Reconsidering the Imaging Evidence Used to Implicate Prediction Error as the Driving Force behind Learning

Jiří Čevora and Richard N. Henson *

MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom

In this paper, we review the evidence that learning is driven by signaling of Prediction Error [PE] by some neurons. We model associative learning in artificial neural networks using Hebbian (non-PE) learning algorithms to investigate whether the data used to implicate PE in learning can arise without actual PE computation. We conclude that the metabolic demands of synaptic change during Hebbian learning would produce a PE-correlated component in functional magnetic resonance imaging (fMRI), which suggests that the research used to imply PE in learning is currently inconclusive.

There is a considerable body of evidence that PE is computed by dopaminergic neurones in ventral midbrain. Single-cell recordings have shown neurons that are excited by unexpected reward, and depressed by unexpected lack of reward (Schultz et al., 1997). This response implies reward PE computation takes place in the brain; however, it does not imply that the PE signal is utilized during learning, and no single-cell study, to our knowledge, has demonstrated this link to learning. Furthermore, these findings have only been obtained with regard to rewarded behavior, while the majority of learning in humans happens in absence of reward (Tolman, 1932).

These concerns can be potentially addressed in fMRI studies by relating a PE-related component of fMRI to subsequent memory, with or without overt rewards. Unfortunately, most fMRI research has focused simply on replicating the single-cell findings by identifying a correlate of PE in the human brain (e.g., McClure et al., 2003; Abler et al., 2006; D'Ardenne et al., 2008), without assessing its effect on behavior. We are only aware of two fMRI studies that attempted to go beyond the single-cell recording findings by demonstrating an effect of a PE-related component in fMRI on learning (Gläscher et al., 2010; McGuire et al., 2014). Both of these studies identify a component of the fMRI signal that is correlated with trial-by-trial estimates of PE from an assumed learning model, and then link that component to subsequent decision-making.

However, PE-correlated fMRI signal does not necessarily originate from PE computation: the BOLD signal measured by fMRI may relate to metabolic changes that are only indirectly related to neural activity. One of the major factors contributing to the BOLD signal is cellular respiration associated mainly with ATP metabolism (Aubert and Costalat, 2002), which is elicited by a large number of cellular processes. Synaptic plasticity has several components working at different timescales (Collingridge et al., 2004), but there are four notable processes that operate at the timescale of these studies: (a) Synaptic transmission of signal, (b) facilitation, which is an important form of short-term synaptic plasticity (Kandel, 2001), (c) migration of receptors, which is crucial components of long-term potentiation and depression (Collingridge et al., 2004), and (d) fast forms of homeostatic activity, which serve as a form of global synaptic scaling and metaplasticity (Pérez-Otaño and Ehlers, 2005). While synaptic transmission (a) is the main energy expense during

signaling (up to 55% of signaling cost, Harris et al., 2012), synaptic plasticity (b–d) can increase signaling efficiency up to hundred-fold (Harris et al., 2012) and therefore be expected to have a significant energy budget. Many of these synaptic processes occur rapidly (Collingridge et al., 2004), and could therefore take place within the same timewindow (resolvable by fMRI) as any neural activity related to PE. Thus while the actual energy consumption of synaptic plasticity is unknown (Harris et al., 2012), we conclude that there is a distinct possibility that it is sufficiently large to contribute to the BOLD response.

The outstanding question for this alternative explanation is why synaptic plasticity would correlate with PE, unless PE were computed and used to update synapses. In what follows, we model synaptic plasticity as the magnitude of Hebbian weight update in associative networks, and demonstrate that this quantity correlates with PE even when the learning algorithm does not compute PE.

We consider a modified Hebbian learning rule that includes weight decay term, also called Oja's rule (Equation 1, Oja, 1982). This learning rule does not use the current state of network (e.g., predictions) to inform learning in any way. The only modification from the classic Hebbian algorithm is that the weights decrease linearly at each time step, which is the minimal modification necessary to obtain stable and biologically plausible learning dynamics. We contrast this variant of Hebbian learning with Widrow-Hoff learning algorithm, also modified to include decay to increase its biological plausibility (e.g., Rumelhart and McClelland, 1988) as shown in Equation (2). The formulation of Widrow-Hoff learning rule used here is essentially Hebbian learning scaled by PE. In these equations, w_{ij} refers to the weight between unit i (representing the cue) and unit j (representing the outcome), a_i/a_j refer to the activity of unit i/j, t_j refers to a desired output of unit j, $0 \leq k < 1$ is the learning rate, $0 < d \leq 1$ is the decay rate and the H and WH superscripts refer to Hebbian or Widrow-Hoff learning rules, respectively.

$$\Delta w_{ij}^{H} = -d^{H} w_{ij} + k^{H} a_i a_j \tag{1}$$

$$\Delta w_{ij}^{WH} = -d^{WH} w_{ij} + k^{WH} a_i \left(t_j - \sum_{i'} (w_{i'j} a_{i'}) \right) \tag{2}$$

First, we address the relationship between learning under Hebbian and Widrow-Hoff rules in an experiment conducted by McGuire et al. (2014). The parameter estimation task they used is effectively associative learning with a single cue, because the participants' task was simply to predict the value of a parameter during each trial. As only one stimulus exists in this paradigm, the i subscript becomes redundant, therefore we can say that $a_j = w_j$ and both H and WH learning rules can be simplified to

$$\Delta w_{j}^{H'} = -d^{H'} w_j + k^{H'} a_j \tag{3}$$

and

$$\Delta w_{j}^{WH'} = -d^{WH'} w_j + k^{WH'} (t_j - w_j). \tag{4}$$

By equating $\Delta w_{j}^{H'} = \Delta w_{j}^{WH'}$, we can see that this statement is true whenever $k^{H'} = k^{WH'}$ and $d^{WH'} + k^{WH'} = d^{H'}$. This

means that in parameter estimation tasks, learning according to the Widrow-Hoff rule can be perfectly mimicked by a Hebbian rule. Therefore, performance on this task cannot be used to argue for PE learning.

This proof cannot be extended to experiments with multiple cues (Čevora, 2017), such as the one by Gläscher et al. (2010). We therefore turn to computational simulations to investigate whether there is a correlation between Hebbian weight update and prediction error. In these simulations we look at whether PE correlates with weight update. Because fMRI observes entire populations of neurons, in contrast to single-cell recordings, we need to specify the variables of interest at the population level too.

We only consider the magnitude of the population weight change, $|\Delta W^H|$, because the fast decreases in synaptic strength are likely to require a similar amount of ATP as increases (Kandel, 2001; Collingridge et al., 2004) thus producing the same BOLD signal. Therefore the change associated with trial τ is:

$$|\Delta W^H(\tau)| = \sum_i \sum_j |w_{ij}(\tau) - w_{ij}(\tau - 1)| . \tag{5}$$

Likewise, we only consider the magnitude of the population PE, given that both positive and negative PE is likely to have metabolic consequences. We define this quantity, $|PE|$, as the sum of the absolute values of differences between predictions for each possible outcome, $\|a_j\|$, and the corresponding target values t_j, on the current trial:

$$|PE| = \sum_j \left| t_j - \|a_j\| \right| , \tag{6}$$
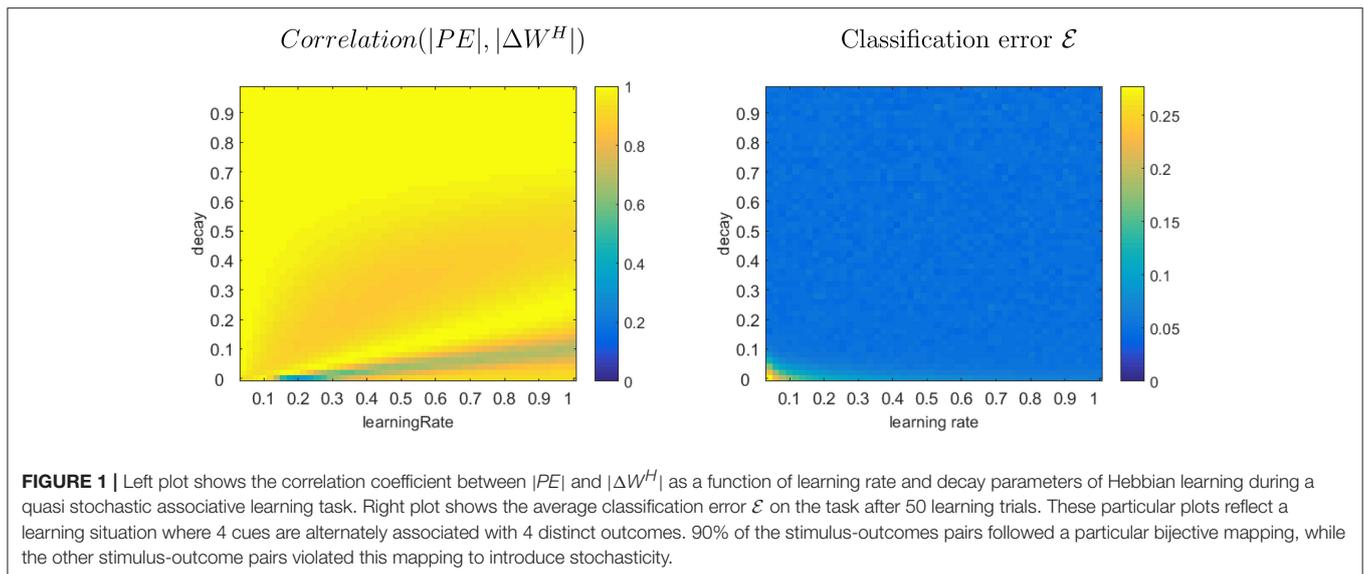
where the prediction $\|a_j\|$:

$$\|a_j\| = \frac{\sum_i a_i w_{ij}}{\sum_j \sum_{i'} a_{i'} w_{i'j}} \tag{7}$$

is a normalized activation vector as parameterizations of Hebbian learning do not produce predictions that can be interpreted directly as probabilities.

Another quantity of interest is resulting classification error after learning $\mathcal{E}$. This is defined as a magnitude of difference between prediction and true (noiseless) outcome for each cue $\mathcal{C}$, thus not only capturing how well the learning model can remember observations, but also how resilient it is to noise during learning:

$$\mathcal{E} = \sum_{\mathcal{C}} \left(\sum_j t_j^{\mathcal{C}} - \|a_j^{\mathcal{C}}\| \right) . \tag{8}$$

Simulations were conducted for a number of possible experimental designs, for both categorical and continuous associative learning, with various degrees of stochasticity and various numbers of cues/outcomes. The simulations were run across the range of values for learning rate and weight decay that produce plausible learning dynamics (**Figure 1**). We recorded

FIGURE 1 | Left plot shows the correlation coefficient between $|PE|$ and $|\Delta W^H|$ as a function of learning rate and decay parameters of Hebbian learning during a quasi stochastic associative learning task. Right plot shows the average classification error $\mathcal{E}$ on the task after 50 learning trials. These particular plots reflect a learning situation where 4 cues are alternately associated with 4 distinct outcomes. 90% of the stimulus-outcomes pairs followed a particular bijective mapping, while the other stimulus-outcome pairs violated this mapping to introduce stochasticity.

$|\Delta W^H|$ and $|PE|$ on each trial, and calculated the correlation between them.

The resulting correlations, plotted as a function of learning rate and weight decay, reveal that most of the parameter space results in strong correlations (**Figure 1**). Moreover, the classification error $\mathcal{E}$ is almost identical across the parameter space (except for a region in bottom left where both parameters are near zero), and therefore almost all parameter combinations are equally plausible for a real learner that tunes its learning parameters to the task. In other words, it is not the case that situations in which $|PE|$ and $|\Delta W^H|$ are highly correlated are non-optimal.

We conclude that, while there is convincing evidence that PE is computed by some neurons, the current evidence used to implicate this neural PE signal in learning has alternative explanations. There are a few fMRI studies that correlate brain activity with PE, a subset of which go further and link this to learning outcomes. However, due to the nature of BOLD signal measured by fMRI, the correlation with PE may not be a result of actual PE signaling, but rather a result of metabolic processes

related to synaptic plasticity: Our computational modeling demonstrates that the magnitude of synaptic plasticity is highly correlated to PE, even when no PE computation takes place during learning. Note that this article has not addressed the role of PE at higher levels of analysis, such as the computational level (Marr, 1982), though one would expect such computation to still have a neural correlate. Further modeling and experimental paradigms are therefore needed to establish the basic form of learning rules underlying human associative learning.

AUTHOR CONTRIBUTIONS

All work was carried out by JČ as a part of his Ph.D. under the supervision of RH.

FUNDING

REFERENCES

Abler, B., Walter, H., Erk, S., Kammerer, H., and Spitzer, M. (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage* 31, 790–795. doi: 10.1016/j.neuroimage.2006. 01.001

Aubert, A., and Costalat, R. (2002). A model of the coupling between brain electrical activity, metabolism, and hemodynamics: application to the interpretation of functional neuroimaging. *Neuroimage* 17, 1162–1181. doi: 10.1006/nimg.2002.1224

Collingridge, G. L., Isaac, J. T., and Wang, Y. T. (2004). Receptor trafficking and synaptic plasticity. *Nat. Rev. Neurosci.* 5, 952–962. doi: 10.1038/nr n1556

D'Ardenne, K., McClure, S. M., Nystrom, L. E., and Cohen, J. D. (2008). Bold responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319, 1264–1267. doi: 10.1126/science.1150605

Gläscher, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595. doi: 10.1016/j.neuron.2010.04.016

Harris, J. J., Jolivet, R., and Attwell, D. (2012). Synaptic energy use and supply. *Neuron* 75, 762–777. doi: 10.1016/j.neuron.2012. 08.019

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294, 1030–1038. doi: 10.1126/science.10 67020

Marr, D. (1982). *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. New York, NY: W. H. Freeman and Company.

McClure, S. M., Berns, G. S., and Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346. doi: 10.1016/S0896-6273(03)00154-5

McGuire, J. T., Nassar, M. R., Gold, J. I., and Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 84, 870–881. doi: 10.1016/j.neuron.2014.10.013

Oja, E. (1982). Simplified neuron model as a principal component analyzer. *J. Math. Biol.* 15, 267–273. doi: 10.1007/BF00275687

Pérez-Otaño, I., and Ehlers, M. D. (2005). Homeostatic plasticity and nmda receptor trafficking. *Trends Neurosci.* 28, 229–238. doi: 10.1016/j.tins.2005.03.004

Rumelhart, D. E., and McClelland, J. L. (1988). *Parallel Distributed Processing*, vol. 1. Cambridge, MA: IEEE.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593

Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. New York, NY: University of California Press.

Čevora, J. (2017). *The role of Prediction Errors in Associative Learning*. Ph.D. thesis, University of Cambridge, Cambridge University Library, West Road, Cambridge.