Reply to 'Can neocortical feedback alter the sign of plasticity?'

Pieter R. Roelfsema and *Anthony Holtmaat*

In our recent Review (Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.* **19**, 166–180 (2018))[1], we reviewed factors that influence synaptic plasticity in sensory and association cortex during reinforcement learning. We asked how the brain can implement powerful learning rules such as error-backpropagation (BP), a rule that is broadly used to train artificial neural networks to perform complex tasks. We described how feedback signals originating from the response-selection stage might tag those synapses at deeper network levels that are responsible for the outcome of actions, thereby gating their plasticity. In their Correspondence (Can neocortical feedback alter the sign of plasticity? *Nat. Rev. Neurosci.* https://doi.org/10.1038/s41583-018-0049-5 (2018))[2], Richards and Lillicrap suggest that feedback connections might not only gate but also change the sign of synaptic plasticity. Here we address this important assertion with a closer examination of the proposed learning rules. Our analysis indicates that the proposed learning rules indeed predict that feedback connections also influence the sign of plasticity.

Previous studies[3,4] demonstrated that the changes in the strengths of synapses ($\Delta w_{i,j}$) required by BP can be split into separable factors, which are available locally, at cortical synapses.

$$\Delta w_{i,j} = \beta \cdot f_i(a_i) \cdot f_j(a_j) \cdot RPE \cdot FB_j \qquad (1)$$

Here, $f_i(a_i)$ is a function of the presynaptic activity level a_i, and $f_j(a_j)$ is a function of the postsynaptic activity level a_j. The RPE is the reward prediction error, which can be mediated by neuromodulators, such as dopamine[5,6]. Previous studies suggested that the RPE steers plasticity: it is positive if the outcome of an action is better than expected and negative if it is worse. In our Review, we suggested that FB_j gates plasticity: it indicates the degree to which a synapse can be held responsible for the outcome of an action. If a neuron does not receive feedback from the response selection stage, its synapses will not change in strength. An important aspect of the learning rule of equation 1 is that it supports trial-and-error learning of new tasks. There is no need for external

supervision that tells the network which units should be switched on and off, unlike several previous models[7,8] that propose a role for feedback connections in modulating synaptic plasticity. Importantly, these biologically plausible plasticity rules, like AuGMEnT[4], enable the learning of high-dimensional functions within a reasonable amount of time because they have better variance properties than previous models such as REINFORCE[3,9]. Indeed, AuGMEnT allows artificial neural networks to learn complex and nonlinear stimulus–response mappings with a learning speed that is faster than monkeys trained on the same tasks[4,10]. However, animals also learn the statistics of stimuli without a specific reward structure; that is, outside of the reinforcement learning contexts. Future studies could investigate whether the beneficial effects of feedback connections in the gating of plasticity also generalize to other learning schemes, such as unsupervised learning, and if and how such learning schemes can train the multilayered networks of the brain within a realistic time frame.

In their Correspondence, Richards and Lillicrap propose to also consider learning rules in which feedback connections not only gate plasticity but also can change the sign of synaptic plasticity. FIGURE 1 illustrates a neural network that selected action 'a' in the output layer. We ask how the learning rule affects the input to the association neurons y_1–y_4, which project to the output layer with excitatory (y_1) or inhibitory connections (y_2). Previous work[3,4] concerning the learning rule of equation 1 suggested that feedback connections should be (or become) proportional in strength to the feedforward connections. Hence, the feedback connection from *a* back to y_1 should be excitatory, and the feedback connection from *a* to y_2 inhibitory (there are no long-range inhibitory connections between remote cortical areas, but long-range excitatory projections can provide inhibition to a cortical column by activating local interneurons; see below). Let us now examine whether this learning rule is compatible with a sign-changing effect of feedback connections.

First, we consider the plasticity of connections onto excitatory neuron y_1. When

the outcome of action *a* is better than expected, all factors in equation 1 are positive, such that $\Delta w_{1,1}$ is also positive: the synapse strengthens. By contrast, if the RPE is negative, the connection would weaken. Units x_3 and y_3 are also active in the example, but unit y_3 does not receive feedback from *a* and thus $w_{3,3}$ is unaltered. Indeed, y_3 did not provide input to *a* and is not responsible for the outcome. This is what we imply with a 'gating effect': feedback connections switch plasticity on or off.

These considerations should not distract from the possibility that the plasticity rule may work differently for inhibitory neurons. Unit y_2 in FIG. 1 inhibits action *a*, and the selected action now sends a negative feedback signal back to y_2. If the RPE is positive, this negative feedback signal causes $\Delta w_{1,2}$ and $\Delta w_{2,2}$ to be negative and excitatory connections $w_{1,2}$ and $w_{2,2}$ therefore weaken. However, $w_{4,4}$ does not change, because it does not receive feedback, representing another instance of a gating effect. Hence, the learning rule of equation 1 actually conforms with the suggestion by Richards and Lillicrap at the level of the cortical column, because the plasticity rules for connections onto excitatory and inhibitory neurons differ.

The neural network model in FIG. 1 provides an abstraction of the complex interconnectivity between cortical neurons, where virtually all long-range connections between brain areas are excitatory. The long-range connections target pyramidal cells to excite a cortical column, particular inhibitory cell types to inhibit the column and other inhibitory cell types to cause disinhibition. It will be of great interest to test if the predicted learning rules for inhibitory and excitatory neurons can be confirmed in experimental work.
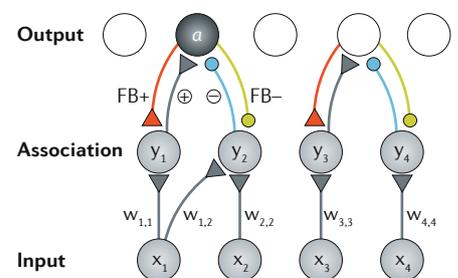

Fig. 1 | **Feedback connections have different influences on the plasticity of inputs onto excitatory and inhibitory neurons.** Neural network with three layers, in which action *a* has been selected in the output layer. Excitatory neurons and connections are shown in black and red, and inhibitory neurons and connections in cyan and green. FB+, excitatory feedback connection; FB−, inhibitory feedback connection.

CORRESPONDENCE

Pieter R. Roelfsema [iD] [1,2,3]* *and Anthony Holtmaat* [4]
[1]*Department of Vision and Cognition, Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences, Amsterdam, Netherlands.*
[2]*Department of Integrative Neurophysiology, VU University, Amsterdam Neuroscience, Amsterdam, Netherlands.*
[3]*Psychiatry Department, Amsterdam University Medical Center, Amsterdam, Netherlands.*
[4]*Department of Basic Neurosciences, Geneva Neuroscience Center, Faculty of Medicine, University of Geneva, Geneva, Switzerland.*
e-mail: p.roelfsema@nin.knaw.nl

1. Roelfsema, P. R. & Holtmaat, A. Control of synaptic plasticity in deep cortical networks. *Nat. Rev. Neurosci.* **19**, 166–180 (2018).
2. Richards, B. A. & Lillicrap, T. P. Can neocortical feedback alter the sign of plasticity? *Nat. Rev. Neurosci.* **19** https://doi.org/10.1038/s41583-018-0049-5 (2018).
3. Roelfsema, P. R. & Van Ooyen, A. Attention-gated reinforcement learning of internal representations for classification. *Neural Comput.* **17**, 2176–2214 (2005).
4. Rombouts, J. O., Bohte, S. M. & Roelfsema, P. R. How attention can create synaptic tags for the learning of working memories in sequential tasks. *PLOS Comput. Biol.* **11**, e1004060 (2015).
5. Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241–263 (2002).
6. Niv, Y. & Schoenbaum, G. Dialogues on prediction errors. *Trends Cogn. Sci.* **12**, 265–272 (2008).
7. Guergiuev, J., Lillicrap, T. P. & Richards, B. A. Towards deep learning with segregated dendrites. *eLife* **6**, e22901 (2016).
8. Urbanczik, R. & Senn, W. Learning by the dendritic prediction of somatic spiking. *Neuron* **81**, 521–528 (2014).
9. Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **8**, 229–256 (1992).
10. Rombouts, J. O. et al. A learning rule that explains how rewards teach attention. *Vis. Cogn.* **23**, 179–205 (2015).

Competing interests
The authors declare no competing interests.