance of chaos. In: Recent advances in artificial life, advances in natural computation, Chapt 15. World Scientific, Singapore, pp 199–213

Poirazi P, Mel BW (2001) Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. Neuron 29(3):779–796

Porr B, Wörgötter F (2003) Isotropic sequence order learning. Neural Comput 15:831–864

Rabinowitch I, Segev I (2006) The endurance and selectivity of spatial patterns of long-term potentiation/depression in dendrites under homeostatic synaptic plasticity. J Neurosci 26(52):13474–13484

Rall W (1967) Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input. J Neurophysiol 30:1138–1168

Redondo R, Morris R (2011) Making memories last: the synaptic tagging and capture hypothesis. Nat Rev Neurosci 12:17–30

Riedel H, Schild D (1992) The dynamics of Hebbian synapses can be stabilized by a nonlinear decay term. Neural Netw 5:459–463

Rioult-Pedotti MS, Friedman D, Hess G, Donoghue JP (1998) Strengthening of horizontal cortical connections following skill learning. Nat Neurosci 1(3):230–234

Rochester N, Holland J, Haibt L, Duda W (1956) Tests on a cell assembly theory of the action of the brain, using a large digital computer. IRE Trans Inf Theory 2:80–93

Rolls ET (2010) Attractor networks. WIREs Cognit Sci 1:119–134

Rolls ET, Kesner RP (2006) A computational theory of hippocampal function, and empirical tests of the theory. Prog Neurobiol 79(1):1–48

Royer S, Paré D (2003) Conservation of total synaptic weight through balanced synaptic depression and potentiation. Nature 422:518–522

Rubin DC, Wenzel AE (1996) One hundred years of forgetting: a quantitative description of retention. Psychol Rev 103(4):734–760

Shimizu E, Tang YP, Rampon C, Tsien JZ (2000) NMDA receptor-dependent synaptic reinforcement as a crucial process for memory consolidation. Science 290(5494):1170–1174

Shouval HZ, Bear MF, Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. Proc Natl Acad Sci USA 99:10,831–10,836

Shouval HZ, Wang SSH, Wittenberg GM (2010) Spike timing dependent plasticity: a consequence of more fundamental learning rules. Front Comput Neurosci 4:19

Sikström S (2002) Forgetting curves: implications for connectionist models. Cognit Psychol 45(1):95–152

Sjöström P, Turrigiano G, Nelson S (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron 32:1149–1164

Song S, Miller K, Abbott L (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat Neurosci 3:919–926

Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol 3(3):e68

Squire LR, Alvarez P (1995) Retrograde amnesia and memory consolidation: a neurobiological perspective. Curr Opin Neurobiol 5(2):169–177

Steingrube S, Timme M, Wörgötter F, Manoonpong P (2010) Self-organized adaptation of a simple neural circuit enables complex robot behaviour. Nat Phys 6(3):224–230

Stocker R, Durham WM (2009) Tumbling for stealth. Science 325:400–402

Suri RE, Schultz W (1999) A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. J Neurosci 91(3):871–890

Sutton R, Barto A (1981) Towards a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88:135–170

Sutton R, Barto A (1998) Reinforcement learning: an introduction. MIT Press, Cambridge

Sutton RS (1988) Learning to predict by the method of temporal differences. Mach Learn 3:9–44

Szatmáry B, Izhikevich EM (2010) Spike-timing theory of working memory. PLoS Comput Biol 6(8):e1000879

Tetzlaff C, Okujeni S, Egert U, Wörgötter F, Butz M (2010) Self-organized criticality in developing neuronal networks. PLoS Comput Biol 6(12):e1001013

Tetzlaff C, Kolodziejski C, Timme M, Wörgötter F (2011) Synaptic scaling in combination with many generic plasticity mechanisms stabilizes circuit connectivity. Front Comput Neurosci 5:47

Triesch J (2007) Synergies between intrinsic and synaptic plasticity mechanisms. Neural Comput 19(4):885–909

Tsodyks M, Pawelzik K, Markram H (1998) Neural networks with dynamic synapses. Neural Comput 10(4):821–835

Turrigiano GG, Nelson SB (2004) Homeostatic plasticity in the developing nervous system. Nat Rev Neurosci 5:97–107

Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature 391(6670):892–896

Van Ooyen A (1994) Activity-dependent neural network development. Comput Neural Syst 5:401–423

van Ooyen A (2011) Using theoretical models to analyse neural development. Nat Rev Neurosci 12(6):311–326

Van Rossum MCW, Bi GQ, Turrigiano GG (2000) Stable hebbian learning from spike-timing-dependent plasticity. J Neurosci 20(23):8812–8821

Vandoorne K, Dierckx W, Schrauwen B, Verstraeten D, Baets R, Bienstman P, Van Campenhout J (2008) Toward optical signal processing using photonic reservoir computing. Opt Express 16(15):11,182–11,192

Varshney LR, Chen BL, Paniagua E, Hall DH, Chklovskii DB (2011) Structural properties of the *Caenorhabditis elegans* neuronal network. PLoS Comput Biol 7(2):e1001066

von der Malsburg C (1973) Self-organization of orientation sensitive cells in the striate cortex. Kybernetik 14:85–100

Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W (2011) Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. Science 334(6062):1569–1573

Walker M, Brakefield T, Hobson J, Stickgold R (2003) Dissociable stages of human memory consolidation and reconsolidation. Nature 425:616–620

Wenner P (2011) Mechanisms of GABAergic homeostatic plasticity. Neural Plast. 2011:489,470

Whitlock JR, Heynen AJ, Shuler MG, Bear MF (2006) Learning induces long-term potentiation in the hippocampus. Science 313:1093–1097

Wixted JT, Ebbesen E (1991) On the form of forgetting. Psychol Sci 2:409–415

Woodin MA, Ganguly K, Poo M (2003) Coincident pre- and postsynaptic activity modifies GABAergic synapses by postsynaptic changes in Cl⁻ transporter activity. Neuron 39(5):807–820

Xu T, Yu X, Perlik AJ, Tobin WF, Zweig Ja, Tennant K, Jones T, Zuo Y (2009) Rapid formation and selective stabilization of synapses for enduring motor memories. Nature 462(7275):915–919

Yang G, Pan F, Gan WB (2009) Stably maintained dendritic spines are associated with lifelong memories. Nature 462(7275):920–924

Yeung LC, Shouval HZ, Blais BS, Cooper LN (2004) Synaptic homeostasis and input selectivity follow from a calcium-dependent plasticity model. Proc Natl Acad Sci USA 101(41):14943–14948

Zamanillo D (1999) Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. Science 284(5421):1805–1811

Zhou Q, Tao HW, Poo MM (2003) Reversal and stabilization of synaptic modifications in a developing visual system. Science 300:1953–1957

Ziv NE, Ahissar E (2009) New tricks and old spines. Nature 462:859–861