



Learning with three factors: modulating Hebbian plasticity with errors

Łukasz Kuśmierz, Takuya Isomura and Taro Toyozumi

Synaptic plasticity is a central theme in neuroscience. A framework of three-factor learning rules provides a powerful abstraction, helping to navigate through the abundance of models of synaptic plasticity. It is well-known that the dopamine modulation of learning is related to reward, but theoretical models predict other functional roles of the modulatory third factor; it may encode errors for supervised learning, summary statistics of the population activity for unsupervised learning or attentional feedback. Specialized structures may be needed in order to generate and propagate third factors in the neural network.

Address

RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

Corresponding author: Toyozumi, Taro (taro.toyozumi@brain.riken.jp)

Current Opinion in Neurobiology 2017, **46**:170–177

This review comes from a themed issue on **Computational neuroscience**

Edited by **Adrienne Fairhall** and **Christian Machens**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 15th September 2017

<http://dx.doi.org/10.1016/j.conb.2017.08.020>

0959-4388/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Associative (Hebbian) learning indicates association between two factors (two sensory inputs or an input and an output), but such a learning is often influenced by a so-called third factor. In a very general framework of three-factor learning, plasticity is realized by changing a synaptic strength w with the following rule

$$\dot{w} = F(pre, post, g, w), \quad (1)$$

where pre and $post$ are some functions of histories of presynaptic and postsynaptic activities, g is a third factor modulating the plasticity (Figure 1), and \dot{w} denotes the time derivative of the synaptic strength w . The third factor may represent, for example, rewards, supervised errors, summary statistics, or attentional feedback, which could be used to facilitate different types of learning by providing more global information about how well the whole network is performing or how important a current

situation is. Often learning rules are written in a more specific form

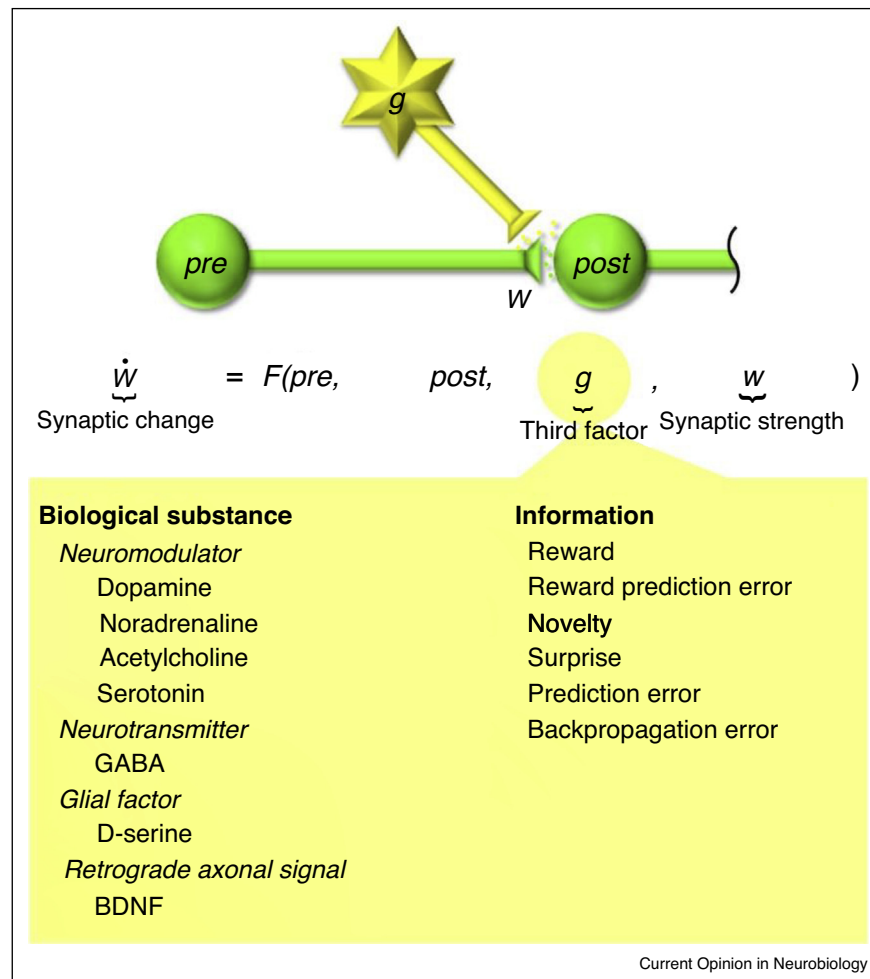
$$\dot{w} = gH(pre, post), \quad (2)$$

where H is a generalized Hebbian term, which includes some measure of correlation between presynaptic and postsynaptic activities. As a simple example, the classical Hebbian learning assumes a rate model of neurons, wherein the activities are described by real valued firing rates f_{pre} and f_{post} . The Hebbian plasticity term then simply involves a product of those firing rates ('fire together wire together').

In more detailed, biologically plausible models, the activity of each neuron is approximated by a point process, that is, it is fully determined by a set of times at which the neuron generated action potentials (spike train). When both pre and $post$ are given by the spike trains of the corresponding neurons, the learning rule based on $H(pre, post)$ is called spike-timing-dependent plasticity (STDP) [1,2]. In the simplest scenario STDP is described by pairwise interactions, that is, it depends only on the relative timing of pairs ($pre-post$) of individual spikes [3,4]. Function $H(pre, post)$ can be in this case determined by the learning window (also STDP function), that is, one dimensional function of the relative time between pre-synaptic and postsynaptic spikes. In the standard STDP, long-term potentiation (LTP, the connection is strengthened) is observed if the presynaptic spike precedes (in some short time window) the postsynaptic spike (pre-before-post), whereas long-term depression (LTD, the connection is weakened) is observed if the postsynaptic spike precedes the presynaptic spike (post-before-pre). This temporally asymmetric STDP is an extension of the original Hebb's postulate and in some limits simplifies to the classical Hebbian term ($f_{pre}f_{post}$). Note that Equation 2 can describe more complicated STDP rules that may involve more than two spikes [5] or more biophysical calcium-based plasticity rules if the calcium concentration is primarily determined by the presynaptic and postsynaptic activity [6,7].

A possible biological implementation of the three-factor learning is provided by neuromodulators. Multiple *in vitro* experimental studies have shown that neuromodulators modulate Hebbian plasticity in various ways. In hippocampus, the activation of the D_1 subunit dopamine receptor reverses LTD to LTP and extends the LTP

Figure 1



A schematic image of modulations of Hebbian plasticity by third factors.

part of the STDP time window [8], leading to temporally symmetric STDP function (LTP for both pre-before-post and post-before-pre). In contrast, the activation of the α subunit adrenaline receptor reverses LTP to LTD [9]. In addition, modulations of synaptic plasticity occur with various neuromodulators, including dopamine [10], noradrenaline [11], acetylcholine [12], and serotonin (5-HT) during the developmental stage [13]. Another biological mechanism that can implement the three-factor learning is inhibition. Recently, it was reported that GABAergic inhibition directly suppresses local dendritic Ca^{2+} signaling and promotes spine shrinkage and elimination of hippocampal dendritic spines [14], and such suppression of dendritic Ca^{2+} is sensitive to precise timing (<5 ms) of inhibitory input [15]. In corticostriatal synapses, with intact physiological GABAergic transmission, the pre-before-post stimulation induces LTD, while the post-before-pre stimulation induces LTP

[16]. However, blockade of GABA_A -receptors converts LTD into LTP, and vice versa [17]. In addition, glial cells may also modulate and coordinate Hebbian plasticity [18].

In this manner, the third factor modulates the original associative learning in various ways, which must play roles in different brain functions. Note that the multiplicative relationship between the third factor and the Hebbian term in Equation 2 is a useful mathematical simplification. The biological third factors described above can, in addition, directly modulate the presynaptic or postsynaptic activities. In the rest of the paper we list some of the hypothetical roles of the third factor proposed in the theoretical literature, as well as possible computational mechanisms of their generation and propagation. Although many of these functions were proposed based on theoretical considerations, the underlying algorithms

are biologically plausible, that is, they could potentially be implemented by the brain.

Functional roles of the three-factor learning rules

In contrast to the two factor Hebbian learning that modifies synapses based on their presynaptic and postsynaptic activity, the three-factor learning is more flexible because synapses are modulated also based on the third factor that can reflect more global information about how well the whole network is performing or how important a current situation is. A three-factor learning rule naturally arises from training synaptic strength to optimize a network function. Examples of such optimization include maximization of predicted reward or minimization of error from preferable activity patterns as we describe below.

The most well-known example of the three-factor learning is the connectionist implementation of the reinforcement learning (RL) [19,20]. Learning in RL is driven by scalar rewards r , received by an agent acting in its environment. Although the reward alone could be in principle used as the third-factor [19,20], effective algorithms for learning the optimal policy are based on a reward prediction error δ (temporal difference [21]). Indeed, it has been observed that δ is encoded by dopaminergic neurons [22,23[•]]. Most of the connectionist RL algorithms can be written in the form of Equation 1, where the reward prediction error δ is used as the third factor g . This type of the three-factor learning is well-established and an interested reader is referred to [24[•],25,26] for comprehensive reviews on the topic.

In supervised learning, in contrast to the reward signal, supervised signals provide full information about the desired output of the neurons. Those third factor signals come in different flavors, depending on the coding scheme used by the network. In rate neural models, the desired and actual outputs are given by continuous firing rates, whereas in spiking neural models the corresponding fully supervised signal should encode the desired spike train. Several schemes for learning spike trains have been proposed recently, see ReSuMe [27], chronotron [28], SPAN [29], PSD [30], MPDP [31], and FILT [32].

One can also consider semi-supervised scenarios, in which the amount of information about the desired output is limited.¹ For example, Gütig proposed aggregate-label learning [33[•]], in which feature-processing neurons are exposed to a (semi-)supervisory signal proportional to a desired number of spikes in a given trial. The

¹ Here, we use semi-supervised learning in a broad sense, in which learning is based on labels that do not fully specify desired output. Note that classical semi-supervised learning deals with a more specific scenario using mixture of labeled and unlabeled data.

corresponding gradient-based learning rule is a multi-spike generalization of the tempotron [34], and can be approximated by a modulated Hebbian learning rule. The aggregate-label learning was shown to work well in a simple speech recognition task. A neuron was trained on utterances with variable number of different words (digits). After training, the neuron would fire exactly one spike every time the desired word crops up in a continuous speech stream. Moreover, each spike would be time-aligned with the corresponding word, so that the spike train encodes not only the number of appearances of the desired word within the trial, but also their timings. Since no temporally annotated training data was used, this amounts to solving the temporal credit-assignment problem.

A three-factor learning rule can also perform unsupervised learning. Recurrently connected networks can learn to generate specific spiking sequences, utilizing a third factor that summarises activity of all recurrently connected neurons [35,36]. Moreover, feedforward networks can learn to separate independent sources when they receive mixtures of the sources as input [37[•]]. As we describe below, a third factor that sums the activity of output neurons is critical for performing this independent component analysis (ICA) [38] in a biologically plausible manner which may be consistent with a recent *in vitro* experiment [39].

Finally, attention may also modulate the learning rate. In the attention-gated RL [40] and the attention-gated memory tagging [41], two factors modulate the Hebbian term: a reward prediction error and an (top-down) attentional feedback. On the basis of the inputs from a hidden layer, an output layer chooses an appropriate action by the winner-take-all mechanism. The rest of the network is then informed about the selected action via feedback connections, which limits the occurrence of plasticity to affect only those synapses that were relevant in the action selection. In other words, the feedback is used to assign a credit to neurons.

Generation of the third factor signals

So far we have seen that the third factors can be useful for learning in many ways and that they are likely used by the real brain. But where do they come from? In the following we list some theoretical models which successfully incorporate internally generated third factors into the learning process.

Let us start again from the reward system described in the machine learning literature. In RL action selection is often based on the value function. While the values of states and actions can be described in a tabular form, this representation is not biological and, in most practical situations, number of possible states and actions is too large to use the tabular representation. A standard way

around this issue is to train a feedforward neural network for encoding the value function. Here again the prediction error drives the learning, but instead of a direct update of the tabular value, the momentary error in the value function is used as a third factor to supervise the neural network. In this way the supervised signal is internally generated based on the reward. The power of this approach has been demonstrated in practice, leading to human-level or superhuman performance in backgammon [42], Atari 2600 games [43], and Go [44]. On a side note, in the standard RL paradigm reward is scalar and is given externally to the agent. In real biological systems, however, rewards are evoked by the sensory inputs and are often non-scalar, that is, they may be related to different goals, like for example, feeding and reproducing. How these rewards are calculated based on the sensory inputs and then combined to produce a scalar reward signal are interesting but highly nontrivial problems [45–47].

In self-supervised spiking neural network models [33^{*}], fully unsupervised learning is achieved by means of internally generated supervisory signals. On the basis of linear summation of spike counts of the processing neurons within a trial, supervisory neurons that are assigned to individual processing neurons determine whether the assigned processing neuron should increase or decrease their number of output spikes. If each supervisory neuron receive input preferentially from geometrically nearby processing neurons, the processing neurons learn topologically organized receptive field.

In addition, some models predict that an average neural activity of a population of neurons could play an important role in the computation of the third factor. In the RL learning of two choice tasks, neural population activity can be used for the computation of the third factor to provide tailormade supervising signals for individual neurons. When a majority-vote of neurons leads to punishment, it is more efficient to provide virtual reward in the form of the third factor to the minority of neurons that suggested the other choice [48]. More generally, many third factors can be calculated as the sum of nonlinearly transformed activities of a population of neurons. For example, surprise is suitable as the third factor and can be calculated locally by summing log firing probabilities of each neuron [35,36].

To see how such a global signal acts, let us consider blind source separation of independent sources, that is, the ICA problem. Suppose the external world mixes underlying hidden independent sources \mathbf{s} by a mixing matrix A and provide to a neural network mixed signals $\mathbf{x} = A\mathbf{s}$ as sensory input (Figure 2). The neural network linearly weights the sensory input \mathbf{x} with synaptic strength matrix W and produces output $\mathbf{u} = W\mathbf{x}$. The goal of ICA is to learn

W so that the components of output \mathbf{u} become as independent as possible. It turns out that the three factor learning rule of Equation 1 can efficiently perform the task [37^{*}], and the common third-factor for all synapses is given by

$$g = \sum_k \log p(u_k) + \text{const.}, \quad (3)$$

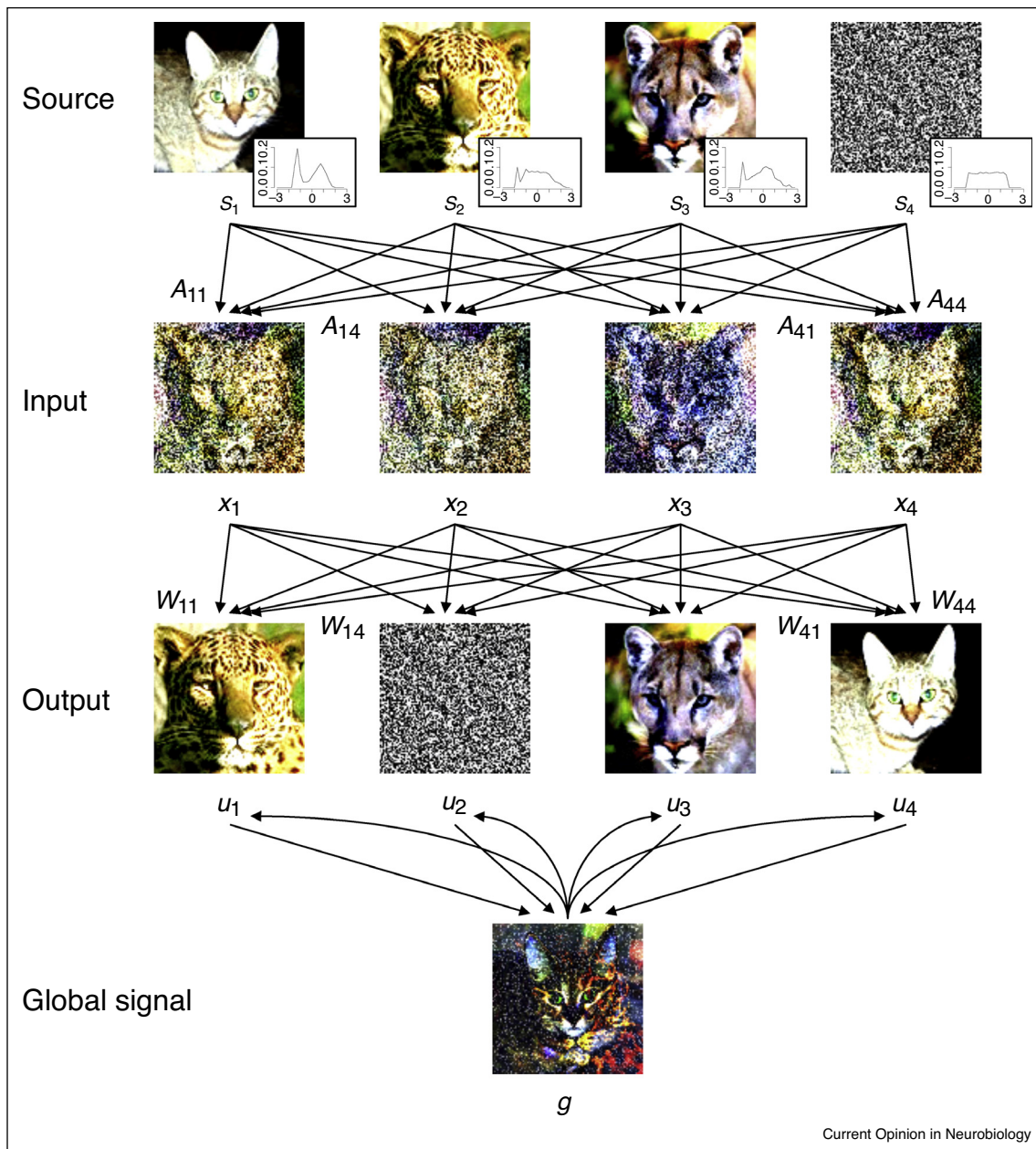
where index k runs over all output neurons and p is the probability distribution of the sources. Thus, the third factor monitors the nonlinear sum of the outputs and gates Hebbian plasticity (Figure 2). This third factor could be biologically encoded by GABA which is consistent with the experimental observation that GABAergic input can invert Hebbian to anti-Hebbian plasticity [17]. Note that the third factor in Equation 3, possibly encoded by inhibition, directly modulates Hebbian plasticity without affecting neural activity. This stands in contrast to other neural network implementations of ICA (see e.g. [49]), where inhibition indirectly affects learning through modulating neural activity.

Propagation of the third factor signals

Currently, many machine learning models have deep hierarchical structures. Typically, supervising signals are provided only to output neurons and, therefore, training of other neurons require solving the spatial credit assignment problem [51]. In order to convert the error of output neurons to that of other neurons the backpropagation (BP) algorithm [52] is often used in machine learning. The standard implementation of BP is based on selective propagation and integration of error signals from postsynaptic to presynaptic neurons through several neural layers (cf. Figure 3a). However, no known biological mechanism could support such coordinated third factor signaling. Because of the successes of BP in diverse artificial intelligence applications [53], the hypothesis that BP is also realized in the brain is alluring and has led many researchers to consider alternative, biologically more plausible implementations of BP. This pursuit has a long history [54–56] and recently many new models have been put forward, some of which are shown in (Figure 3b–d) and discussed in the following.

A new learning rule introduced in [57] draws on the two-layer perceptron interpretation of a biologically plausible neuron model [58]. A model neuron consists of a soma and multiple active dendritic branches. Each dendritic branch has its own local voltage and can generate NMDA-spikes, which is motivated by *in vitro* [59–61] and *in vivo* [62] studies. The resulting multi-compartment neuron is equivalent to a two-layer network of point neurons, in which the hidden units correspond to the dendritic branches (Figure 3b).

Figure 2

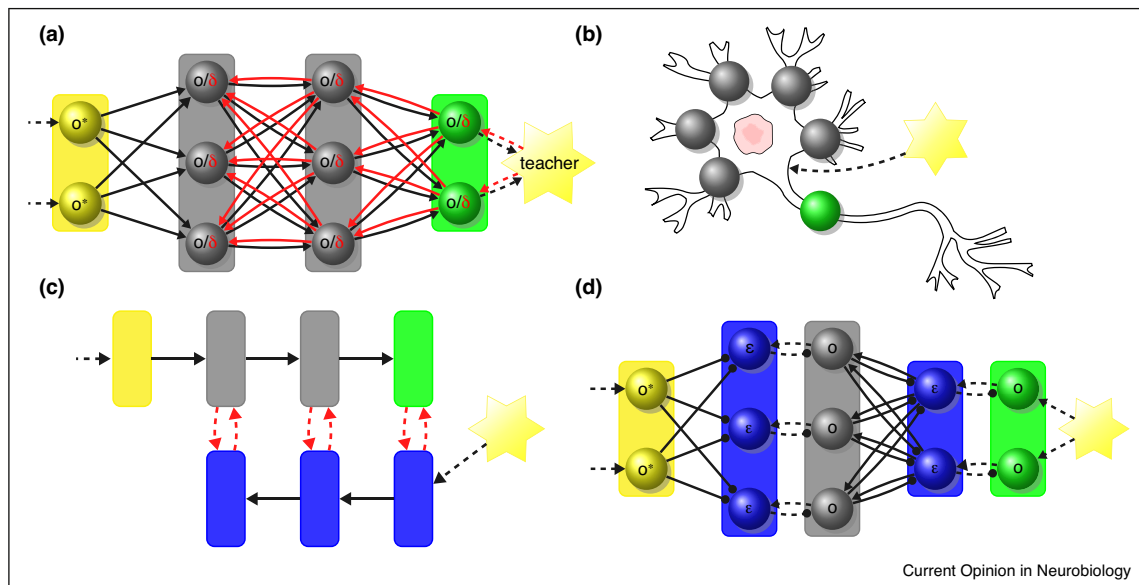


A three-factor learning rule for ICA (reprinted and modified from [37]). The mathematical model consists of hidden sources s_i ($1 \leq i \leq 4$) (the highest layer), sensory inputs $x_i = \sum_j A_{ij}s_j$ (the second layer), output neural activities $u_i = \sum_j W_{ij}x_j$ (the third layer), and a global signal g as the third-factor (the lowest layer). Note that i and j are indices of sources, inputs, and outputs; A_{ij} is an element of the mixing matrix of the generative model; W_{ij} is an element of the synaptic strength matrix. The third-factor prevents outputs from correlating with each other, through a modification of Hebbian plasticity. The figure illustrates ICA of natural images, where sensory inputs (the second layer) are generated by mixing hidden sources (the highest layer; three natural images and a noise image), and outputs of the neural network can extract hidden sources (the third layer) in the absence of supervision. We retrieved these pictures from the Caltech101 dataset [50] (http://www.vision.caltech.edu/Image_Datasets/Caltech101/) and processed them accordingly.

The supervised learning rule introduced in [57] is equivalent to BP. However, this approach cannot be directly generalized to propagate signals back across different neurons.

A more general machine learning approach with a long history is based on the idea of target propagation [63,64]. Here, during training each neuron in the trained network has access to its individual target output value, rather than

Figure 3



Different ways of propagating supervised signals. **(a)** Standard backpropagation: during forward pass a multilayer network receives signals (o) from input nodes (yellow) and propagates them via learnable weights (black solid arcs) through hidden layers (gray) to the output layer (green). In the backward pass, a teacher (yellow star) compares the outputs to their desired values and feeds back the errors (red arrows, δ). Errors are propagated back using the same weights as in the forward pass (red solid arcs). Each neuron computes its error (δ) in a different manner than its activity (o). **(b)** Multi-compartment neuron: it is equivalent to a two layer perceptron. Each dendritic tree performs nonlinear operations on its inputs and is equivalent to a point neuron. Soma and axon correspond to the output neuron. Teacher signal is provided by direct somatic inputs. **(c)** Target propagation: for training a multilayered network (top) an auxiliary network (bottom) is used. The auxiliary network is trained to match inverse of the main network. Given the desired outputs, the auxiliary network can then generate target values for hidden layers of the main network. **(d)** Predictive coding: a recurrent neural network with auxiliary nodes encoding prediction errors (ϵ). They are connected with their corresponding nodes via fixed connections (dashed arcs) with weights $+1$ (arrows) or -1 (circles). Learnable weights come in anti-symmetric pairs, arcs with circles correspond to the learnable parameters in the feedforward network with a reversed sign with respect to arcs with arrows in the feedback network. During learning both input and output nodes are clamped at the desired values. No third factor is involved here.

the propagated error. Target propagation can thus in principle be applied in nondifferentiable networks, given that we know how to guess the target value. In a recently proposed variant [63,64], target values are generated by an auxiliary neural network, which learns to implement a top-down inverse mapping with respect to the trained bottom-up network (Figure 3c). One can think of it as a way of learning the backpropagation algorithm.

Further, it has recently been proven [65] that weights of middle layer neurons established by learning in a hierarchical predictive coding model [66,67] are approximately the same as weights established by the BP algorithm. Thus, BP might be biologically plausibly implemented in the brain using such a predictive coding model (Figure 3d). Other recent models involve random feedback connections [68] and energy-based networks [56,69], see [70] for a comprehensive review of recent results.

In addition, some experimental evidence supports the existence of backpropagation in actual neural networks. Poo and colleagues have shown that a retrograde axonal

signal, for example, BDNF, generated by a pairwise firing led to changes in synaptic strengths in upstream neurons without activation [71,72]. This retrograde axonal signal could be propagated from postsynaptic to presynaptic neurons and is hypothesised to represent information required for backpropagation [73].

Conclusion

Living organisms face challenging environments in which they have to generate complex behaviors in order to survive and reproduce. This complexity can be achieved by an optimization of a limited number of internally generated cost functions [70]. Each cost function should have a corresponding error that can be broadcasted within the brain. Neuromodulators may serve as the implementation of this broadcasting mechanism. They can directly affect neural activity as well as modulate synaptic plasticity in the form of the third factor. The observed diversity of substances possibly implementing the third factor may provide the required dimensions for encoding distinct errors at different time-scales and spatial-scales. Although a lot is already known about the effects of

neuromodulators on neural activity and synaptic plasticity [74], the relation between plasticity modulation and behavior is still not fully understood. In the hope of providing guidance for experimental studies, we have reviewed putative computational roles of recently conjectured third factors as well as possible mechanisms of their calculation and propagation through the neural network. We believe that the recent advances in experimental techniques will pave the way for examining those models.

Conflict of interest statement

Nothing declared.

Acknowledgement

This work was supported by RIKEN Brain Science Institute.

References and recommended reading

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

- of special interest

1. Bi GQ, Poo MM: **Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type.** *J Neurosci* 1998, **18**:10464-10472.
 2. Markram H, Lübke J, Frotscher M, Sakmann B: **Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs.** *Science* 1997, **275**:213-215.
 3. Song S, Miller KD, Abbott LF: **Competitive Hebbian learning through spike-timing-dependent synaptic plasticity.** *Nat Neurosci* 2000, **3**:919-926.
 4. van Rossum MC, Bi GQ, Turrigiano GG: **Stable Hebbian learning from spike timing-dependent plasticity.** *J Neurosci* 2000, **20**:8812-8821.
 5. Pfister J-P, Gerstner W: **Triplets of spikes in a model of spike timing-dependent plasticity.** *J Neurosci* 2006, **26**:9673-9682.
 6. Shouval HZ, Bear MF, Cooper LN: **A unified model of NMDA receptor-dependent bidirectional synaptic plasticity.** *Proc Natl Acad Sci U S A* 2002, **99**:10831-10836.
 7. Graupner M, Brunel N: **Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location.** *Proc Natl Acad Sci U S A* 2012, **109**:3991-3996.
 8. Zhang J-C, Lau P-M, Bi G-Q: **Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses.** *Proc Natl Acad Sci U S A* 2009, **106**:13028-13033.
 9. Salgado H, Köhr G, Treviño M: **Noradrenergic "tone" determines dichotomous control of cortical spike-timing-dependent plasticity.** *Sci Rep* 2012, **2** <http://dx.doi.org/10.1038/srep00417>.
 10. 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.
 11. Johansen JP, Diaz-Mataix L, Hamanaka H, Ozawa T, Ycu E, Koivumaa J *et al.*: **Hebbian and neuromodulatory mechanisms interact to trigger associative memory formation.** *Proc Natl Acad Sci U S A* 2014, **111**:E5584-E5592.
 12. Seol GH, Ziburkus J, Huang S, Song L, Kim IT, Takamiya K *et al.*: **Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity.** *Neuron* 2007, **55**:919-929.
 13. Mansour-Robaey S, Mechawar N, Radja F, Beaulieu C, Descarries L: **Quantified distribution of serotonin transporter and receptors during the postnatal development of the rat barrel field cortex.** *Brain Res Dev Brain Res* 1998, **107**:159-163.
 14. Hayama T, Noguchi J, Watanabe S, Takahashi N, Hayashi-Takagi A, Ellis-Davies GCR *et al.*: **GABA promotes the competitive selection of dendritic spines by controlling local Ca^{2+} signaling.** *Nat Neurosci* 2013, **16**:1409-1416.
 15. Müllner FE, Wierenga CJ, Bonhoeffer T: **Precision of inhibition: dendritic inhibition by individual GABAergic synapses on hippocampal pyramidal cells is confined in space and time.** *Neuron* 2015, **87**:576-589.
 16. Schulz JM, Redgrave P, Reynolds JNJ: **Cortico-striatal spike-timing dependent plasticity after activation of subcortical pathways.** *Front Synaptic Neurosci* 2010, **2**:23.
 17. Paille V, Fino E, Du K, Morera-Herreras T, Perez S, Kotaleski JH *et al.*: **GABAergic circuits control spike-timing-dependent plasticity.** *J Neurosci* 2013, **33**:9353-9363.
 18. Ben Achour S, Pascual O: **Glia: the many ways to modulate synaptic plasticity.** *Neurochem Int* 2010, **57**:440-445.
 19. Williams RJ: **Simple statistical gradient-following algorithms for connectionist reinforcement learning.** *Mach Learn* 1992, **8**:229-256.
 20. Mazzoni P, Andersen RA, Jordan MI: **A more biologically plausible learning rule for neural networks.** *Proc Natl Acad Sci U S A* 1991, **88**:4433-4437.
 21. Sutton RS: **Learning to predict by the methods of temporal differences.** *Mach Learn* 1988, **3**:9-44.
 22. Schultz W: **Neuronal reward and decision signals: from theories to data.** *Physiol Rev* 2015, **95**:853-951.
 23. Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N: **Arithmetic and local circuitry underlying dopamine prediction errors.** *Nature* 2015, **525**:243-246.
- This work reports that dopamine neurons in the ventral tegmental area perform arithmetic subtraction for encoding reward prediction error and that local GABA neurons enable this computation.
24. Frémaux N, Gerstner W: **Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules.** *Front Neural Circuits* 2015, **9**:85.
- A review of STDP three-factor learning rules in reward and novelty based learning.
25. Pennartz CMA, Ito R, Verschure PFMJ, Battaglia FP, Robbins TW: **The hippocampal-striatal axis in learning, prediction and goal-directed behavior.** *Trends Neurosci* 2011, **34**:548-559.
 26. Glimcher PW: **Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis.** *Proc Natl Acad Sci U S A* 2011, **108**(Suppl 3):15647-15654.
 27. Ponulak F, Kasiński A: **Supervised learning in spiking neural networks with ReSuMe: sequence learning, classification, and spike shifting.** *Neural Comput* 2010, **22**:467-510.
 28. Florian RV: **The chronotron: a neuron that learns to fire temporally precise spike patterns.** *PLoS ONE* 2012, **7**:e40233.
 29. Mohemmed A, Schliebs S, Matsuda S, Kasabov N: **Span: spike pattern association neuron for learning spatio-temporal spike patterns.** *Int J Neural Syst* 2012, **22**:1250012.
 30. Yu Q, Tang H, Tan KC, Li H: **Precise-spike-driven synaptic plasticity: learning hetero-association of spatiotemporal spike patterns.** *PLOS ONE* 2013, **8**:e78318.
 31. Albers C, Westkott M, Pawelzik K: **Learning of precise spike times with homeostatic membrane potential dependent synaptic plasticity.** *PLOS ONE* 2016, **11**:e0148948.
 32. Gardner B, Grüning A: **Supervised learning in spiking neural networks for precise temporal encoding.** *PLOS ONE* 2016, **11**:e0161335.
 33. Güttig R: **Spiking neurons can discover predictive features by aggregate-label learning.** *Science* 2016, **351**:aab4113.
- A new (semi-)supervised learning rule for spiking neurons which can solve the temporal credit-assignment problem. The learning rule is based on

- the gradient of spike-threshold-surface. Unsupervised learning model with internally generated third factors.
34. Gütig R, Sompolinsky H: **The tempotron: a neuron that learns spike timing-based decisions.** *Nat Neurosci* 2006, **9**:420-428.
 35. Jimenez Rezende D, Gerstner W: **Stochastic variational learning in recurrent spiking networks.** *Front Comput Neurosci* 2014, **8**:38.
 36. Brea J, Senn W, Pfister J-P: **Matching recall and storage in sequence learning with spiking neural networks.** *J Neurosci* 2013, **33**:9565-9575.
 37. Isomura T, Toyozumi T: **A local learning rule for independent component analysis.** *Sci Rep* 2016, **6**:28073.
A biologically plausible model of a linear neural network performing ICA. The network learns to separate sources by a local learning rule with a scalar global third factor.
 38. Hyvärinen A, Karhunen J, Oja E: *Independent Component Analysis*. 2001.
 39. Isomura T, Kotani K, Jimbo Y: **Cultured cortical neurons can perform blind source separation according to the free-energy principle.** *PLoS Comput Biol* 2015, **11**:e1004643.
 40. Roelfsema PR, van Ooyen A: **Attention-gated reinforcement learning of internal representations for classification.** *Neural Comput* 2005, **17**:2176-2214.
 41. Rombouts JO, Bohte SM, Martinez-Trujillo J, Roelfsema PR: **A learning rule that explains how rewards teach attention.** *Vis Cogn* 2015, **23**:179-205.
 42. Tesauro G: **TD-Gammon, a self-teaching backgammon program, achieves master-level play.** *Neural Comput* 1994, **6**:215-219.
 43. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG et al.: **Human-level control through deep reinforcement learning.** *Nature* 2015, **518**:529-533.
 44. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G et al.: **Mastering the game of Go with deep neural networks and tree search.** *Nature* 2016, **529**:484-489.
 45. Van Moffaert K, Drugan MM, Nowe A: **Scalarized multi-objective reinforcement learning: novel design techniques.** *2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*. 2013 <http://dx.doi.org/10.1109/adpdl.2013.6615007>.
 46. Berridge KC, Kringelbach ML: **Neuroscience of affect: brain mechanisms of pleasure and displeasure.** *Curr Opin Neurobiol* 2013, **23**:294-303.
 47. Schultz W, Carelli RM, Wightman RM: **Phasic dopamine signals: from subjective reward value to formal economic utility.** *Curr Opin Behav Sci* 2015, **5**:147-154.
 48. Urbanczik R, Senn W: **Reinforcement learning in populations of spiking neurons.** *Nat Neurosci* 2009, **12**:250-252.
 49. Brito CSN, Gerstner W: **Nonlinear Hebbian learning as a unifying principle in receptive field formation.** *PLoS Comput Biol* 2016, **12**:e1005070.
 50. Fei-Fei L, Fergus R, Perona P: **Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories.** *2004 Conference on Computer Vision and Pattern Recognition Workshop*. 2004 <http://dx.doi.org/10.1109/cvpr.2004.383>.
 51. Barto AG, Sutton RS: **Reinforcement learning in artificial intelligence.** *Adv Psychol* 1997:358-386.
 52. LeCun Y, Bengio Y, Hinton G: **Deep learning.** *Nature* 2015, **521**:436-444.
 53. Schmidhuber J: **Deep learning in neural networks: an overview.** *Neural Netw* 2015, **61**:85-117.
 54. Stork DG: **Is backpropagation biologically plausible?** *International Joint Conference on Neural Networks*. 1989 <http://dx.doi.org/10.1109/ijcnn.1989.118705>.
 55. Fausett DW: **Strictly local backpropagation.** *1990 IJCNN International Joint Conference on Neural Networks*. 1990 <http://dx.doi.org/10.1109/ijcnn.1990.137834>.
 56. Xie X, Seung HS: **Equivalence of backpropagation and contrastive Hebbian learning in a layered network.** *Neural Comput* 2003, **15**:441-454.
 57. Schiess M, Urbanczik R, Senn W: **Somato-dendritic synaptic plasticity and error-backpropagation in active dendrites.** *PLoS Comput Biol* 2016, **12**:e1004638.
 58. Brunel N, Hakim V, Richardson MJE: **Single neuron dynamics and computation.** *Curr Opin Neurobiol* 2014, **25**:149-155.
 59. Schiller J, Major G, Koester HJ, Schiller Y: **NMDA spikes in basal dendrites of cortical pyramidal neurons.** *Nature* 2000:285-289 <http://dx.doi.org/10.1038/35005094>.
 60. Nevian T, Larkum ME, Polsky A, Schiller J: **Properties of basal dendrites of layer 5 pyramidal neurons: a direct patch-clamp recording study.** *Nat Neurosci* 2007, **10**:206-214.
 61. Branco T, Clark BA, Häusser M: **Dendritic discrimination of temporal input sequences in cortical neurons.** *Science* 2010, **329**:1671-1675.
 62. Lavzin M, Rapoport S, Polsky A, Garion L, Schiller J: **Nonlinear dendritic processing determines angular tuning of barrel cortex neurons in vivo.** *Nature* 2012, **490**:397-401.
 63. Lee D-H, Zhang S, Fischer A, Bengio Y: **Difference target propagation.** *Lecture Notes in Computer Science*. 2015:498-515.
 64. Bengio Y: **How auto-encoders could provide credit assignment in deep networks via target propagation.** 2014arXiv:1407.7906.
 65. Whittington J, Bogacz R: **An approximation of the error back-propagation algorithm in a predictive coding network with local Hebbian synaptic plasticity.** *bioRxiv* 2016 <http://dx.doi.org/10.1101/035451>.
Authors show equivalence of learning in a predictive coding scheme to BP. The proof is based on the assumption that pairs of excitatory and inhibitory weights have the same strength, although the authors report that weakening this assumption does not preclude the network from learning. This seems to be in line with a recent report showing that back-propagating the error through fixed random weights (different from feedforward weights) does not harm the learning [68].
 66. Rao RP, Ballard DH: **Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects.** *Nat Neurosci* 1999, **2**:79-87.
 67. Friston K: **The free-energy principle: a unified brain theory?** *Nat Rev Neurosci* 2010, **11**:127-138.
 68. Lillicrap TP, Cownden D, Tweed DB, Akerman CJ: **Random synaptic feedback weights support error backpropagation for deep learning.** *Nat Commun* 2016, **7**:13276.
 69. Scellier B, Bengio Y: **Equilibrium propagation: bridging the gap between energy-based models and backpropagation.** 2016arXiv:1602.05179.
 70. Marblestone AH, Wayne G, Kording KP: **Towards an integration of deep learning and neuroscience.** 2016 <http://dx.doi.org/10.1101/058545>.
A comprehensive review aiming at integrating neuroscience and machine learning. It brings together models from both communities and discusses them in a common framework. It includes a comparison of biologically plausible implementations of backpropagation.
 71. Fitzsimonds RM, Song HJ, Poo MM: **Propagation of activity-dependent synaptic depression in simple neural networks.** *Nature* 1997, **388**:439-448.
 72. Du J-L, Poo M-M: **Rapid BDNF-induced retrograde synaptic modification in a developing retinotectal system.** *Nature* 2004, **429**:878-883.
 73. Harris KD: **Stability of the fittest: organizing learning through retroaxonal signals.** *Trends Neurosci* 2008, **31**:130-136.
 74. Nadim F, Bucher D: **Neuromodulation of neurons and synapses.** *Curr Opin Neurobiol* 2014, **29**:48-56.