

Perspective

Mathematical models of learning and what can be learned from them

 Omri Barak¹ and Misha Tsodyks^{2,3}


Abstract

Learning is a multi-faceted phenomenon of critical importance and hence attracted a great deal of research, both experimental and theoretical. In this review, we will consider some of the paradigmatic examples of learning and discuss the common themes in theoretical learning research, such as levels of modeling and their corresponding relation to experimental observations and mathematical ideas common to different types of learning.

Addresses

¹ Rappaport Faculty of Medicine and Network Biology Research Laboratories, Technion - Israeli Institute of Technology, Haifa, Israel

² School of Natural Sciences, Institute for Advanced Study, Princeton, USA

³ Department of Brain Sciences, Weizmann Institute of Studies, Rehovot, Israel

Corresponding author: Tsodyks, Misha (mtsodyks@gmail.com)

Current Opinion in Neurobiology 2023, **80**:102721

This review comes from a themed issue on **Neurobiology of Learning and Plasticity 2023**

Edited by **Muming Poo** and **Thomas John McHugh**

For complete overview of the section, please refer the article collection - [Neurobiology of Learning and Plasticity 2023](#)

Available online 10 April 2023

<https://doi.org/10.1016/j.conb.2023.102721>

0959-4388/© 2023 Elsevier Ltd. All rights reserved.

What is learning?

Learning is a vital aspect of our life. Most of our activities, including social and professional, are learned throughout life. By reading this article, you will hopefully learn something. Nevertheless, defining learning is not a trivial task. Some types of learning are conceptually quite clear, such as “learning” how to ride a bicycle. “Learning” that you had bread for breakfast however, is a less obvious example. Even less obvious is “learning” that most English verbs acquire the ending “ed” when used in the past tense, or which words often appear together (e.g. “eat” and “food”) and which do not (e.g. “plane” and “animal”). In all cases, there is something you “know” (explicitly or implicitly) that

you did not know sometime in the past (before you started biking or before you learned English to a fluent level), and this new knowledge resulted from some kind of experience rather than inherited (see e.g. [1]). Often the result of learning is a new “memory” (such as the second example above), even though this is not always the case. In order to demonstrate that learning has indeed occurred, it is often desirable to see a change in behavior, and thus many definitions emphasize this aspect. In contrast, the abstract concept of knowledge lends itself to definitions involving the processing of information [2,3,4•]. Just as the borders between fast (what you had for breakfast) and slow (how to ride a bicycle) instances of learning in humans are not clear, so are the borders between learning in different systems. Learning is evident in virtually all the animal kingdom, with invertebrates providing some of the classic biochemical foundations for learning [5]. It is not only whole organisms that can improve their performance through experience. Natural selection can be described as a learning process, the immune system learns to identify pathogens, artificial networks learn how to play chess, and the list goes on. There are, of course, changes to systems that are less likely to earn the title of learning, such as leaving a dent in a table with a hammer strike—even though they involve changes to a system through experience that can be read out at a later time. The ease with which the concept of learning seems to encapsulate multiple phenomena on the one hand, and the difficulty of converging to a precise definition on the other was neatly captured by Dyal almost 50 years ago: “... the amount of effort ... debating the criteria for the existence of learning has been exceeded only by our philosophical brethren in their debates over the criteria for establishing the existence of god.” [6].

Instead of delving into this debate, we will adopt an operative approach. That is, we will study learning even if we do not agree on what it is precisely. Specifically, we are interested in cases where it is scientifically beneficial to consider different systems under the umbrella of learning. Such benefits can arise, for instance, when the same mathematical tool can be applied across different manifestations of learning. Looking for analogies

between systems can reveal blind spots—phenomena that are studied in one system and ignored in another.

We will try to illustrate the richness of the phenomenon of learning and the various modes in which experiments can interact with models of learning (see Figure 1). The review is structured around common themes in the study of learning. We hope this will encourage the reader to find places of interaction—between levels and systems—and that at the end of the day they will learn from it.

Paradigmatic examples

In this section, we briefly introduce three examples of learning, which will be used to illustrate the various themes: classical conditioning, associative memory, and free recall.

As mentioned above, to show learning has happened typically requires observing a change in behavior. One of the foundations of learning is behavioral change—*classical conditioning* [7]. At the most basic level, repeated temporal proximity of a ‘conditional stimulus’ (CS, e.g. a sound), which is originally neutral for the animal, and an ‘unconditional stimulus’ (US, e.g. food or electric shock), which has an inherent unconditional response (UR, e.g. freezing), results in the CS alone leading to the UR for the US with which it was paired. This is a very general example of “implicit” learning, i.e. the animal is not explicitly instructed or rewarded to exhibit the US response to CS. Nevertheless, one can reasonably argue that learning to associate different

stimuli that often appear together is useful in many situations (e.g. climbing a tree when you hear the leaves rustling and not waiting for the tiger to arrive).

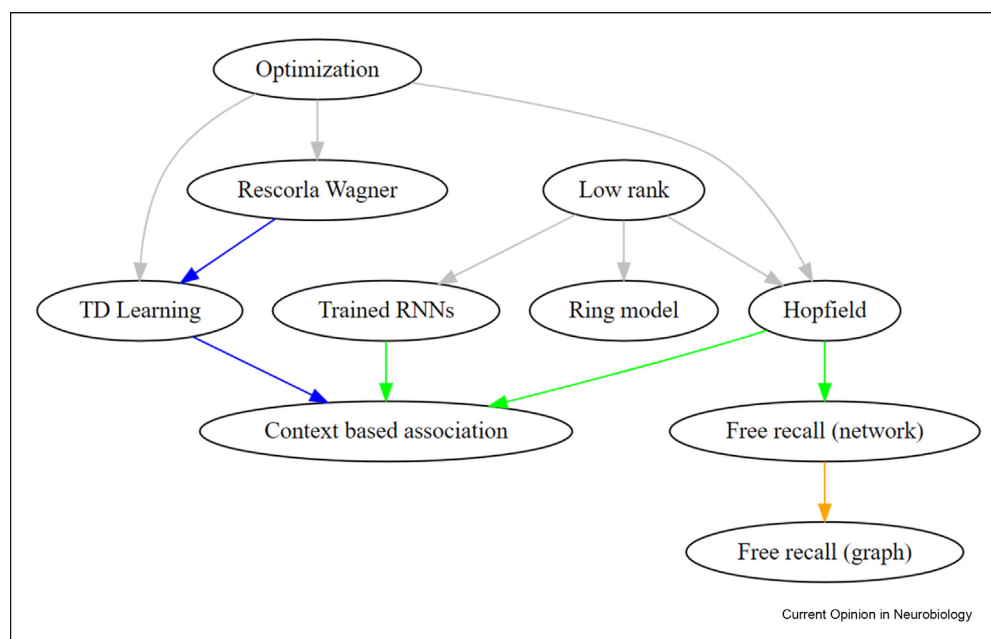
One of the prominent learning algorithms that were proposed to implement this learning is the famous Rescorla–Wagner theory [8]. This theory introduces an “associative strength,” w , between stimuli (CSs) and responses (US), which can be thought of as predicting the probability of occurrence of a US given the CSs. The change to associative strength follows the simple equation:

$$\begin{aligned}\Delta W &= \eta(r - V)u \\ V &= w^T u\end{aligned}\quad (1)$$

where r is whether the US was present at a given trial, u is a binary vector (0,1) denoting the presence or absence of stimuli, and V is the predicted presence of the US. These equations predict, for instance, that if one of the CSs (say u_1) was previously paired with the US such that it fully predicts its appearance (i.e. $w_1 = r$), the addition of another CS (u_2) will not result in any new association, i.e. w_2 will stay at zero. This effect is well known and is called “blocking” [9].

A second example is *associative recall*, in which a partial cue can trigger the recall of an item in memory. This has been studied in tasks, such as visual memory [10]. A classical model to describe this phenomenon is the Hopfield model [11], in which a recurrent neural

Figure 1



The concept of learning encompasses many systems, and insights can be drawn from many sources. Arrows denote types of links: abstraction (orange), elaboration (blue), component (green), and mathematical (gray).

network with N neurons is created with the p desired memories as fixed points of the dynamics. Specifically, the dynamics are given by

$$\begin{aligned} r(t+1) &= \varphi(Jr(t)) \\ J &= \frac{1}{Np} SS^T \end{aligned} \quad (2)$$

where φ is a threshold nonlinearity, and the connectivity J is an outer product of the desired memories S (an $N \times p$ matrix). Initializing the network at a state that has a partial overlap with one of the memories leads to convergence to the corresponding fixed point.

A third example comes from the paradigm of *free recall*, where subjects learn lists of randomly assembled items, such as single words or short sentences for subsequent recall (see e.g. [12]). Numerous important results were obtained on this topic over the century of research, but we will focus here on the issue of performance in terms of the number of items that can be recalled from a list of a given length. The consensus in the field was reached that the average performance increases with the list length (i.e. there is no upper bound on the number of items recalled); however, the rate of increase is decelerating, such that smaller and smaller proportion of items can be recalled [13–15], with power-law relation between the recall and list length as a plausible behavior [16]. It was recently shown [17] that if M is the number of items retained in memory (independent of the ability to recall them), a phenomenological theory results in the universal, parameter-free relation between the number of items in memory (M) and the average number of recalled items (R) that closely matched experimental results of [17]:

$$\langle R \rangle = \sqrt{\frac{3\pi}{2}} M \quad (3)$$

To derive this simple formula, the authors assumed that the recall is fully controlled by a random symmetric matrix of size $M \times M$ that represents the pair-wise similarity in item internal representations (e.g. overlaps in their neuronal encoding in long-term memory). In particular, each time a particular item is recalled, the next item is chosen to be the one with the largest similarity measure to the current one (excluding the item that was recalled just before the current one). This elementary and abstract recall algorithm, which is assumed to be completely deterministic, invariably leads to the recall trajectory across the items that after some number of steps converges to a cycle, when the same items are recalled again and again, shielding the remaining items from ever being recalled. As a result, the average length of recall trajectory (i.e. the number of items recalled) can be calculated mathematically and

shown to quickly converge to the expression 3 as M increases.

Levels of modeling and their interaction with biology

All models are an abstraction of reality. The level of abstraction, however, varies between models and also partially determines the mode of interaction between the model and experiments.

At the behavioral level, the Rescorla–Wagner model offers an explanation for phenomena such as blocking and shadowing that were absent from previous accounts. Testing the model is, therefore, naturally done at the behavioral level. It is this behavioral level that also challenged the model, primarily by its inability to explain secondary conditioning. This phenomenon occurs when a second CS, which was never seen together with the US, predicts it due to its association with the first CS. This failure prompted extensions that consider the temporal proximity of stimuli, such as temporal difference (TD) learning [18]. In this model, as opposed to equation (1), all variables are taken to depend on time during the trial, introducing the dependence of the learning on temporal aspects of the stimuli. The modification in the associative strengths w is proportional to the TD error as

$$\begin{aligned} \delta_t &= r_t + V_t(u_t) - \gamma V_t(u_t - 1) \\ V_t(u) &= w_t^T u \end{aligned} \quad (4)$$

where γ is a discounting factor between 0 and 1. The modification of associative weights is determined by

$$\begin{aligned} w_{t+1} &= w_t + \eta \delta_t e_t \\ e_t &= \lambda e_{t-1} + u_t \end{aligned} \quad (5)$$

with e_t standing for *eligibility traces* that integrate inputs over time with decay constant λ , and η defines the rate of learning. This model accounts for secondary conditioning, because if one of the CSs (say, u_1) is consistently presented before the other CS (u_2) that was previously paired with US and hence developed a strong associative weight w_2 , the TD error is positive at the onset of u_2 and hence drives the associative strength w_1 up even if no US are presented at these trials.

While the model is phrased in an abstract manner, it lends itself to a simple mechanistic interpretation. The association strength corresponds to a group of synapses, and these are modified by a simple correlative Hebbian rule: $\Delta w \propto xy$. In order to provide stability, saturation is also added. The mechanistic aspect suggests another link between the model and experiments—at the neuronal level. This link manifested most clearly when studying the extended model—TD learning. Here, w_t is

a function of time since the stimulus, and the error that drives learning is that of the expected future reward (in the case of positive US). Long after the model was suggested, dopamine neurons with a strong resemblance to this reward prediction error were discovered [19]. We see that this model connected behavioral level observations to a computational model and to a neural correlate of these computational elements. In recent years, the link at the neural level is making its own contributions back to the computational level. Specifically, the model assumes that the prediction error is a scalar—because the expected reward is a scalar, and so is the deviation from it [20]. Dopamine neurons, on the other hand, were recently shown to be more heterogeneous [21]. This observation has led to proposing a vector representation of the prediction error [22].

The free recall model also offers two degrees of abstraction. While the model described above is naturally expressed as a trajectory along an abstract graph, it also affords a natural implementation in Hopfield-like networks if a mechanism for transitions between attractors is added, such as neuronal adaptation and/or modulation of inhibition [23,24•]. [24•] studied such networks for their recall properties. Loading some items in short-term memory was achieved by an additional strengthening of synaptic connections between the neurons encoding these items. It was indeed observed that if network dynamics are deterministic, under some conditions, the network enters a cyclic activity after retrieving some of the attractors, and the number of retrieved attractors scaled sublinearly with the number of stored attractors, which could be approximated by power-law. The exponent of the power-law scaling depends on the parameters of the network but generally was found to be above 0.5. Moreover, if a weak noise is added to the network, the exponent is increasing so that more items are recalled.

Comparison between these two models raises some important issues on the level of theorizing that is most appropriate for making contacts with experiments. The phenomenological theory provides a very good and robust match to data, indicating that it captures at least some of the basic processes occurring in the brain during recall. However, beyond the behavioral performance, it is very difficult to directly relate this model to any type of recordings. Moreover, some of the assumptions of the theory appear downright unrealistic. Most strikingly, the assumption of deterministic dynamics is contradicting the common observations of the extremely stochastic nature of neural activity observed in the data. In this respect, the network model appears to be much more realistic; however, given its many parameters, it cannot account for the data in a robust way. In the face of this dichotomy, one should either admit that no realistic models of cognitive phenomena that would at the same time robustly match the data are possible or that

realistic models should not be attempted at the level of neurons (see, however, recent attempts to close this gap [25,26]).

Normative versus mechanistic

The classical conditioning models also offer a case study of normative versus mechanistic approaches. As stated above, the RW rule can be seen as a Hebbian correlation-based strengthening of links. A different view is obtained when considering a normative goal—minimizing the prediction error. If we consider this error $\frac{1}{2}(r - v)^2$ and then update weights according to gradient descent, we obtain the very same rule. Viewing the problem through the lens of optimization offers a way to both analyze learning rules and suggest new ones. This is true not only for reinforcement learning but also for learning associative memories. The synaptic strengths in the classic Hopfield model can be obtained by a Hebbian updating rule $\Delta W = S_\mu S_\mu^T$, where S_μ is the memory to be stored. This rule is also the result of gradient descent on a cost function $S^T W S$, where S is the matrix whose columns are S_μ . This cost function measures the angle between the S_μ and $W S_\mu$, two terms that should be correlated to obtain the desired fixed points. Viewing the problem from this perspective allows defining multiple measures of distance as a cost function, each one leading to a different (and usually well known) learning rule[27•].

It remains to be seen whether the free recall model also admits an energy-based optimization view.

Mathematical tools in learning

The examples above showed two cases where learning rules can be formulated either at the mechanistic level or as steps in an optimization procedure. This is, perhaps, not surprising as learning can be defined as an improvement in the performance of some task. Thus, optimization is a natural mathematical language for its description. Indeed, mathematical insights from the optimization viewpoint appear in many learning models. As mentioned above, optimization often manifests as moving down an energy gradient. This is expected to converge to the best solution only in simple scenarios—a single global minimum of a convex function. The Hopfield model represents one deviation from this scenario—there are local minima, representing spurious attractors. Because these minima are shallower than the global ones (corresponding to memories), adding noise to the dynamics can improve performance. This picture of several local minima does not necessarily hold in more complex energy landscapes. Gradient descent may be delayed by saddle points, rather than trapped in local minima [28]. Perhaps more important is the fact that in over-parametrized networks, there are many global minima—an infinite number of them arranged in manifolds. This picture implies that learning algorithms

implicitly bias convergence to certain minima and not others. The similarity of learning in the brain implies that such biases may be at play in neuroscience as well. Indeed, recent work shows that efficient coding and psychophysical biases can be explained by this effect [29].

Another mathematical hallmark of learning is low-rank connectivity. This type of connectivity implies that while all neurons can be connected to each other, they effectively do so via a small number of channels. Networks that were designed for specific tasks, such as the Hopfield model or the ring model, are often endowed with low-rank connectivity. In the Hopfield case, for instance, the effective channels are the memorized patterns. The field of reservoir computing uses a “trick” of only training readout weights and feeding the output back into the network [30,31]. This is effectively a rank-1 perturbation to the connectivity, where the network’s output is the effective channel. It was recently shown that training the full connectivity matrix of networks without any constraints, nevertheless, leads to a low-rank perturbation [32•].

Learning outside the brain

The ubiquity of the phenomenon of learning suggests that insights can be gained by also looking outside of the brain. This can be seen, for instance, within the optimization framework. Gradient descent algorithms move along the steepest direction. In the brain, it seems implausible that synapses can compute exact gradients, as this would require to propagate learning signals backward with weights that are exactly the same as the forward ones used in computation. It turns out, however, that approximate gradients also suffice to increase performance. For instance, using random weights to back-propagate the learning signal is sufficient to reach good solutions. This is because the forward weights have enough freedom to both perform the task and align to the random backward weights, thereby making the learning steps correlated with (but not equal to) the true gradient [33]. In recurrent networks, there are also families of rules that are only correlated to the true gradient [34]. Such convergence with approximate gradients is well known from the phenomenon of bacterial chemotaxis. Bacteria climb the nutrient gradient with only two actions at their disposal—continue straight, or tumble around and exit in a random direction. By increasing the probability of tumble with decreasing concentration, bacteria move in a direction that is correlated with the true gradient. An even weaker form of learning was proposed to explain the adaptation of yeast cells to unforeseen challenges. In this scenario, stress causes a random walk in the weights space of the gene regulatory network [35]. Remarkably, this procedure only works for networks with a heavy-tailed distribution of outgoing weights—a topology that was

found to be equivalent to a low-rank perturbation [36]. It is worth emphasizing that learning is not always for the better, as in the case of cancer cells, which seem to adapt to challenges by mechanisms that are highly reminiscent of learning [37]. Finally, recent studies have shown analogs of learning in complex systems outside biology, such as mechanical spring networks [38,39], and it remains to be seen which principles will translate between these domains.

Context

Just as the behavioral level properties of classical conditioning prompted the Rescorla–Wagner model and its extension into the TD model, there are behavioral observations that are at odds with both of them. A prominent one is extinction: it turns out that it does not constitute a complete erasure of conditioning, rather it sometimes reappears either spontaneously or after a brief additional exposure to CS [40]. These and other observations indicate that describing the animal’s internal model of the situation with one variable (the strength of CS-US association) is inadequate, rather it appears that the animal keeps both learned behaviors (with and without association) in memory and can rapidly switch from one to another. This interpretation, surprising as it may sound, can be also reasoned to be important for behavior. For example, when a tourist comes to London, she has to adapt to the left-side traffic in order to escape being hit by cars. It would be unreasonable to assume though that the tourist completely forgets the traffic condition in her home country. Instead, she will quickly switch back to her usual behavior almost immediately upon returning from London.

A phenomenological generative theory of conditioning that is more compatible with the idea of simultaneous representation of competing learned behaviors was proposed by Gershman et al. [41]. According to this theory, instead of simply learning to approximate the conditional probabilities between the appearance of CS and US, an animal assumes that the statistics of both of them are defined by a generative process with hidden, mutually exclusive, “latent causes” that are not directly perceived but have to be inferred from sensory observations. In this theory, extinction is accompanied by learning a new conjunction between the CS and US (that CS is not followed by US) that corresponds to the new hidden cause added to the model in the animal’s brain. Crucially, this new hidden cause is *added* to the previously learned one (according to which CS is followed by US) and does not erase it. At the end of extinction, two competing hidden causes coexist, and the animal can switch from one to another without new learning, similar to how a tourist does not have to learn right-side traffic after returning from a trip to London. As described in [41], this phenomenological theory

accounts for many (but not all) salient features of classical conditioning observed over the years, indicating that it may correctly capture some underlying processes responsible for conditioning. However, as opposed to the Rescorla–Wagner model, it does not have an obvious biophysical implementation. In particular, it is not obvious how two different behaviors could coexist in a neural network with fixed connectivity.

Two examples, using designed and trained networks implement elements of context-based behaviors. Blumenfeld et al. [42] modified the Hopfield model so that plasticity is modulated by a novelty signal—similar to the abstract suggestion in [41]. Mante et al. [43] studied a decision task where a monkey has to respond to a stimulus, consisting of randomly moving dots of two different colors, by moving the gaze according to either the dominant direction of movement or the dominant color, depending on the contextual signal presented to the monkey concurrently with the moving dots. Instead of building in the mechanism for solving this task into the connectivity of the network, an initially random recurrent network was trained by optimizing its connectivity until the network could solve the decision task at the levels comparable to those of monkeys. After training, it was found that population activity in the network was qualitatively similar to electrical recordings from the Prefrontal Cortex of monkeys. Crucially, the contextual input, indicating whether the decision should be made based on motion or color signals, was applied simultaneously with the sensory input. This is unlike our hypothetical tourist in London, who infers context from stimuli that change on much slower time scales than the sensory inputs. Hence, having a context-dependent input would mean that somewhere in the brain we have persistent activity of context-encoding neurons, which does not seem reasonable for very long durations.

An attempt to propose a biologically realistic model of fear conditioning was made at [44]. Instead of having one synapse (or a homogeneous group of synapses) to encode the association strength between CS and US, the model assumes a heterogeneous recurrent network of excitatory and inhibitory neurons in Amygdala receiving inputs from auditory thalamus carrying the CS (auditory tone) and inputs from the cortex carrying the US (electric shock), as well as non-specific spontaneous inputs. All the recurrent and afferent connections undergo activity-dependent plasticity of a Hebbian type. At the baseline state, the application of CS leads to weak activation of pyramidal neurons which is assumed to not result in an unconditional response (freezing). During conditioning, both afferent and recurrent inputs to pyramidal neurons in the network become stronger, leading to their stronger response, which is assumed to result in the unconditional response. Importantly, also inputs to inhibitory neurons become stronger, but the

effect of inhibition is not strong enough to suppress the enhanced pyramidal response. When extinction phase begins, connections targeting pyramidal cells weaken, and their response to CS is correspondingly weakened to pre-conditioning levels; hence, no US response to CS emerges. However, the state of the network after extinction is different from the baseline, namely connections to both pyramidal and inter-neurons are stronger, and hence the network can easily recover conditioning due to spontaneous activity in the network. While representing an interesting attempt at biological modeling of conditioning, this model does not allow for rapid switches between two regimes since the time it takes for the network to spontaneously recover the conditioning response is quite large, and the model was also not reproducing multiple other effects that were experimentally observed.

The coexistence of different learned behaviors discussed above is not specific to classical conditioning but relates to a general feature of biological learning. Artificial neural network approaches to learning suffer from “catastrophic forgetting,” i.e. they cannot efficiently learn multiple tasks [45]. When the network is trained for a new task, its performance on a previously learned task tends to drastically suffer. This contrasts sharply with our ability to exhibit different behaviors that apparently involve the same cortical networks. For example, the same soccer player tries to put the ball into the net when attacking the opposing team but does his best to prevent the same event from happening when defending [46]. It is obvious that the player has no time to learn a new skill each time the game switches from attack to defense. More generally, it appears that many skills (like biking or playing piano), when properly mastered, are extremely resilient to very long periods in which they are not practiced. In one recent example, subjects were playing computer games, and their performance was tracked over the time course of years [47•]. Interestingly, even after a very long period of no practice, the subject’s level quickly (often after just one trial) reached the level of performance from the last session. Several approaches were put forward to deal with this issue. In one of them [48], it is proposed that when training the network for a particular task, each synapse not only acquires a particular strength but also “computes” the information about how sensitive the performance of this task is to changes in the strength (how “important” this synapse is for the task). When the subsequent task is learned, the network tries to only modify synapses that are less “important” for the previous tasks, keeping the more important synapses stable (“synaptic consolidation”), resulting in much-reduced interference. The authors illustrated this idea with training a feed-forward network on a sequence of tasks of distinguishing pairs of consecutive digits (say, ‘1’ from ‘2’, then ‘2’ from ‘3’, etc). With the standard stochastic gradient descent technique, by the time the last task is

learned, the performance on the first one is almost reduced to chance, whereas the new learning algorithm leads to close to perfect execution of all the tasks. The estimation of the importance for each synapse concurrent with stochastic gradient descent for synaptic strengths entails quite a complex set of calculations that go well beyond anything that is currently known about synaptic plasticity experimentally. On the other hand, as the authors correctly note, synaptic transmission and plasticity in the brain utilize an intricate molecular machinery on different temporal and spatial scales, and are vastly more complex than simple synaptic description with one parameter (strength) that has a simple modification algorithm. Whether some of this unexplored (and unknown) machinery is responsible for “intelligent” synapses utilized in [48] is of course an open question.

The studies described above illustrate that having a biologically realistic model of learning that allows for multiple coexisting behaviors is challenging. The most natural explanation would be to do away with the idea of single weight for each synaptic connection and assume that connections can be characterized by multiple context-dependent weights, but this assumption goes beyond anything that we currently know about synapses in the brain. We are left with the dichotomy: either the difficulties in proposing realistic models of learning are due to some yet unknown phenomena in the brain, or we should try harder to develop better models with our current knowledge about the neuronal hardware. One under-explored direction, in particular, could be the effects of neuromodulators on information processing (see e.g. [49,50]). Neuromodulators influence processing in many ways, including by changing the neuronal excitability and synaptic strengths [51], which could potentially underlie quick changes in network states without the need for synaptic plasticity (see e.g. [52]) or directly affect learning in neural circuits (see [53] for a recent review on putative effects of neuromodulation on learning).

Conclusions

Despite decades of studies, learning remains a mysterious trait. In this brief review, we touched on some of the most puzzling observations and their theoretical descriptions. Many more learning phenomena exist that were not discussed here. To mention just one, there is a vast difference in individual learning abilities in particular skills, such as music and math. To some degree at least, this is genetically predetermined, but how are these differences expressed in neuronal hardware? Are there specialized “math” neurons/regions in the brain? The first evidence for this begins to emerge [54], but we are clearly just at the beginning of the road. We discussed in some detail the ability to learn multiple behaviors in similar situations and rapidly switch

between them depending on the context. While there are good phenomenological models of this phenomenon, translating them into biologically realistic network models is still challenging. This raises an important issue of levels of modeling, in particular, which ones are most suitable for comparing them with experiments. Some behavioral observations can be described quite precisely in phenomenological models (such as free recall), but translating phenomenological models into realistic implementation requires much tuning and hence appears to be not robust. This may imply that there are some yet unknown mechanisms in the brain that guarantee neural networks can robustly perform desired tasks. We, therefore, believe that designing robust realistic models of learning is the main direction for future research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

We thank Dr. Mikhail Katkov for helpful comments on the manuscript. OB is supported by the Israel Science Foundation grant 1442/21 and an HFSP research grant (RGP0017/2021). MT is supported by the Israel Science Foundation grant 1657/19 and Foundation Adelis.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- 1. Gross C: *Psychology: the science of mind and behavior*. London: Hodder Education; 2010.
- 2. De Houwer J, Barnes-Holmes D, Moors A: **What is learning? On the nature and merits of a functional definition of learning.** *Psychonomic Bull Rev* 2013, **20**:631–642.
- 3. Barron AB, Hebets EA, Cleland TA, Fitzpatrick CL, Hauber ME, Stevens JR: **Embracing multiple definitions of learning.** *Trends Neurosci* 2015, **38**:405–407.
- 4. De Houwer J, Hughes SJ: **Learning in individual organisms, genes, machines, and groups: a new way of defining and relating learning in different systems.** *Perspect Psychol Sci* 2022.
- A review that highlights similarities between learning in different systems. The relation to the version from 2013 is interesting.
- 5. Coming WC, Dyal JA, Willows AOD: *Invertebrate learning*. Boston, MA: Springer US; 1975.
- 6. Dyal JA, Coming WC: **Invertebrate learning and behavior taxonomies.** In *Invertebrate learning*, W.C. Coming. Edited by Dyal JA, Willows AOD, Springer US; 1973:1–48.
- 7. Pavlov I: *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford Univ. Press; 1927.
- 8. Rescorla RA, Wagner AR: **A theory of Pavlovian conditioning: variations on the effectiveness of reinforcement and**

- nonreinforcement.** In *Classical conditioning II: current research and theory*. Edited by Black AH, Prokasy WF, New York: Appleton-Century-Crofts; 1972:64–99.
9. Kamin LJ: **Predictability, surprise, attention, and conditioning.** In *Symp. on punishment*. Edited by Campbell BA, Church RM, New York: Appleton-Century-Crofts; 1969:279–296.
 10. Bosch SE, Jehee JFM, Fernández G, Doeller CF: **Reinstatement of associative memories in early visual cortex is signaled by the Hippocampus.** *J Neurosci* 2014, **34**:7493–7500 [Publisher: Society for Neuroscience Section: Articles].
 11. Hopfield JJ: **Neural networks and physical systems with emergent collective computational abilities.** *Proc Natl Acad Sci USA* 1982, **79**:2554.
 12. Kahana MJ: *Foundations of human memory*. OUP USA; 2012.
 13. Binet A, Henri V: **Memoire des mots.** *L'Année psychologique* 1894, **1**:1–23.
 14. Standing L: **Learning 10000 pictures.** *Q J Exp Psychol* 1973, **25**: 207–222.
 15. Murray D, Pye C, Hockley W: **Standing's power function in long-term memory.** *Psychol Res* 1976, **38**:319–331.
 16. Waugh NC: **Presentation time and free recall.** *J Exp Psychol* 1967, **73**:39.
 17. Naim M, Katkov M, Romani S, Tsodyks M: **Fundamental law of memory recall.** *Phys Rev Lett* 2020, **124**. 018101.
 18. Sutton RS, Barto AG: *Reinforcement learning: an introduction, volume 1*. Cambridge Univ Press; 1998.
 19. Schultz W, Dayan P, P.R. M: **A neural substrate of prediction and reward.** *Science* 1997, **14**:1593–1599.
 20. Glimcher PW: **Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis.** In *Proceedings of the national academy of sciences*, **108**. Publisher: Proceedings of the National Academy of Sciences; 2011: 15647–15654.
 21. Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, et al.: **Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons.** *Nature* 2019, **570**:509–513. Number: 7762 Publisher: Nature Publishing Group.
 22. Lee RS, Engelhard B, Witten IB, Daw ND: **A vector reward prediction error model explains dopaminergic heterogeneity.** *bioRxiv* 2022. 2022.02.28.482379.
 23. Romani S, Pinkoviezky I, Rubin A, Tsodyks M: **Scaling laws of associative memory retrieval.** *Neural Comput* 2013, **25**: 2523–2544.
 24. Boboeva V, Pezzotta A, Clopath C: **Free recall scaling laws and short-term memory effects in a latching attractor network.** *Proc Natl Acad Sci USA* 2021, **118**, e2026092118.
- A network model of free recall implemented as latching dynamics of neural activity
25. Finkelstein A, Fontolan L, Economo MN, Li N, Romani S, Svoboda K: **Attractor dynamics gate cortical information flow during decision-making.** *Nat Neurosci* 2021. 1–8Publisher: Nature Publishing Group.
 26. Chen G, Scherr F, Maass W: **A data-based large-scale model for primary visual cortex enables brain-like robust and versatile visual processing.** *Sci Adv* 2022, **8**, eabq7592 [Publisher: American Association for the Advancement of Science].
 27. Tolmachev P, Manton JH: **New insights on learning rules for Hopfield networks: memory and objective function minimisation.** In *2020 international joint conference on neural networks (IJCNN)*. IEEE; 2020:1–8.
- An elegant connection of normative and mechanistic views of learning in Hopfield networks.
28. Dauphin YN, Pascanu R, Gulcehre C, Cho K, Ganguli S, Bengio Y: **Identifying and attacking the saddle point problem in high-dimensional non-convex optimization.** *Adv Neural Inf Process Syst* 2014, **4**.
 29. Benjamin AS, Zhang LQ, Qiu C, Stocker A, Kording KP: *Efficient neural codes naturally emerge through gradient descent learning.* *bioRxiv*; 2022.
 30. Jaeger H: *The "echo state" approach to analysing and training recurrent neural networks-with an erratum note*. German National Research Center for Information Technology; 2001. Technical report, Technical Report GMD Report 148.
 31. Sussillo D, Abbott LF: **Generating coherent patterns of activity from chaotic neural networks.** *Neuron* 2009, **63**:544–557.
 32. Schuessler F, Mastrogiuseppe F, Dubreuil A, Ostojic S, Barak O: **The interplay between randomness and structure during learning in RNNs.** *Adv Neural Inf Process Syst* 2020, **33**. The authors show that low-rank connectivity naturally emerges via training, and is not just a design principle.
 33. Lillicrap TP, Cownden D, Tweed DB, Akerman CJ: **Random synaptic feedback weights support error backpropagation for deep learning.** *Nat Commun* 2016, **7**:1–10.
 34. Marschall O, Cho K, Savin C: **A unified framework of online learning algorithms for training recurrent neural networks.** *J Mach Learn Res* 2020.
 35. Schreier HI, Soen Y, Brenner N: **Exploratory adaptation in large random networks.** *Nat Commun* 2017, **8**.
 36. Rivkind A, Schreier H, Brenner N, Barak O: **Scale free topology as an effective feedback system.** *PLoS Comput Biol* 2020, **16**, e1007825 [Publisher: Public Library of Science].
 37. Shomar A, Barak O, Brenner N: **Cancer progression as a learning process.** *iScience* 2022, **25**, 103924.
 38. Keim NC, Paulsen JD, Zeravcic Z, Sastry S, Nagel SR: **Memory formation in matter.** *Rev Mod Phys* 2019, **91**. 035002.
 39. Stern M, Arinze C, Perez L, Palmer SE, Murugan A: **Supervised learning through physical changes in a mechanical system.** In *Proceedings of the national academy of sciences*, **117**. Publisher: Proceedings of the National Academy of Sciences; 2020: 14843–14850.
 40. Ahmadian Y, Fumarola F, Miller KD: **Context and behavioral processes in extinction.** *Learn Mem* 2004, **11**:485–494.
 41. Gershman SJ, Monfils MH, Norman KA, Niv Y: **The computational nature of memory modification.** *Elife* 2017, **6**, e23763.
 42. Blumenfeld B, Preminger S, Sagi D, Tsodyks M: **Dynamics of memory representations in networks with novelty-facilitated synaptic plasticity.** *Neuron* 2006, **52**:383–394.
 43. Mante V, Sussillo D, Shenoy KV, Newsome WT: **Context-dependent computation by recurrent dynamics in prefrontal cortex.** *Nature* 2013, **503**:78–84.
 44. Li G, Nair S, Quirk G: **A biologically realistic network model of acquisition and extinction of conditioned fear associations in lateral amygdala neurons.** *J Neurophysiol* 2009, **101**: 1629–1646.
 45. McCloskey M, Cohen N: **Catastrophic interference in connectionist networks: the sequential learning problem.** *Physiology of Learning and Motivation* 1989, **24**:109–165.
 46. Tsuda B, Tye KM, Siegelmann HT, Sejnowski TJ: **A modeling framework for adaptive lifelong learning with transfer and savings through gating in the prefrontal cortex.** In *Proceedings of the national academy of sciences*, **117**. Publisher: Proceedings of the National Academy of Sciences; 2020: 29872–29882.
 47. Kumar A, Benjamin AS, Heathcote A, Steyvers M: **Comparing models of learning and relearning in large-scale cognitive training data sets.** *Science of Learning* 2022, **7**:24.
- Long-term monitoring of learning showing that even after very long period with no practice people retain some skills at a high level.

48. Zenke F, Poole B, Ganguli S: **Continual learning through synaptic intelligence**. In *International conference on machine learning*; 2017:3987–3995.
49. Yu A, Dayan P: **Uncertainty, neuromodulation, and attention**. *Neuron* 2005, **46**:681–692.
50. Smith SJ, Hawrylycz M, Rossier J, Sömböl U: **New light on cortical neuropeptides and synaptic network plasticity**. *Curr Opin Neurobiol* 2020, **63**:176–188.
 Review of new developments on the role of neuropeptides in neural communication. The authors show data-derived connectivity matrices of neuromodulation, with specificity stemming from ligand–receptor pairs in various neuron types.
51. Nadim F, Bucher D: **Neuromodulation of neurons and synapses**. *Curr Opin Neurobiol* 2014, **29**:48–56.
52. Mark S, Tsodyks M: **Population spikes in cortical networks during different functional states**. *Front Comput Neurosci* 2012, **6**.
53. Van Damme S, De Fruyt N, Watteyne J, Kenis S, Peymen K, Schoofs L, Beets I: **Neuromodulatory pathways in learning and memory: lessons from invertebrates**. *J Neuroendocrinol* 2021, **33**, e12911.
54. Kutter EF, Boström J, Elger CE, Nieder A, Mormann F: **Neuronal codes for arithmetic rule processing in the human brain**. *Curr Biol* 2022, **32**:1275–1284.e4.