Can neocortical feedback alter the sign of plasticity?

Blake A. Richards and *Timothy P. Lillicrap*

The recent Review by Roelfsema and Holtmaat (Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.* **19**, 166–180 (2018))[1] provides a much-needed guide to learning in deep cortical networks. The importance of credit assignment for deep cortical networks has come into focus recently with the success of deep learning in artificial intelligence[2]. The learning algorithm that is typically used for credit assignment in deep artificial neural networks, the backpropagation-of-error algorithm[3], is biologically infeasible[4]. Yet, its successful application to complicated tasks suggests that credit assignment is important for learning in non-trivial circumstances, whether in artificial or biological neural networks.

As Roelfsema and Holtmaat note, the focus in neuroscience to date has been either on Hebbian plasticity mechanisms[5], or three-factor Hebbian plasticity rules that incorporate a global reward prediction error[6]. They argue that a more powerful approach is to use feedback signals that can determine credit on a neuron-by-neuron basis. We agree with the authors that this is a likely role for feedback connections in learning. A question that is left open, though, is whether feedback signals act purely as a gating mechanism, and if so, is that enough to solve the credit assignment problem?

According to one framework that Roelfsema and Holtmaat explore in their paper, changes (Δ) in the strength of a synapse between presynaptic neuron i and postsynaptic neuron j (w_{ij}) are guided by a four-term equation (which we simplify here slightly):

$$\Delta w_{ij} = f_i \cdot f_j \cdot RPE \cdot FB_j \tag{1}$$

where f_i and f_j are functions of presynaptic and postsynaptic activity, respectively, RPE is a global reward prediction error communicated via neuromodulators and FB_j is the feedback received by the postsynaptic neuron. In the Review, Roelfsema and Holtmaat suggest that the feedback signal, FB_j, could be a gating signal ranging from 0 to 1, such that

it can turn synaptic plasticity in a neuron on or off but cannot alter the sign of synaptic plasticity (for example, whether synapses potentiate or depress). Instead, the term RPE determines the sign of plasticity. However, in our opinion it may be important for FB_j to determine whether neurons potentiate or depress their synapses.

Roelfsema and Holtmaat state that the weight changes from equation 1 are equivalent to those prescribed by backpropagation-of-error, but the equivalence is on the weight changes on average, and this point is crucial. Notably, even random search algorithms, such as REINFORCE[7], also agree with backpropagation-of-error on average. This means that for an individual stimulus, algorithms like REINFORCE or those prescribed by equation 1 do not follow the true gradient. Instead, these algorithms only follow the true gradient when their synaptic weight updates are averaged across many repetitions of the same stimuli. Put another way, equation 1 uses an estimator of the true gradient followed by backpropagation-of-error. There are two key questions pertinent to this approach. First, what is the variance of this estimator? Second, how long does it take to reach good performance for a given task[8]? We speculate that if a task requires learning a high-dimensional, complex function, the variance of the estimator will be high and it will take an intractably long time to reach reasonable levels of performance. For example, to the best of our knowledge, there are no examples in the literature of successfully training a good ImageNet classifier using REINFORCE-like algorithms. Algorithms like AGREL[9], which uses feedback-based gating, can have better variance properties than REINFORCE, but whether the variance in the estimator is small enough to learn high-dimensional, complex tasks in a reasonable amount of time remains to be determined.

Given these considerations, we propose that neuroscientists should consider how feedback in the neocortex may do more than act as a gating mechanism. In other words, we

postulate that neocortical feedback may be set up to communicate signed credit signals that cause some neurons to potentiate and others to depress. One possibility is to use the temporal order of feedback onto specific dendrites as a signal of sign[10,11]. Another possibility is to use inhibitory interneuron circuits to calculate a difference[12]. Ultimately, we believe that neuroscientists should not assume that feedback acts only as a gating mechanism. Importantly, we are not arguing that feedback never acts as a gating signal. Indeed, recent evidence from the Holtmaat group shows feedback-based gating of plasticity[13], although this does not preclude signed credit assignment. Prejudging that possibility could lead our investigations on this important topic astray.

There is a reply to this letter by Roelfsema, P. R. & Holtmaat, A. *Nat. Rev. Neurosci.* https://doi.org/10.1038/s41583-018-0048-6 (2018).

Blake A. Richards [1,2]* *and Timothy P. Lillicrap*[3]
[1]*Department of Biological Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada.*
[2]*Learning in Machines and Brains Program, Canadian Institute for Advanced Research, Toronto, Ontario, Canada.*
[3]*DeepMind Inc., London, UK.*
*e-mail: blake.richards@utoronto.ca

1. Roelfsema, P. R. & Holtmaat, A. Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.* **19**, 166–180 (2018).
2. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
3. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
4. Crick, F. The recent excitement about neural networks. *Nature* **337**, 129–132 (1989).
5. Bi, G. & Poo, M. Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* **24**, 139–166 (2001).
6. Frêmaux, N. & Gerstner, W. Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Front. Neural Circuits* **9**, 85 (2016).
7. Williams, R. J. *Proc. IEEE Int. Conf. Neural Networks* **1**, 263–270 (San Diego, CA., 1988).
8. Werfel, J., Xie, X. & Seung, H. S. Learning curves for stochastic gradient descent in linear feedforward networks. *Neural Comput.* **17**, 2699–2718 (2005).
9. Roelfsema, P. R. & van Ooyen, A. Attention-gated reinforcement learning of internal representations for classification. *Neural Comput.* **17**, 2176–2214 (2005).
10. Guerguiev, J., Lillicrap, T. P. & Richards, B. A. Towards deep learning with segregated dendrites. *eLife* **6**, e22901 (2017).
11. Bono, J. & Clopath, C. Modeling somatic and dendritic spike mediated plasticity at the single neuron and network level. *Nat. Commun.* **8**, 706 (2017).
12. Sacramento, J., Costa, R. P., Bengio, Y. & Senn, W. Dendritic error backpropagation in deep cortical microcircuits. *ArXiv* 1801.00062 (2017).
13. Williams, L. E. & Holtmaat, A. Higher-order thalamocortical inputs gate synaptic long-term potentiation via disin-hibiton. *bioRxiv* https://doi.org/10.1101/281477 (2018).