Frozen algorithms: how the brain’s wiring facilitates learning
Dhruva V Raman and Timothy O’Leary

Synapses and neural connectivity are plastic and shaped by experience. But to what extent does connectivity itself influence the ability of a neural circuit to learn? Insights from optimization theory and AI shed light on how learning can be implemented in neural circuits. Though abstract in their nature, learning algorithms provide a principled set of hypotheses on the necessary ingredients for learning in neural circuits. These include the kinds of signals and circuit motifs that enable learning from experience, as well as an appreciation of the constraints that make learning challenging in a biological setting. Remarkably, some simple connectivity patterns can boost the efficiency of relatively crude learning rules, showing how the brain can use anatomy to compensate for the biological constraints of known synaptic plasticity mechanisms. Modern connectomics provides rich data for exploring this principle, and may reveal how brain connectivity is constrained by the requirement to learn efficiently.

Address
Department of Engineering, University of Cambridge, United Kingdom

Corresponding author: O’Leary, Timothy (tso24@cam.ac.uk)

Introduction
The brain learns by modifying itself in response to information from the external world. Since the earliest experimental demonstration of how learning can be implemented by synaptic modification [1,2], experimental attention has focused on mechanisms that link experience-dependent signals to modifications in synaptic strength. Over many decades of research we have witnessed an explosion of different synaptic learning rules, from timing-dependent mechanisms, to neuromodulator induced plasticity as well as combinations of both [3,4].

Paradoxically we are still unable to connect these many kinds of synaptic modification mechanisms to high-level biological learning at the circuit or behavioral level. Basic theory from optimization sheds light on the problem of connecting biological learning to low-level plasticity rules [5**,6,7,8**,9*,10,11*]. A key insight is the so-called credit assignment problem. In a large network with many interconnected cells there is a complex relationship between a change in any given connection and a change in overall behavior. This difficulty means that each synapse may require detailed information in order to modify in the correct way during learning [12**].

Recent progress in AI shows that conceptually simple optimization processes can allow artificial neural networks to achieve superhuman performance on cognitive tasks. The high-level design of such AI systems has informed systems-level hypotheses on biological learning (e.g. [13]). However, the training step requires exchange across the network of vast amounts of information, depending upon the state of each synapse.

To a greater extent than artificial neural networks, biological neural circuits face limiting constraints on the speed, accuracy, and quality of information that can be exchanged between synapses and from the external world. This limits the set of plausible learning mechanisms, and, we argue here, imposes heavy constraints on circuit architecture that are beginning to be revealed in recent connectomics work [14**,15**, 16].

We will outline key theoretical aspects of learning from an optimization perspective. This will give a concrete framework for understanding why learning is hard computationally, and how this problem can be partially solved with specific circuit connectivity. This topic does not address all aspects of learning, nor does it address all ways in which circuit architecture aids or hinders learning. We will cite recent experimental connectomics work that provides striking examples of circuit motifs that make sense in light of concrete theory. We will also review recent theoretical work that shows how circuit architecture can aid learning, providing hypotheses for future experimental studies.

Theoretical background: why is learning hard in large neural circuits?
Consider a neural circuit with $N$ synapses. Abstractly, learning minimises some measure of behavioural error (a ‘loss function’) that depends on all $N$ synapses. Any change in the synapses can be represented as an $N$-dimensional vector, endowing collective synaptic
changes with a direction and a magnitude. The credit assignment problem boils down to three general factors that determine the difficulty of learning:

Factor 1 The sensitivity of behavioural performances to ‘good’ synaptic changes.
Factor 2 The robustness of behavioural performance to ‘bad’ synaptic changes.
Factor 3 The relative proportions of ‘good’ and ‘bad’ changes induced by a synaptic learning rule.

What do we mean by ‘good’ and ‘bad’ changes? At any point in time during learning, there is a direction of maximal sensitivity of the loss function to changes in the synapses. This corresponds to the gradient of the loss function, which defines the ‘good’ direction of change.

Artificial learning algorithms compute gradients of loss functions directly to solve the credit assignment problem. In optimization theory, this is known as a ‘first-order’ method [17], which explicitly computes the sensitivity of a circuit-level property to a change in synaptic weights, before adjusting all weights collectively to steer the network to a desired state. However, computing the full gradient generally requires knowledge of all synaptic weight values. While possible on a computer, there is no feasible means of exchanging this amount of weight information explicitly in the brain [12**].

Biologically, this means that synaptic changes may only correlate with the gradient, with some proportion of synaptic changes that are not related to the gradient. These correspond to ‘bad’ changes.

An example of a class of biologically plausible learning rules that suffer from a problematic proportion of bad changes is as follows. Apply a random perturbation to all \( N \) synapses, then retain the perturbation or reverse it depending on the observed effect on behavioural performance. It is easy to see how a biological synapse might implement such a rule: all it requires is access to the recent error history, for example by integrating a global neuromodulatory signal over time. Learning rules of this form are known as ‘zero-order’, as they do not directly approximate the gradient. Examples include the MIT rule [18], the REINFORCE algorithm [19], and node perturbation [19,20]. Biologically plausible implementations have been proposed as the underlying mechanisms for birdsong learning [21] and cerebellar learning [22**].

Unfortunately, although plausible biologically, such learning rules do not work well if they operate on large populations of synapses. A calculation outlined in Box 1 shows that the correlation of such learning rules with the gradient drops in proportion to \( 1/\sqrt{N} \) where \( N \) is the number of synapses. Put another way, biologically plausible zero-order rules only provide feedback on a single direction in weight space, while the full gradient provides information on all directions. Intuitively, we can see that as \( N \) increases, the chance that a random direction in weight space aligns with the gradient drops. For zero-order rules this means that the proportion of bad

---

**Box 1 Mathematical analysis of learning**

Denote behavioural performance in a task by a loss function \( L[\mathbf{w}] \), where \( \mathbf{w} \) is the \( N \)-dimensional vector of synaptic weights in a network. Learning involves descending this loss function. ‘Gradient-based’ learning rules explicitly form an approximation of the negative gradient \( -\nabla L[\mathbf{w}(t)] \) and adjust synaptic weights in this direction during learning. However, *any* learning rule, even ‘gradient-free’ learning rules, must adjust synapses in direction that anticorrelates with \( \nabla L[\mathbf{w}] \) on average. Consider a small time interval \( \Delta t \), in which the weights change by an amount \( \Delta \mathbf{w} \). We can write the change in behavioural error over time \( t + \Delta t \) as

\[
L[\mathbf{w} + \Delta \mathbf{w}] - L[\mathbf{w}] \approx \nabla L[\mathbf{w}]^T \Delta \mathbf{w} + \frac{1}{2} \Delta \mathbf{w} \nabla^2 L[\mathbf{w}] \Delta \mathbf{w},
\]

where \( \nabla L[\mathbf{w}] \) is the gradient of the loss function, and \( \nabla^2 L[\mathbf{w}] \) is the Hessian (second derivative). If learning occurs in \( \Delta t \), the change in error in 1 will be negative. First, consider the average effect of a completely random weight change, \( \Delta \mathbf{w} \), on loss. In this case,

\[
\mathbb{E}[L[\mathbf{w} + \Delta \mathbf{w}] - L[\mathbf{w}] = 0 + \frac{1}{2} \| \Delta \mathbf{w} \|_2^2 \text{Tr}(\nabla^2 L[\mathbf{w}]) / N],
\]

where \( \text{Tr}(\cdot) \) is the trace of a matrix (sum of eigenvalues). We can see that

- A random perturbation is uncorrelated, in expectation, with the gradient. So the first term on the RHS is zero.
- Therefore \( \text{Tr}(\nabla^2 L[\mathbf{w}]) \) quantifies the robustness of the learning system to random perturbations (Factor 2 in the main text).
- In a partially trained network, random perturbations are expected to negatively affect behavioural performance, that is, \( \text{Tr}(\nabla^2 L[\mathbf{w}]) > 0 \).

Now consider a perturbation-based (0 order) learning rule. Let \( \Delta \mathbf{w} \) be a random Gaussian perturbation that is kept or reversed so that it anticorrelates with the gradient. In this case, \( \Delta \mathbf{w} \nabla L[\mathbf{w}] \) has a half-normal distribution. In expectation, the first term of the RHS of Equation 1 reduces to

\[
\mathbb{E}[\Delta \mathbf{w} \nabla L[\mathbf{w}] = \sqrt{\frac{2}{N \pi}} \| \Delta \mathbf{w} \|_2 \| \nabla L[\mathbf{w}] \|_2. \tag{2a}
\]

Thus the correlation with the direction of steepest descent decreases at a rate proportional to \( 1/\sqrt{N} \). At the other extreme, exact (batch) backpropagation uses the update

\[
\Delta \mathbf{w} = -c \nabla L[\mathbf{w}],
\]

for some constant of proportionality \( c \). In this case, the correlation with the gradient is

\[
\mathbb{E}[\Delta \mathbf{w} \nabla L[\mathbf{w}] = -c \| \nabla L[\mathbf{w}] \|_2^2. \tag{2b}
\]

In either case, \( \| \nabla L[\mathbf{w}] \|_2 \) defines the sensitivity of behavioural performance to plasticity in ‘good’ directions, and is a mathematical analogue to Factor 1 in the main text.

Intermediate to Equations 2a and 2b, a noisy first-order rule might align only a small amount with the gradient, like 2a, but have a constant degree of alignment that is independent of \( N \), like 2b. For these learning rules (or for anywhere \( \mathbb{E}[\Delta \mathbf{w} \nabla L[\mathbf{w}] \) degrades at a slower rate than \( 1/\sqrt{N} \) adding apparently redundant neurons can increase the speed and eventual degree of learning \( 5^{**} \).
synaptic changes is likely to grow with network size. There are two ways to mitigate this problem:

1. Increase the dimensionality of feedback signals. A better correlation between weight changes and the full gradient can be achieved if the feedback signal itself has multiple components. This can be achieved in the brain via multiple, independent feedback pathways that carry vectorised error information to neural circuits. Biologically, error signals can be encoded in neural activity and, in particular, in neuromodulatory signals that convey behavioural valence directly to synapses where learning occurs. We will review recent connectomic studies, particularly in invertebrates, that have uncovered the existence of such precise modulatory feedback in learning centers in the brain.

2. Use circuit architectures that are robust to ‘bad’ perturbations (i.e. improve Factor 2), while maintaining sensitivity to the ‘good’ component (i.e. maintain Factor 1). It turns out, perhaps surprisingly, that redundant connectivity in a neural circuit can reduce behavioural sensitivity to bad synaptic changes while maintaining sensitivity to good changes. This means that even crude but biologically plausible synaptic learning rules may achieve efficient learning with the appropriate circuit architecture.

In the following sections we review recent work that suggests these principles are at work in the brain.

**Connectomic evidence for vectorised learning signals in the brain**

Imagine repetitively practicing and attempting to improve a tennis serve. A coach evaluates each serve with a mark out of 10. This is an example of scalar feedback, since the entire behaviour is reduced to a single number. With scalar feedback and multiple behavioural parameters, it is hard to determine which behavioural aspects should be altered, and how.

Now suppose you can correlate aspects of the behaviour (the serve) to the coach’s mark, and gradually build up a picture of which behavioural aspects contribute to a good mark (hit the ball harder, keep head raised higher, etc.). Such multifaceted error information is an example of vector feedback.

**Figure 1a** depicts the architecture of feedback signals in a generic neural circuit, from a coarse (scalar) feedback error signal to a fine, vector signal that decomposes error across different behavioural readouts and/or subsets of circuit connections. Higher dimensional feedback eases the difficulty of learning through factor 3. This is because the more connections that are affected by a given component of an error signal, the worse the ratio of ‘good’ changes to ‘bad’ changes becomes (see Equation 2).

Regardless of the specific learning mechanism, more targeted information on behavioural contributions to a task cannot be a bad thing. This benefit can be enhanced further by the partitioning of subcircuits according to the behaviour they influence. For example, suppose separate neural subcircuits control head direction and racket force during the tennis serve in the example above. If these behavioural variables are uncorrelated and targeted feedback on how to change each variable reaches separate neural circuits, then the credit assignment problem need only be solved in each separate subcircuit. On the other hand if head direction and racket force influence each other than this benefit disappears, as both subcircuits collectively influence both behavioural variables.

There is a large body of literature detailing how abstract properties of behaviour, such as reward prediction error, are represented in the brain and inform learning [3,5,23,24]. Minimally, such error information can correspond to a scalar error signal from bulk release of neuromodulators such as dopamine.

However, more recently, detailed connectomics in *Drosophila* suggests that evolution has exploited the advantages of vector neuromodulatory feedback for learning. Feedback pathways necessary for learning target discrete neural subcircuits, each responsible for different aspects of behaviour. Moreover, these separate pathways seem to provide weakly correlated signals.

The *Drosophila* mushroom body (MB) mediates associative learning across sensory modalities, especially olfaction. This circuit has an architecture that enables weighted combinations of stimuli, such as a combination of odour signals, to become associated with behavioural actions in an experience-dependent manner [25,26]. Distinct modulatory subpopulations represent different error signals. For example, the optogenetic activation of specific modulatory subpopulations, in concert with an odour stimulus, is sufficient to form an aversive or appetitive association [27].

Even for a particular valence (e.g. appetitive), distinct modulatory subpopulations encode a more finely tuned, multidimensional reward signal. As an example, [28*] showed that different subsets of modulatory neurons were required for the associative learning of an odour with two different kinds of reward (fructose versus aspartic acid), even while both triggered the same behaviour of learned odour attraction.

Is the functional specificity of these modulatory neurons encoded anatomically? Evidence points to different modulatory neurons receiving different sets of sensory afferents, and therefore potentially computing somewhat independent error signals. The mushroom body has a segregated architecture consisting of largely independent...
neural subpopulations learning associations from different neuromodulatory feedback signals. In [29], the MB was conceptually divided into 15 compartments, defined by the presence of specific dopaminergic (modulatory) and output neuron cell types. Tracing studies including [14**,15**] found that each modulatory neuron received a unique combination of sensory afferents, with same-type modulatory neurons receiving more similar sensory input. Moreover, many modulatory neurons receive over half of their dendritic input from recurrent connections [14**]. This recurrence links different MB compartments, such as those responsible for aversive and appetitive associations (see Figure 2b), thus allowing for reward signals incorporating both sensory input and integrated past experienced from multiple associative systems. The recurrent input is highly diverse [15**] across different modulatory neuron types, emanating from different regions including gustatory interneurons and the lateral horn.

Together, this work supports a hypothesis that different neuromodulatory pathways are affected by different aspects of past experience, as well as unique combinations of sensory stimuli. This may provide non-redundant, vector feedback signals to independent neural subpopulations in the MB. This circuit architecture is expensive in terms of developmental complexity and wiring cost, yet the theory discussed in the previous section tells us that this price may be necessary for efficient learning. It would be interesting, on the other end, to see whether different neuromodulatory pathways are associated with the formation of distinct behavioural responses.

We have reviewed the anatomical features in the *Drosophila* mushroom body that could ease the burden of the credit assignment problem. This was made possible by the detailed connectomic information available on the *Drosophila*. The existence of zero-order learning mechanisms has been proposed in higher organisms as well, notably the song-learning system in Zebra finches [16,21,30]. The songbird provides an excellent model system for a difficult learning problem because learning to mimic and recognise birdsong corresponds to learning...
complex temporal sequences of muscle movements and auditory stimuli: a rich feedback signal. A proposed implementation of local credit assignment, using the node-perturbation algorithm [21] lead to specific anatomical predictions subsequently verified by connectomics studies [16].

Redundant connectivity can boost learning efficiency of noisy, local synaptic plasticity mechanisms

Clearly, the difficulty of a specific learning problem is influenced by the architecture of the neural circuit used to solve it. We now review some recent theoretical results that provide mechanistic reasons as to why particular, observed architectures are favourable for the learning problems they solve.

A generic architecture that recurs across organisms is the ‘input expansion’. Here a large, horizontal ‘mixed-layer’ of cells receives low numbers of inputs, and projects to low numbers of outputs. For example, this architecture occurs in Kenyon cells in the insect mushroom body, or cerebellar granule cells. One advantage of such an architecture is that it can ‘separate out’ correlated inputs by embedding them in a higher dimensional space, as proposed decades ago in [31]. Recent modelling work has justified more detailed aspects of these architectures [6,7,8**]. In [8**], a simplified circuit model of the input expansion motif was constructed. The sparsity (i.e. number of inputs projecting onto each mixed-layer neuron), and degree of feedforward inhibition onto the mixed layer were set as tunable parameters, as was the number of neurons in the two layers. They investigated how these tunable parameters related to dimensionality of the representation. This latter quantity is important as the higher it is, the greater the efficacy of associative learning using a Hebbian readout neuron. Optimal sparsity of the input connections to the mixed layer was calculated, and found to correspond exactly to the degree of anatomically observed sparsity in the mossy fibre inputs to cerebellar granule cells to the human brain, as well as antennal lobe inputs to Kenyon cells in Drosophila.

We have discussed how a particular anatomical circuit motif supports pattern separation, an unsupervised learning problem. The specific anatomical predictions of [8**] did not apply in the case where synapses were subject to supervised plasticity, although it was noted that dense connectivity seemed more supportive of fast learning in this case. More recently, [5**] showed quantitatively how and why adding neurons in densely connected circuits could speed learning, even when smaller circuits could theoretically solve the learning problem to perfect accuracy. Thus, adding ‘redundant’ neurons can compensate...
for an inaccurate learning rule that cannot induce plasticity well-aligned with the gradient. The key insight was that adding redundant neurons could increase sensitivity of behavioural performance to ‘good’ directions of plasticity, without decreasing the robustness of behavioural performance to ‘bad’ directions. Mathematically, the sensitivity of behavioural performance corresponded to the magnitude of the gradient, and the robustness depended upon the eigenvalues of the Hessian (see Box 1 for insight and some details). Note that this benefit of redundancy is counteracted, for zero-order learning rules, by the increased burden of the credit assignment problem in large networks (i.e. inability to accurately find the gradient).

The result of [5**] suggests that learning rules could be quantified not only by how well they perform in *in silico* circuits, but on how their performance scales as network size increases to biologically realistic levels. A given learning rule may provide a highly inaccurate approximation of the gradient. Nevertheless, if the quality of this approximation does not decrease with increasing network size (or decreases more slowly than zero-order, perturbation-based learning rules), then the described benefits of redundantly large circuits could compensate to nevertheless enable fast, accurate learning. However, these benefits might only emerge at extremely large network sizes.

The previous point is relevant given the plethora of proposed, biologically plausible approximations to back-propagation that have emerged in the literature [12**]. Backpropagation itself requires individual synapses to know the states of their downstream counterparts. This is known as the ‘weight transport’ problem. Various biologically plausible surrogates to weight transport partially circumvent this. One solution is to demand parallel feedback connections that instead provide synapses with information on downstream neural activities, which serves as a rough proxy [32–41]. It is worth noting that these feedback connections effectively provide an error signal with vectorised information on the errors of different output neurons. Another solution is to let individual neurons take the role of an abstract ‘network’, with internal ‘layers’ [9*]. This makes weight transport an intracellular problem, and allows for the separation of feedback pathways respectively providing reward and weight transport signals [10,11*,42].

All of these proposed algorithms explicitly, albeit inaccurately, compute the sensitivities of behavioural performance to individual synapses. This inaccuracy has led to mixed success when benchmarking their performance against artificial learning rules not subject to biological plausibility constraints [43**]. However, the results of [5**] suggest that learning performance could improve as network size increases to biological levels.

**Conclusion**

It is difficult, if not impossible, to understand how synaptic-level learning rules enable behavioural level learning without studying the structure of neural circuits and the abstract mathematical nature of learning as an optimisation process. Even with intricate synaptic plasticity mechanisms, there is a fundamental algorithmic difficulty in implementing learning at the whole circuit or behavioural level due to the credit assignment problem. We have reviewed two very plausible classes of solutions to this problem: vector feedback and redundant connectivity.

Recent advances in quantitative neuroanatomy [25,44–48] have made it possible to measure and quantify circuit architecture, revealing striking evidence that these solutions are used biologically. We need principles to make sense of the mass of complex data that connectomics reveals. The question of how learning might work algorithmically was successfully addressed long before the era of modern connectomics in Marr and Albus’s groundbreaking analysis of cerebellar architecture [31,49]. New theoretical insights and insights from artificial learning, coupled with extraordinarily detailed and comprehensive connectomics data, places us in an exciting era that will leverage experiment and theory to understand biological learning.

**Conflict of interest statement**

Nothing declared.

**Acknowledgements**

This work is supported by ERC grant 716643 FLEXNEURO.

**References and recommended reading**

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

- of special interest
- of outstanding interest

1. Pinsker H, Kupfermann I, Castellucci V, Kandel ER: Habituation and disinhibition of the GM-Withdrawal reflex in Aplysia. Science 1970, 167:1740-1742.
2. Castellucci V, Kandel ER: Presynaptic facilitation as a mechanism for behavioral sensitization in Aplysia. Science 1976, 194:1176-1178.
3. He K, Huertas M, Hong SZ, Tie XX, Hell JW, Shouval Ha. Kirkwood A: Distinct eligibility traces for LTP and LTD in cortical synapses. Neuron 2015, 88:528-538.
4. Brzosko Z, Zannone S, Schultz W, Clopath C, Paulsen O: Sequential neuromodulation of Hebbian plasticity offers mechanism for effective reward-based navigation. Learning 2017, 39:40.
5. Raman DV, Rotondo AP, O’Leary TS: Fundamental bounds on learning performance in neural circuits. Proc Natl Acad Sci U S A 2019.

This theoretical study shows that neural circuits can compensate for highly inaccurate learning rules by adding ‘redundant’ neurons and connections that are not required to perform the assigned task. This implies that excess circuit connectivity can be traded for learning performance. However, if the connections are inherently noisy there is an
upper bound to the number of excess connections, above which learning performance degrades.

6. Babadi B, Sompolinsky H: Sparserness and expansion in sensory representations. Neuron 2014, 83:1213-1226.

7. Cayco-Gajic AN, Clopath C, Silver AR: Sparse synaptic connectivity is required for decorrelation and pattern separation in feedforward networks. Nat Commun 2017, 8:1116.

8. Kumar AL, Harris KD, Axel R, Sompolinsky H, Abbott LF; ** Optimal degrees of synaptic connectivity. Neuron 2017, 93:1153-1164

This theoretical study shows that differing neural connectivity patterns across circuits can be explained by the common goal of producing high dimensional stimulus representations for effective associative learning.

9. Richards BA, Lillicrap TP: Dendritic solutions to the credit • assignment problem. Curr Opin Neurobiol 2019, 54:28-36

This review points out that feedback signals responsible for credit assignment can be extracted from other feedback signals and reviews how dendritic processing can achieve this.

10. Guergueïev J, Lillicrap TP, Richards BA: Towards deep learning with segregated dendrites. eLife 2017, 6:e22901.

11. Sacramento J, Costa RP, Bengio Y, Senn W: Dendritic cortical microcircuits approximate the backpropagation algorithm. Advances in Neural Information Processing Systems 31 2018:8721-8732

This study presents a model of a biologically plausible, approximate implementation of backpropagation. Both learning and task performance operate continuously in time, by exploiting spatially segregated dendritic computations.

12. Lillicrap TP, Santoro A, Marris L, Akerman CJ, Hinton GE; ** Backpropagation and the brain. Nat Rev Neurosci 2020:1-12

This review highlights how key bottlenecks in biologically plausible implementation of backpropagation could be circumvented. This would result in high-performance, biologically plausible learning rules.

13. Wang JX, Kurch-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, Hassabis D, Botvinick M: Prefrontal cortex as a meta-reinforcement learning system. Nat Neurosci 2018, 21:860-868.

14. Eschbach C, Fushiki A, Windling M, Schneider-Mizell CM, Shao M, ** Arruda R, Eichler K, Valdes-Alemán J, Ohyama T: Recurrent architecture for adaptive regulation of learning in the insect brain. Nat Neurosci 2020, 23:544-555

An extensive experimental connectomics study that uncovers circuit motifs supporting associative learning in Drosophila. A range of distinct, distributed teaching signals are regulated by the output of disparate brain regions.

15. Otto N, Pleijjier MW, Morgan IC, Edmondsson-Stait AJ, Heinz KJ, ** Stark I, Dempsey G, Ito M, Kapoor I, Hsu J, Schlegel PM, Bates AS, Fang L, Costa M, Ito K, Bock DD, Rubin GM, Jeffries GSX, Waddell S: Input connectivity reveals additional heterogeneity of dopaminergic reinforcement in Drosophila. Curr Biol 2020, 30:3200–3211.68

This study reveals that separate types of dopaminergic reward neurons respond to different, specific, stimuli and project onto different Drosophila mushroom body compartments. This provides explicit evidence of vector learning signals in this brain region.

16. Komfeld J, Januszewski M, Schubert P, Jain V, Denk W, Fee MS: An anatomical substrate of credit assignment in reinforcement learning. bioRxiv 2020, 2020.02.18.954354.

17. Polya BT: Introduction to Optimization. New York: Optimization software. Inc., Publications Division; 1987, 1.

18. Whitaker HP: An Adaptive System for Control of the Dynamics Performance of Aircraft and Spacecraft. Inst. Aeronautical Sciences; 1959:59–100.

19. Williams RJ: Simple statistical gradient-following algorithms for connectionist reinforcement learning. Mach Learn 1992, 8:229-256.

20. Flower B, Jabri M: Summed weight neuron perturbation: an O(n) improvement over weight perturbation. Neural Information Processing Systems 1993:212-219.

21. Fiete IR, Fee MS, Seung HS: Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. J Neurophysiol 2007, 98:2038-2057.

22. Bouvier G, Aljadeff J, Clopath C, Bimbad C, Ranft J, Blot A; ** Nadal J-P, Brunel N, Hakim V, Barbour B: Cerebellar learning using perturbations. eLife 2018, 7:e31599

The authors propose an explicit mechanism for implementing a classic perturbation-based learning rule in the cerebellar architecture. Verified surprising predictions on synaptic plasticity rules with electrophysiological recordings.

23. Schultz W: Neuronal reward and decision signals: from theories to data. Physiol Rev 2015, 95:853-951.

24. Yagishita S, Hayashi-Takagi A, Ellis-Davies GCR, Urakubo H, Ishii S, Kasai H: A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science 2014, 345:1616-1620.

25. Thum AS, Gerber B: Connectomics and function of a memory network: the mushroom body of larval Drosophila. Curr Opin Neurobiol 2019, 54:146-154.

26. Eichler K, Li F, Litwin-Kumar A, Park Y, Andrade I, Schneider-Mizell CM, Baumweber T, Husar A, Eschbach C, Gerber B et al.: The complete connections of a learning and memory centre in an insect brain. Nature 2017, 548:175.

27. Schroll C, Riemsponger T, Bucher D, Emmer J, Völler T, Erburg K, Gerber B, Hendel T, Nagel G, Buchner E, Fiala A: Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in Drosophila larvae. Curr Biol 2006, 16:1741-1747.

28. Baumweber T, Rohweder A, Schleyer M, Eichler K, Chen Y-c, ** Aso Y, Cardona A, Eschbach C, Kobler O, Voigt A, Duraraja A, Mancini N, Zlatic M, Truman JW, Thum AS, Gerber B: Functional architecture of reward learning in mushroom body extrinsic neurons of larval Drosophila. Nat Commun 2018, 9:1104

This study characterises the identity, morphology, and functionally analyses all 44 input/output neurons of the Drosophila mushroom body. The data reveal that teaching signals for different associative learning tasks (e.g. odour preference, taste preference) are specialised to different modulatory neurons.

29. Aso Y, Hattori D, Yu Y, Johnston RM, Iyer NA, Ngo T-TB, Dione H, Abbott LF, Axel R, Tanimoto H, Rubin GM: The neuronal architecture of the mushroom body provides a logic for associative learning. eLife 2014, 3:e04577.

30. Fee MS, Goldberg JH: A hypothesis for basal ganglia-dependent reinforcement learning in the songbird. Neuroscience 2011, 198:152-170.

31. Albus JS: A theory of cerebellar function. Math Biosci 1971, 10:25-61.

32. LeCun Y, Bengio Y, Hinton GE: A theoretical framework for back-propagation.. In Proceedings of the 1988 Connectionist Models Summer School, vol 1. CMU, Pittsburgh, PA: Morgan Kaufmann; 1988, 21-28.

33. Bengio Y: How Auto-Encoders Could Provide Credit Assignment in Deep Networks Via Target Propagation. 2014arXiv:1407.7906.

34. Lillicrap TP, Cownden D, Tweed DB, Akerman CJ: Random synaptic feedback weights support error backpropagation for deep learning. Nat Commun 2016, 7.

35. Lee D-H, Zhang S, Fischer A, Bengio Y: Difference target propagation. Joint European Conference on Machine Learning and Knowledge Discovery in Databases 2015:498-515.

36. Liao Q, Leibo JZ, Poggio T: How Important is Weight Symmetry in Backpropagation? 2016arXiv:1510.05067 [cs].

37. Whittington JCR, Bogacz R: An approximation of the error backpropagation algorithm in a predictive coding network with local Hebbian synaptic plasticity. Neural Comput 2017, 29:1229-1262.

38. Meulemans A, Carzaniga FS, Suykens JAK, Sacramento J, Grewe BF: A Theoretical Framework for Target Propagation. 2020arXiv:2006.14331 [cs, stat].
39. Nøkland A: Direct feedback alignment provides learning in deep neural networks. In Advances in Neural Information Processing Systems, volume 29. Edited by Lee D, Sugiyama M, Luxburg U, Guyon I, Garnett R. Curran Associates, Inc.; 2016:1037-1045.

40. Murray JM: Local online learning in recurrent networks with random feedback. eLife 2019, 8:e43299.

41. Bellec G, Scherr F, Subramoney A, Hajek E, Salaj D, Legenstein R, Maass W: A solution to the learning dilemma for recurrent networks of spiking neurons. Nat Commun 2020, 11:3625.

42. Payeur A, Guerguiev J, Zenke F, Richards BA, Naud R: Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. bioRxiv 2020. 2020.03.30.015511.

43. Bartunov S, Santoro A, Richards BA, Marris L, Hinton GE, Lillicrap TP: Assessing the scalability of biologically-motivated deep learning algorithms and architectures. Advances in Neural Information Processing Systems 31 2018:9368-9378

This paper benchmarks performance of a suite of ‘biologically plausible’ learning rules from the literature on standard machine learning datasets (MNIST, CIFAR, Imagenet). Finds limited success on learning the latter two datasets.

44. Kevin MB, Boergens M, Helmstaedter M: SegEM: efficient image analysis for high-resolution connectomics. Neuron 2015, 87:1193-1206.

45. Kornfeld J, Denk W: Progress and remaining challenges in high-throughput volume electron microscopy. Curr Opin Neurobiol 2018, 50:261-267.

46. Motta A, Berning M, Boergens KM, Staffler B, Beining M, Loomba S, Hennig P, Wissler H, Helmstaedter M: Dense connectomic reconstruction in layer 4 of the somatosensory cortex. Science 2019, 366.

47. Schneider-Mizell CM, Gerhard S, Longair M, Kazimiers T, Li F, Zwart MF, Champion A, Midgley FM, Fetter RD, Saalfeld S et al.: Quantitative neuroanatomy for connectomics in Drosophila. Elife 2016, 5:e12059.

48. Zheng Z, Lauritzen JS, Perlman E, Robinson CG, Nichols M, Milkie D, Torrens O, Price J, Fisher CB, Sharifi N et al.: A complete electron microscopy volume of the brain of adult Drosophila melanogaster. Cell 2018, 174:730-743.

49. Marr D: A theory of cerebellar cortex. J Physiol 1969, 202:437-470.