



Models of Acetylcholine and Dopamine Signals Differentially Improve Neural Representations

Raphaël Holca-Lamarre^{1,2*}, Jörg Lücke^{3,4†} and Klaus Obermayer^{1,2†}

¹ Neural Information Processing Group, Fakultät IV, Technische Universität Berlin, Berlin, Germany, ² Bernstein Center for Computational Neuroscience, Berlin, Germany, ³ Cluster of Excellence Hearing4all and Research Center Neurosensory Science, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, ⁴ Machine Learning Lab, Department of Medical Physics and Acoustics, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

OPEN ACCESS

Edited by:

Sander Bohte,
Centrum Wiskunde & Informatica,
Netherlands

Reviewed by:

Malte J. Rasch,
Beijing Normal University, China
Wouter Kruijine,
VU University Amsterdam,
Netherlands

*Correspondence:

Raphaël Holca-Lamarre
raphael@bccn-berlin.de

† Joint last authorship

Received: 01 December 2016

Accepted: 07 June 2017

Published: 22 June 2017

Citation:

Holca-Lamarre R, Lücke J and Obermayer K (2017) Models of Acetylcholine and Dopamine Signals Differentially Improve Neural Representations. *Front. Comput. Neurosci.* 11:54. doi: 10.3389/fncom.2017.00054

Biological and artificial neural networks (ANNs) represent input signals as patterns of neural activity. In biology, neuromodulators can trigger important reorganizations of these neural representations. For instance, pairing a stimulus with the release of either acetylcholine (ACh) or dopamine (DA) evokes long lasting increases in the responses of neurons to the paired stimulus. The functional roles of ACh and DA in rearranging representations remain largely unknown. Here, we address this question using a Hebbian-learning neural network model. Our aim is both to gain a functional understanding of ACh and DA transmission in shaping biological representations and to explore neuromodulator-inspired learning rules for ANNs. We model the effects of ACh and DA on synaptic plasticity and confirm that stimuli coinciding with greater neuromodulator activation are over represented in the network. We then simulate the physiological release schedules of ACh and DA. We measure the impact of neuromodulator release on the network's representation and on its performance on a classification task. We find that ACh and DA trigger distinct changes in neural representations that both improve performance. The putative ACh signal redistributes neural preferences so that more neurons encode stimulus classes that are challenging for the network. The putative DA signal adapts synaptic weights so that they better match the classes of the task at hand. Our model thus offers a functional explanation for the effects of ACh and DA on cortical representations. Additionally, our learning algorithm yields performances comparable to those of state-of-the-art optimisation methods in multi-layer perceptrons while requiring weaker supervision signals and interacting with synaptically-local weight updates.

Keywords: acetylcholine, dopamine, neuromodulator, sensory representations, neural networks, biology-inspired learning, representation learning

1. INTRODUCTION

Neurons in the cortex represent countless features of sensory signals, from the frequencies of photons falling on the retina to high-level attributes like quantities and numbers. The particular form a sensory representation takes is critical to perception. For instance, experienced musicians display enhanced sensory representations which putatively explain their finer perceptual abilities

(Elbert et al., 1995; Pantev et al., 1998, 2001). This view is further supported by the observation that, following discrimination training, improvements in perceptual sensitivity correlate with the degree of reorganization in cortical representations (Recanzone et al., 1992, 1993; Weinberger, 2003; Polley et al., 2006). On the other hand, perceptual disorders like phantom limb pain (Ramachandran et al., 1992; Halligan et al., 1993; Flor et al., 2006) or tinnitus (Eggermont and Roberts, 2004) appear to be correlates of degenerate sensory representations.

In animals, sensory representations undergo modifications in various circumstances, for instance following extensive perceptual training (Weinberger and Bakin, 1998; Harris et al., 2001; Schoups et al., 2001; Fletcher and Wilson, 2002; Fritz et al., 2003; Wang et al., 2003; Bao et al., 2004; Yang and Maunsell, 2004; Polley et al., 2006; Poort et al., 2015), repeated sensory exposure (Han et al., 2007; Kim and Bao, 2009), cortical stimulation (Godde et al., 2002; Dinse et al., 2003; Tegenthoff et al., 2005), or sensory deprivation (Calford and Tweedale, 1988; Allard et al., 1991; Gambino and Holtmaat, 2012). Additionally, the neuromodulators acetylcholine (ACh) and dopamine (DA) bear potent effects on cortical representations. In particular, repeated efflux of either ACh (Kilgard and Merzenich, 1998a; Froemke et al., 2007, 2013; Gu, 2003; Weinberger, 2003) or DA (Bao et al., 2001; Frankó et al., 2010) coinciding with a stimulus strengthens the responses of neurons to this stimulus and enlarges its cortical representation.

ACh and DA are critical to forms of learning which require modifications of sensory representations. For instance, lesion of the cholinergic (Butt and Hodge, 1995; Fletcher and Wilson, 2002; Conner et al., 2003; Wilson et al., 2004; Conner et al., 2010) or dopaminergic (Kudoh and Shibuki, 2006; Molina-Luna et al., 2009; Hosp et al., 2011; Luft and Schwarz, 2009; Schicknick et al., 2012) system disrupts perceptual and motor learning as well as the associated plasticity in cortical maps. These observations suggest that the neuromodulators orchestrate plastic changes that refine cortical representations and give rise to perceptual and motor learning.

In physiological conditions, ACh transmission appears to signal attentional effort, a construct reflecting both the relevance and difficulty of a task (Himmelheber et al., 2000; Arnold et al., 2002; Kozak et al., 2006; Sarter et al., 2006). DA carries information relative to reward-prediction errors (RPEs) (Schultz et al., 1997; Schultz, 2007, 2010). Although their release properties are relatively well defined, the functional roles these signals serve in shaping neural representations is unclear.

Much like the cortex, artificial neural networks (ANNs) represent input data in the form of neural activation. As for other machine learning algorithms, the performance of ANNs critically depends on the representation data take. The most widely used learning rule for ANNs, the error back-propagation algorithm (Werbos, 1974; Rumelhart et al., 1985), learns representations optimised for specific tasks. Although, the back-propagation algorithm yields remarkable performances, it is unlikely to be implemented in biological neural structures and it also bears its own limitations. For instance, in order to compute the error function, a target output must be specified

for each training example, making training data expensive to acquire. Additionally, weight updates require information not available locally at the weights which limits the use of the back-propagation algorithm in physical devices like neuromorphic chips.

In the present work we explore the use of signals inspired from ACh and DA for learning in a neural network model. This effort serves two aims: first, to shed light on the functional roles of ACh and DA in shaping cortical representations and, second, to provide inspiration for novel training methods for ANNs.

Previous studies examine the roles of ACh and DA in neural information processing. Weinberger and Bakin (1998) develop a model of ACh signaling to investigate its function in classical conditioning. Li and Cleland (2013) present a detailed biophysical model of ACh neuromodulation in the olfactory bulb. However, these studies do not see to the perceptual benefits of long-term plasticity induced by ACh. Other work tackle the question of DA-modulated plasticity in neural networks. Roelfsema and colleagues show that a signal inspired from DAergic signaling allows a network to learn various classification tasks (Roelfsema and Ooyen, 2005; Roelfsema et al., 2010; Rombouts et al., 2012). Similarly, other models make use of DA-like reinforcement signals to learn stimulus-response associations (e.g., Law and Gold, 2009; Liu et al., 2010). In these cases, however, the models for the plastic effects of DA were chosen to carry out reinforcement learning rather than to tally with experimental observations.

In contrast with previous work, we base our modeling effort on the well-documented observation that pairing ACh or DA release with a stimulus boosts neural responses to the stimulus. We use this model to study the perceptual benefits of ACh- and DA-induced plasticity under natural release conditions. In more details, we make use of a Hebbian-learning neural network and simulate the physiological release schedules of ACh and DA. In the model, ACh activation approximates attentional demand while DA activation arises from RPEs. We find that the neuromodulators trigger distinct changes in representations that both improve the network's classification performance. Specifically, ACh leads to changes in synaptic weights such that more neurons are dedicated to stimuli that are challenging for the network. DA adapts synaptic weights to the reward contingencies of a task, thereby sharpening neural tuning with respect to the classes of the task. These results provide a functional explanation for the roles of cholinergic and dopaminergic signals in refining cortical representations.

Our learning algorithm offers several advantages from a practical perspective. First, the network achieves performances comparable to those of state-of-the-art optimisation methods used to train multi-layer perceptrons (MLPs) while requiring weaker supervision signals. Second, learning takes place even in the absence of environmental feedback. And third, weight updates are based on synaptically-local information and on two signals broadcasted identically to all neurons. These features may make the algorithm interesting for

functional applications such as learning in neuromorphic processors.

2. METHODS

2.1. Hebbian Network Model

For our study, we make use of a Hebbian-learning neural network model introduced by Keck et al. (2012). The learning mechanisms implemented in this model achieve approximately optimal learning in terms of maximum likelihood estimation (see original publication for a detailed discussion). As a theoretically well-founded and biologically realistic model, this network is a natural starting point for our work. In this section, we briefly present the original model and then describe our simulation of the neuromodulators ACh and DA.

The network consists of three layers, an input, a representation, and a classification layer (Figure 1). Input values activate neurons in the first layer; activity then propagates through the network in the following steps.

2.1.1. Feedforward Inhibition

In mammals, the responses of sensory neurons are largely invariant to contrast in sensory stimuli (Sclar et al., 1990; Stopfer et al., 2003; Mante et al., 2005; Assisi et al., 2007; Olsen and Wilson, 2008), in part due to rapid feedforward inhibition (Pouille and Scanziani, 2001; Swadlow, 2003; Mittmann et al., 2005; Wehr and Zador, 2005; Pouille et al., 2009; Isaacson and Scanziani, 2011). To emulate this process, neural activations in the input layer are normalized:

$$y_d = (A - D) \frac{\tilde{y}_d}{\sum_{d'=1}^D \tilde{y}_{d'}} + 1, \quad (1)$$

where \tilde{y} are input data, A is a normalization constant, and D is the number of input neurons. This form of normalization yields contrast-invariant responses in representation neurons. For the dataset used in this work, $D = 28 \times 28 = 784$ input neurons.

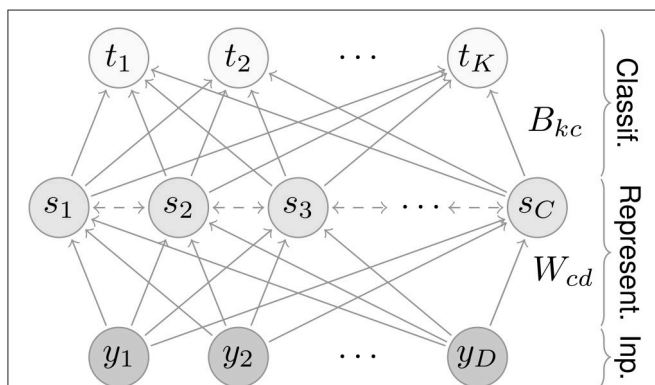


FIGURE 1 | Network architecture. The network contains three layers: an input, a representation, and a classification layer. For the MNIST dataset, the input and classification layers contain $D = 28 \times 28 = 784$ and $K = 10$ neurons, respectively. The number of representation neurons is variable; for most results we use $C = 7 \times 7 = 49$ neurons.

For other hyper-parameters, values are determined through grid search to maximize classification performance (see Table A1 in Appendix section).

2.1.2. Input Integration

Neurons in the representation layer integrate their input through a weighted sum:

$$I_c = \sum_{d=1}^D S(W_{cd})y_d, \quad (2)$$

where W is the weight matrix between the input and representation layers and $S(\cdot)$ is a linearised logarithm function given by:

$$S(W_{cd}) = \begin{cases} W_{cd} & \text{if } W_{cd} < 1 \\ \log(W_{cd}) + 1 & \text{if } W_{cd} \geq 1. \end{cases} \quad (3)$$

Taking the logarithm of W_{cd} guarantees approximate optimal learning of the weights, with the linearisation ensuring that the function is never negative for $W_{cd} \geq 0$.

2.1.3. Lateral Inhibition

The integrated input is fed through a softmax function that models global lateral inhibition:

$$s_c = \frac{\exp(I_c)}{\sum_{c'} \exp(I_{c'})}. \quad (4)$$

2.1.4. Hebbian Learning

Hebbian learning takes place between the input and representation neurons:

$$\Delta W_{cd} = \epsilon \cdot (s_c y_d - s_c W_{cd}), \quad (5)$$

where ϵ is the learning rate.

2.1.5. Classification

We subject the network to a classification task of images of handwritten digits from the MNIST dataset (LeCun et al., 1998b). These input images provide stimuli of intermediate complexity and high-dimensionality akin to natural sensory stimuli, making them a popular dataset to study neural information processing (Nessler et al., 2013; Schmuken et al., 2014). These data consist of gray-scale images with pixel values in the range $[0, 255]$ fed as input \tilde{y} to the first layer.

In the classification layer, we use statistical inference to decode activity in the representation layer. Given an input pattern \tilde{y} and the model parameters Θ , we want to infer the class of the input pattern, that is, to compute the posterior $\Pr(k | \tilde{y}, \Theta)$. Here, we approximate the posteriors using the labels of the input images. We first compute a value B_{kc} :

$$B_{kc} := \frac{1}{N_m} \sum_{n=1}^{N_m} \Pr(c | \tilde{y}^{(n)}, W) = \frac{1}{N_m} \sum_{n=1}^{N_m} s_c^{(n)}, \quad (6)$$

with N_m input patterns $\tilde{y}^{(n)}$ bearing a label $m = k$. The matrix B can be interpreted as the weights between the representation

and classification layers. This matrix is updated after every presentation of 100 images, or roughly 600 times during one iteration over the dataset. The posteriors are approximated as:

$$\Pr(k | \vec{y}, \Theta) \approx t_k = \frac{B_{kc} s_c}{\sum_{k'=1}^K B_{k'c}}. \quad (7)$$

As a classification result \hat{m} , we take the unit with the largest value of approximation to the posterior:

$$\hat{m} = \underset{k=1}{\operatorname{argmax}}^K(t_k). \quad (8)$$

This hierarchical formulation allows to decode activity in the representation layer, providing a probabilistic classification of the input images.

Previous work based on a fully probabilistic description of the Hebbian-learning network model (Forster et al., 2016; Forster and Lücke, 2017) shows that local Hebbian learning converges to the weight matrix B without requiring the non-local summation over k . This is true also when using a small fraction ($\approx 1\%$) of labeled training examples. Learning the classification weights can therefore be achieved while respecting biological constraints. For this work, we mainly focus on the standard fully labeled setting, as is customary (Keck et al., 2012; Nessler et al., 2013; Schmuker et al., 2014; Diehl and Cook, 2015; Neftci et al., 2015), but also provide results for experiments with very few labels.

2.2. Model of the Neuromodulators

2.2.1. Effects on Plasticity

We extend the network model described above to emulate the effects of ACh and DA on neural representations. Specifically, we simulate the impact of the neuromodulators as a modulation of the network's learning rate:

$$\text{acetylcholine:} \quad \Delta W_{cd} = \epsilon \cdot ACh \cdot (s_c y_d - s_c W_{cd}), \quad (5a)$$

$$\text{dopamine:} \quad \Delta W_{cd} = \epsilon \cdot DA \cdot (s_c y_d - s_c W_{cd}), \quad (5b)$$

where ACh and DA represent the activation of the corresponding neuromodulatory system. This model is in general agreement with experimental observations in that both ACh (Bröcher et al., 1992; Chun et al., 2013) and DA (Blond et al., 2002; Sun et al., 2005; Matsuda et al., 2006) are reported to promote synaptic plasticity. This model for the neuromodulators was chosen so as to reproduce the results of pairing experiments in mammals (see Results section).

2.2.2. Acetylcholine and Attentional Efforts

ACh release in the mammalian neocortex is tightly linked with attentional processes. For instance, as rats detect a behaviorally meaningful sensory cue, a spike in cortical ACh accompanies the reorientation of their attention towards the cue (Parikh et al., 2007). Additionally, when rats perform a task requiring sustained attention, the concentration of ACh in their prefrontal cortices more than doubles compared to control (Arnold et al., 2002; Kozak et al., 2006). In the course of such tasks, distractors that further tax the animals' attentional systems trigger supplemental

ACh release (Himmelheber et al., 2000; Kozak et al., 2006). These observations indicate that the cholinergic system responds to events demanding an animal's attention such as relevant stimuli or challenging tasks. In this sense, ACh transmission reflects the cognitive construct of attentional effort defined as a subject's motivated effort to maintain performance under challenging conditions (Sarter et al., 2006).

In the present work, we model ACh activation to approximate attentional demand. To quantify how demanding a stimulus is for the network, we use the network's classification confidence. Classification confidence is measured as the classifier's maximal posterior over the digit classes, $\kappa = \max_{k=1}^K(t_k)$. Classification confidence strongly correlates with classification accuracy ($r = 0.89$, **Figure 2A**) indicating that this measure is suitable to quantify stimulus demand. For each stimulus, the value of the ACh variable is given by:

$$ACh = \frac{\alpha}{1.0 + \exp(\beta \cdot (\bar{\kappa}_{\hat{m}}/\bar{\kappa} - 1.0))} \quad (10)$$

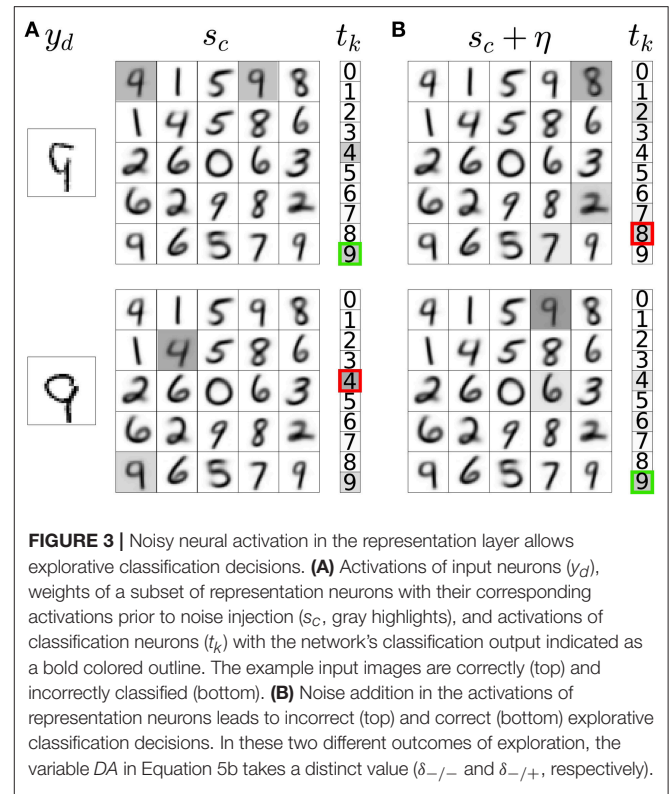
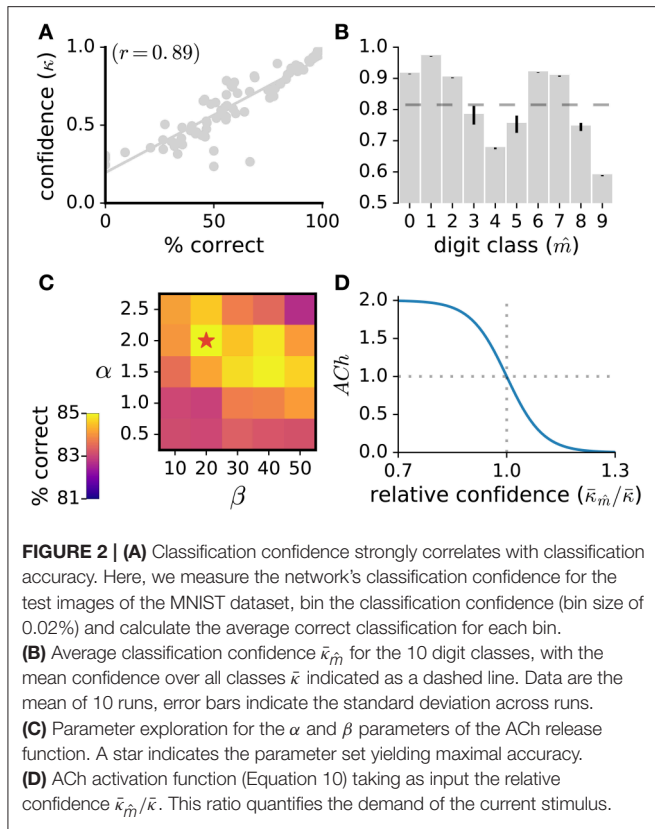
where $\bar{\kappa}_{\hat{m}}$ is the network's average classification confidence for the inferred class of the current stimulus, $\bar{\kappa}$ is the average classification confidence for all stimuli, and α and β are hyper-parameters of the sigmoid function whose values are determined through grid search (**Figure 2**). According to this formulation, the lower the classification confidence (i.e., the greater the stimulus difficulty), the larger the ACh activation. Note that, to compute the average classification confidence over the digit classes, we use the network's inferred classification (\hat{m}) and not the stimulus label. Thus, for a given stimulus, ACh activation is evaluated without requiring immediate environmental information. Also note that the classification confidence for the same stimulus may vary during training as the network's weight matrices W and B are updated.

2.2.3. Dopamine and Reward Prediction Errors

DA efflux in animals follows RPEs (Schultz et al., 1997; Satoh et al., 2003; Tobler et al., 2005; Schultz, 2010). We reproduce this release schedule in the model as follows. First, we allow explorative decision making by injecting additive noise in the activation of representation neurons (**Figure 3**):

$$I_c = \sum_{d=1}^D S(W_{cd}) y_d + \eta_c, \\ \eta_c \sim \mathcal{N}(0, \nu),$$

where \mathcal{N} is a normal distribution with zero mean and variance ν . This method for exploration approximates the softmax rule for action selection in reinforcement learning (Sutton and Barto, 1998). Following this rule, actions are selected stochastically with the probability of selecting an action proportional to its expected reward. The parameter ν corresponds to the temperature parameter of the softmax rule: for $\nu \rightarrow \infty$, all classification decisions have equal probabilities; for $\nu \rightarrow 0^+$, classification is purely exploitative. We find the optimal value for ν through grid search.



We then compute the classification output for each \vec{y} with and without the addition of noise η . If noise addition results in a classification decision that is different from the decision without noise addition, the classification is labeled as explorative; otherwise it is labeled as exploitative. If the network takes an exploitative decision it is said to predict a reward (+pred); if it takes an explorative decision it is said to not predict a reward (-pred). The network is rewarded for taking correct classification decisions (+rew) and not rewarded for incorrect decisions (-rew). The difference between the predicted and delivered rewards gives rise to a RPE. There are four possible RPE scenarios. In each of these cases, the DA variable in Equation 5b takes a distinct value:

$$DA = \begin{cases} \delta_{+/+} & \text{if } +pred \text{ and } +rew \\ \delta_{+/-} & \text{if } +pred \text{ and } -rew \\ \delta_{-/+} & \text{if } -pred \text{ and } +rew \\ \delta_{-/-} & \text{if } -pred \text{ and } -rew \end{cases} \quad (12)$$

where $\delta_{i,j}$ are constants whose values are determined through 4-dimensional parameter search to maximize classification performance.

2.2.4. Critical period

We are interested in changes in sensory representations triggered by neuromodulators in adult animals. Adult animals possess stable neural representations of their environment learned in early life during a brief window of heightened plasticity. During

this so-called critical period, the response properties of neurons rapidly adjust to the statistical structure of sensory stimuli (Sengpiel et al., 1999; de Villiers-Sidani et al., 2007; Han et al., 2007; Barkat et al., 2011).

As a model of this critical period, we pre-train the network solely through Hebbian learning (Equation 5). The network then learns synaptic weights based on correlations in the activation of input neurons, with weights that resemble the different digit classes. The weights in the representation layer are then learned solely through the statistics of the input images and do not reflect the task to be performed. As learning progresses, performance on the classification task increases and eventually saturates. Once performance reaches a plateau, we allow the release of ACh or DA. As an additional control condition, we also continue training the network through Hebbian learning. Omitting the pre-training results in the same functional performance but, without it, the optimal DA activation values found through parameter search differ (see Figure 6).

3. RESULTS

3.1. Pairing Experiment

In animals, coupling a stimulus with the release of either ACh (Kilgard and Merzenich, 1998a; Weinberger, 2003; Froemke et al., 2007, 2013) or DA (Bao et al., 2001; Frankó et al., 2010) triggers long-lasting changes in sensory representations. Specifically, sensory neurons increase their responses to the paired stimulus, resulting in more neurons preferring this

stimulus. To test whether our model of ACh and DA is in agreement with this observation, we perform a similar experiment. The experiment consists of coupling all stimuli of a target class with ACh or $DA = \rho$ in Equations 5a or 5b, where ρ is a constant > 1 (Figure 4A). Stimuli of all other classes have ACh and $DA = 1$. We then examine the distribution of class preferences in the network. The preferred digit class of a neuron is determined by taking $\text{argmax}_{k=1}^K(B_{kc})$ which gives the class to which neuron c maximally responds to. We find that the pairing protocol increases the responses of individual neurons to the paired stimulus class and augments the number of neurons preferring this class, in agreement with experimental data (Figures 4B–G). Furthermore, the procedure reduces the number of units tuned to classes close to the paired one (class closeness is measured as the Euclidean distance between the averages of all training examples of each class). These findings are in line with pairing experiments with DA showing that the cortical representations of frequencies neighboring a paired tone shrink as a result of the pairing procedure (Figure 4G; Bao et al., 2001). This observation however contrasts with pairing experiments with ACh which result in enlargements of the cortical representations of both the paired frequency and adjacent ones (Kilgard and Merzenich, 1998a). For this work, this difference in the effects of ACh and DA is not taken into account.

3.2. Physiological Release Schedule

3.2.1. Optimal Release Values

With our model in general agreement with the results of pairing experiments, we can now study the effects of the natural release schedules of ACh and DA. We first pre-train the network through Hebbian learning. As training progresses, performance saturates (Figure 5, inset). After this point, we allow the release of ACh or DA. We perform parameter search to identify the optimal values for parameters α and β in Equation 10 (Figure 2C) and for the $\delta_{i,j}$ constants in Equation 12 (Figure 6). In the case of the $\delta_{i,j}$ constants, we find that for surprising rewards ($-pred, +rew$) the optimal $\delta_{-/+}$ is positive while in the absence of an expected reward ($+pred, -rew$) the optimal $\delta_{+/-}$ is negative. For correctly predicted rewards (either $+pred, +rew$ or $-pred, -rew$) the optimal $\delta_{+/+}$ and $\delta_{-/-}$ are close to zero. This optimal activation profile matches that observed in primates (Schultz et al., 1997; Tobler et al., 2005), Figures 6B–C).

3.2.2. Effects of ACh

Visual inspection of the weights of the network (Figure 7A) indicates that ACh alters the number of neurons dedicated to the different digit classes. For instance, there are more neurons resembling a “4” and fewer neurons resembling a “1” after training with ACh. We quantify this redistribution by determining the preferred class of a representation neuron. For Hebbian learning, the distribution of preferred classes is close to uniform but not entirely so (Figure 7B). There is a positive correlation between the number of neurons dedicated to a class and the network’s performance on this class ($r = 0.22$,

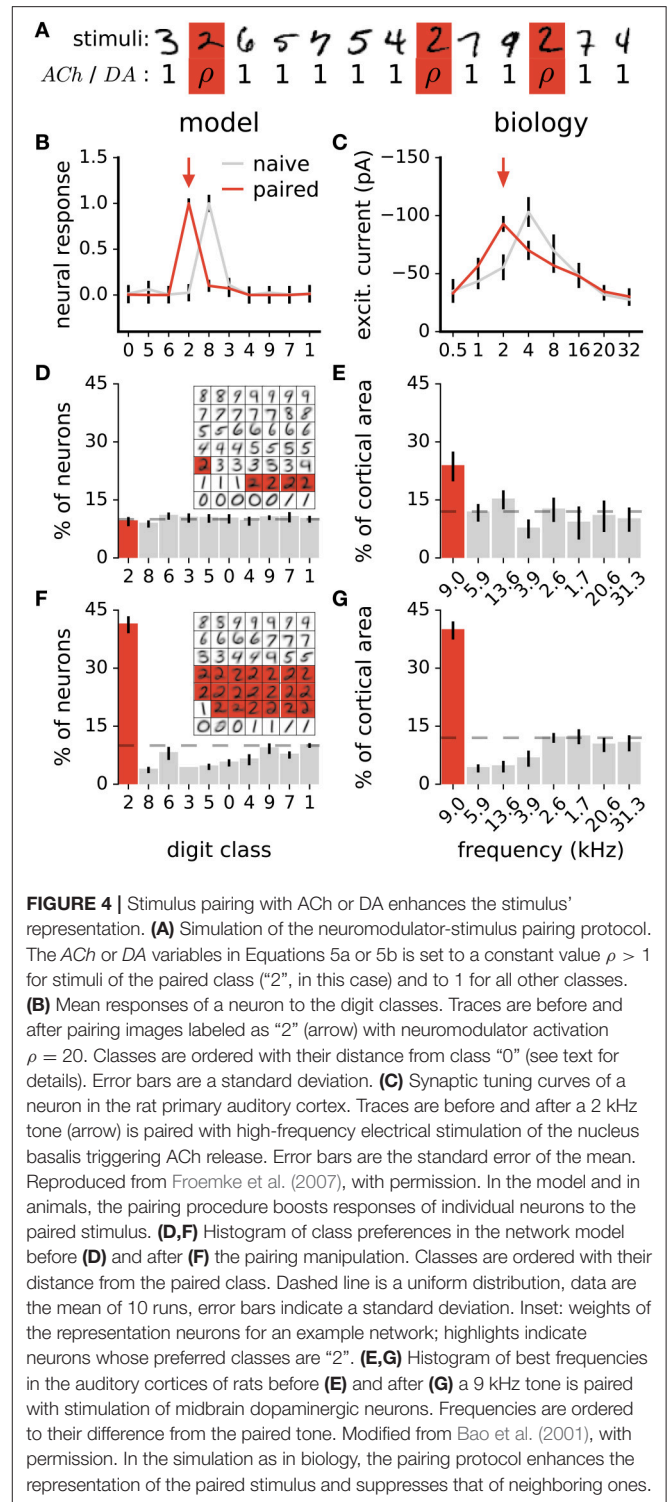
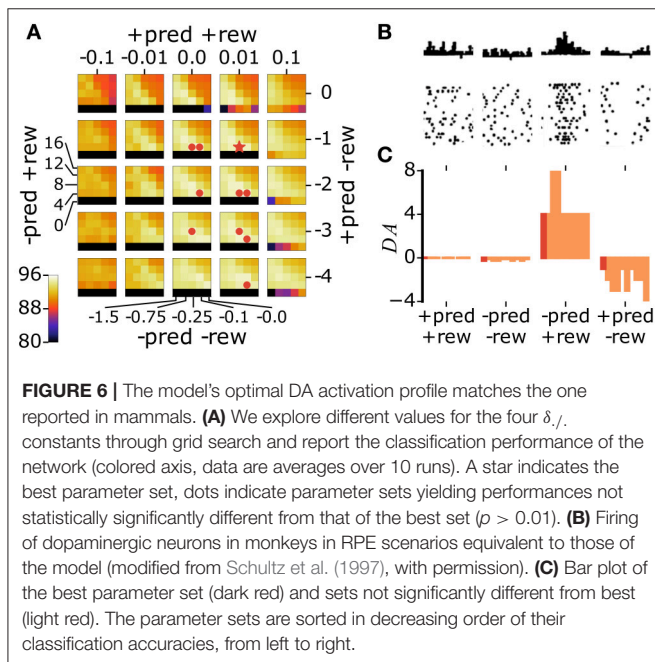
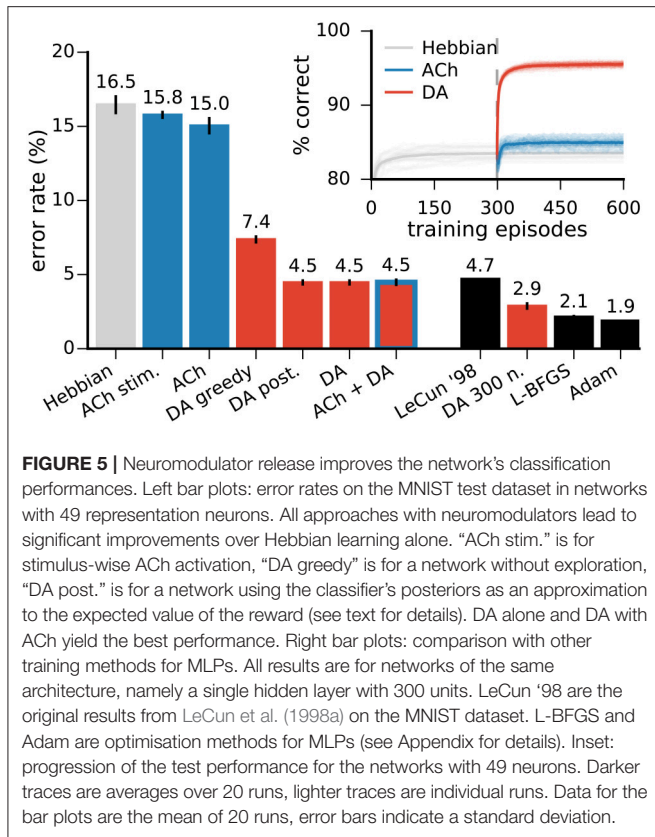


FIGURE 4 | Stimulus pairing with ACh or DA enhances the stimulus' representation. **(A)** Simulation of the neuromodulator-stimulus pairing protocol. The ACh or DA variables in Equations 5a or 5b is set to a constant value $\rho > 1$ for stimuli of the paired class (“2”, in this case) and to 1 for all other classes. **(B)** Mean responses of a neuron to the digit classes. Traces are before and after pairing images labeled as “2” (arrow) with neuromodulator activation $\rho = 20$. Classes are ordered with their distance from class “0” (see text for details). Error bars are a standard deviation. **(C)** Synaptic tuning curves of a neuron in the rat primary auditory cortex. Traces are before and after a 2 kHz tone (arrow) is paired with high-frequency electrical stimulation of the nucleus basalis triggering ACh release. Error bars are the standard error of the mean. Reproduced from Froemke et al. (2007), with permission. In the model and in animals, the pairing procedure boosts responses of individual neurons to the paired stimulus. **(D, F)** Histogram of class preferences in the network model before **(D)** and after **(F)** the pairing manipulation. Classes are ordered with their distance from the paired class. Dashed line is a uniform distribution, data are the mean of 10 runs, error bars indicate a standard deviation. Inset: weights of the representation neurons for an example network; highlights indicate neurons whose preferred classes are “2”. **(E, G)** Histogram of best frequencies in the auditory cortices of rats before **(E)** and after **(G)** a 9 kHz tone is paired with stimulation of midbrain dopaminergic neurons. Frequencies are ordered to their difference from the paired tone. Modified from Bao et al. (2001), with permission. In the simulation as in biology, the pairing protocol enhances the representation of the paired stimulus and suppresses that of neighboring ones.

Figure 7D), suggesting that representing a class with more neurons is beneficial to performance.

Training with ACh redistributes class preferences in the network, leading to a less uniform distribution. Specifically, ACh increases the number of neurons dedicated to challenging classes while easier classes are represented with fewer units.



Consider for example the classes "1" and "4," the stimuli on which the network performs best and worst, respectively (Figure 7C, top row). ACh release leads to a respective decrease and increase in the number of neurons preferring these classes

(Figure 7B). The redistribution of neurons elicited by ACh raises the network's accuracy on the difficult classes (e.g., "4") and lowers performance on the easy classes (e.g., "1," Figure 7C, middle row). ACh thus reverses the correlation between neuron count and performance ($r = -0.79$, Figure 7D). On average over all classes, performance rises from $83.5 \pm 0.7\%$ with Hebb's rule alone to $85.0 \pm 0.6\%$ when supplemented with ACh, corresponding to a relative decrease of 12% in the error rate.

In addition to ACh activation computed as an average over the classes \hat{m} , we experiment with stimulus-wise ACh activation. Here, the value of the ACh variable is determined for each individual stimulus based on the classifier's posterior for this stimulus (specifically, we use the term κ instead $\bar{\kappa}_{\hat{m}}$ in Equation 10). Although this approach also improves performance, the gains in accuracy are of smaller magnitude than if ACh activation is computed as an average over the classes (Figure 5, "ACh stim."). We explain this outcome as the learning mechanism attributing a too great representational importance to demanding but detrimental data, for instance miss-labeled or outlier data points.

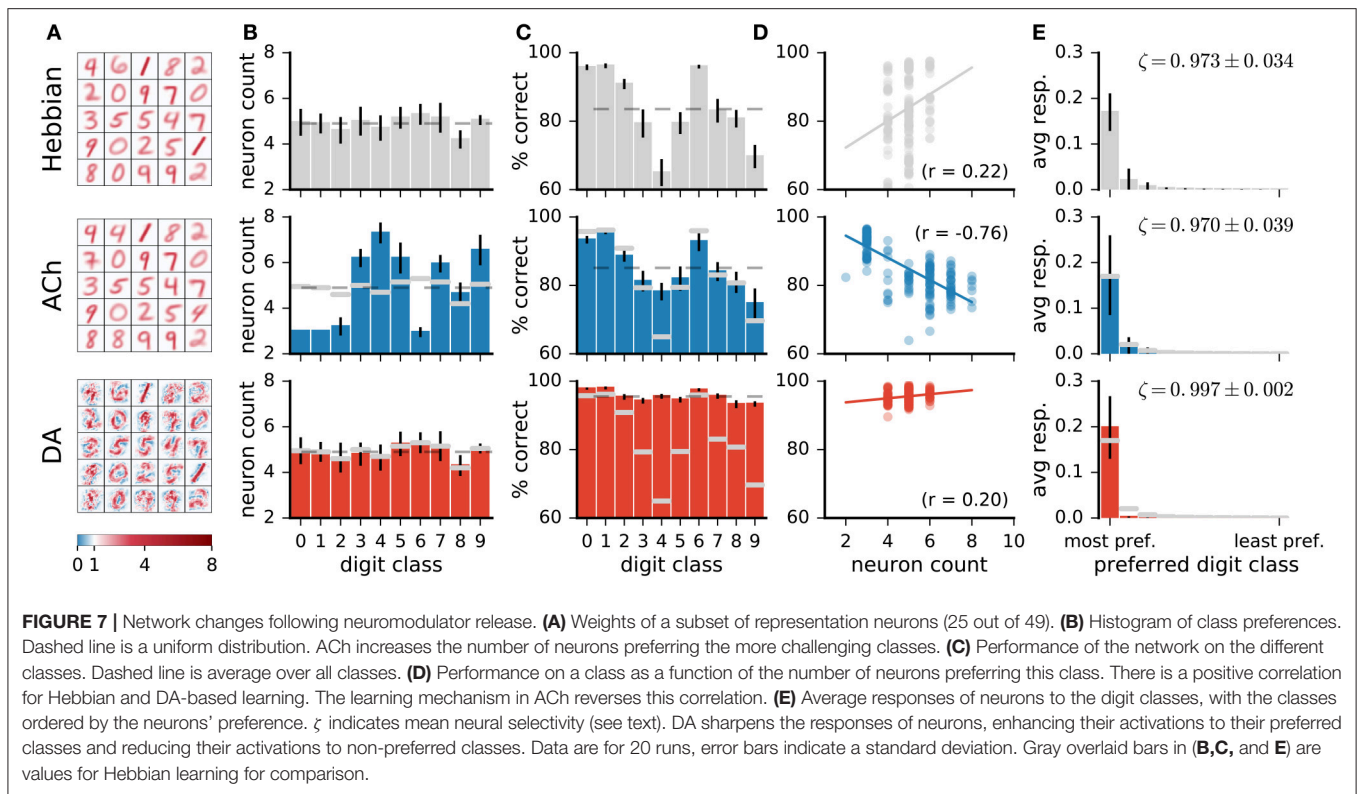
3.2.3. Effects of DA

In contrast with ACh signalling, DA bears little effect on the number of neurons responsive to the different classes (Figure 7B). For both Hebbian and DA-based learning, the distribution of the neurons' preferred digit class is close to uniform. The positive correlation between neuron count and classification performance also remains after training with DA (Hebbian: $r = 0.22$, DA: $r = 0.20$).

Visual inspection of the weights suggests that DA makes neurons' weights more selective to specific digit classes. Consider the example weights shown in Figure 8A. Weights in one column are for corresponding neurons in a Hebbian and DA network (the networks were initialised with the same random seed). Weights in the Hebbian model are rather poorly tuned to the digit classes (e.g., the neuron resembling a "3," "5," and "8" in the second column of Figure 8A). On the other hand, DA-based learning leads to weights that more closely correspond to specific digits. This observation can be quantified by measuring the average responses of neurons to the different classes (first and third rows in Figure 8A). The measure shown indicates that Hebbian learning yields neurons exhibiting strong responses to multiple stimulus classes, i.e., with a broad tuning. Training with DA yields more sharply tuned weights as units respond almost exclusively to a single digit category.

On average over all neurons, DA generates a 17% increase in neurons' activations to their preferred classes, accompanied by a 84% reduction to non-preferred classes (Figure 7E). These modifications amount to neuron weights being more selective to specific digits, or having a sharper tuning. We quantify such neural selectivity as the difference between a neuron's mean response to stimuli of its preferred class and its mean response to stimuli of all other classes:

$$\zeta_c = \frac{\bar{s}_c^\bullet - \bar{s}_c^\circ}{\bar{s}_c^\bullet}, \quad (13)$$



where \bar{s}_c^* and \bar{s}_c° are the average responses of neuron c to stimuli of its preferred and non-preferred classes, respectively. Here, $\zeta_c = [0, 1]$, where $\zeta_c = 0$ is a neuron that responds equally strongly to all stimuli and $\zeta_c = 1$ is a neuron that responds exclusively to one digit category. Selectivities of individual neurons are indicated on **Figure 8A**; selectivities averaged over all neurons of a network, ζ , are indicated on **Figure 7E**. We can also quantify a neural network's selectivity for a specific digit class m as the sum of the selectivity of the neurons whose preferred stimulus class is m , ζ_m (see **Figure 8-B,C**). Training with DA statistically significantly boosts neural selectivity ($p < 0.001$).

DA induces large improvements in classification accuracy ($95.53 \pm 0.05\%$ for DA compared to $83.5 \pm 0.7\%$ for Hebbian learning, $p < 0.0001$), corresponding to a 72.7% reduction in the error rate. Performance for a class strongly correlates with neural selectivity for this class, for both the Hebbian and DA networks ($r = 0.996$ and $r = 0.920$, respectively, **Figures 8B,C**). These strong correlations suggest that enhanced neural selectivity explains the rise in correct responses following training with DA.

We can further visualize the outcome of DA learning by reducing the dimensionality of input images to 2 features (using t-SNE, Maaten and Hinton, 2008) and train the network on these data (**Figure 9**). In Hebbian learning, the neural network acts as a clustering algorithm and, as the learning mechanism is agnostic to the labels of the stimuli, the classification boundaries miss some aspects of the data classes. In particular, boundaries are poorly defined between close-by clusters such as “3,” “5,” and “8.” Following DA signalling, weights adjust to

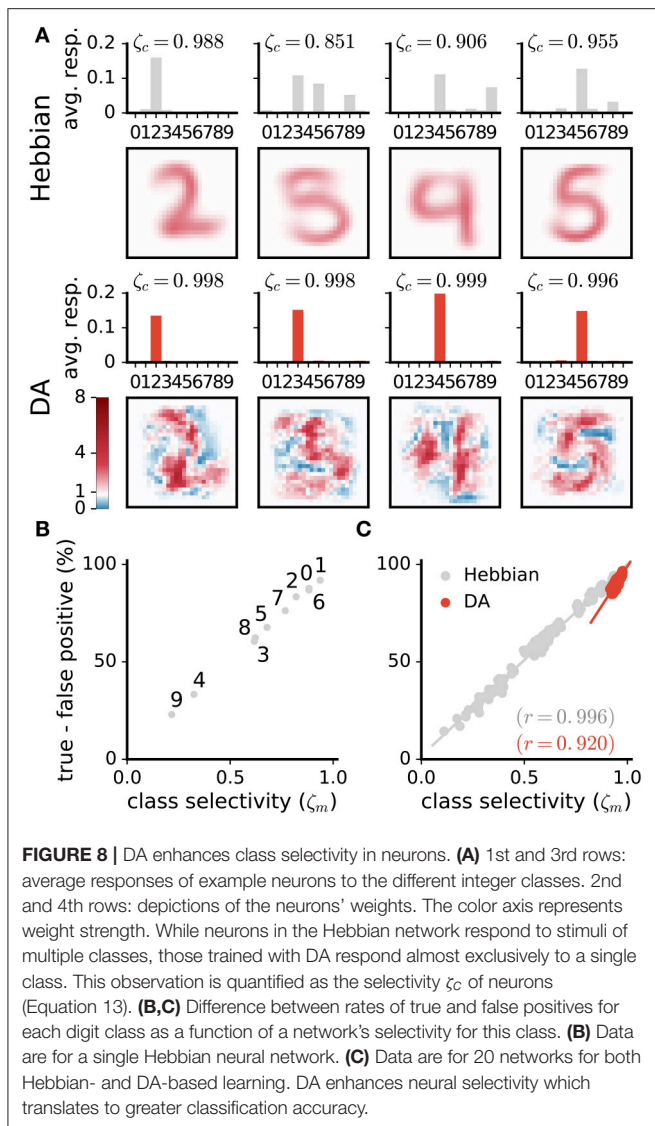
match the boundaries for the conditions for reward delivery of the task.

In the model for DA activation presented above, reward predictions are binary, reflecting solely whether a decision is explorative or not. An alternative approach is to use the classifier's posterior for the output class (i.e., its classification confidence) as an approximation to the expected value of the predicted reward. This posterior probability strongly correlates with the empirically-measured reward probability ($r = 0.98$), validating the approximation. However, we find that this approach does not improve the network's accuracy over binary reward predictions (**Figure 5**, “DA post”).

In order to assess the role of exploration in DA-based learning, we train a network without allowing explorative decision making. This greedy network achieves a classification score of $92.51 \pm 0.07\%$ (**Figure 5**, “DA greedy”), compared with $95.53 \pm 0.05\%$ with exploration. Exploration thus accounts for a further 18% relative drop in the error rate.

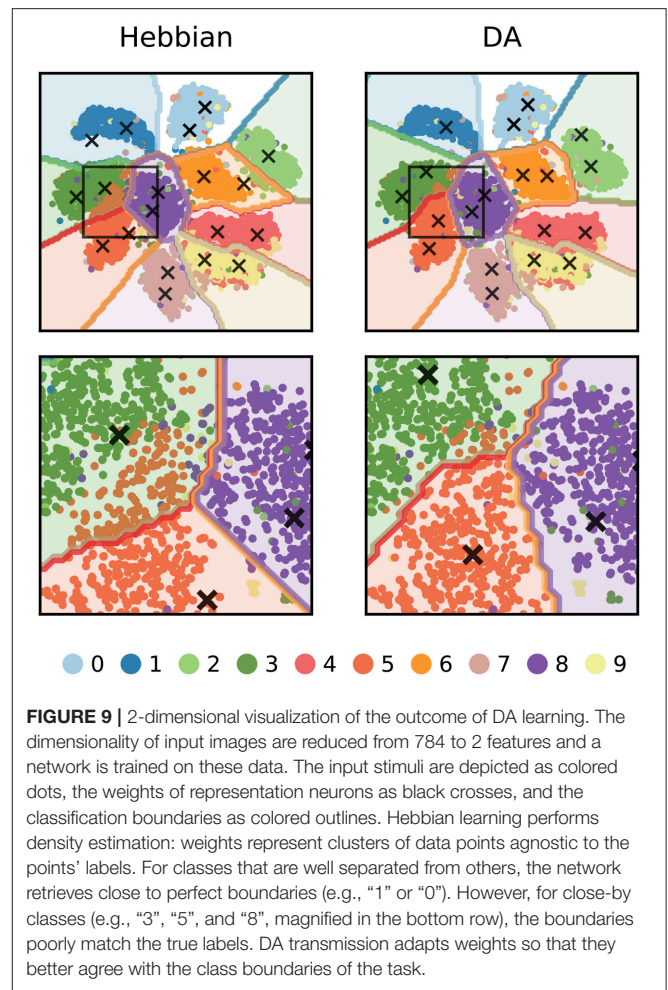
3.2.4. Learning on Non-uniformly Distributed Data

For the results on the MNIST dataset, ACh yields modest reductions in error rates relative to DA. This less important effect may be explained in part by the almost even distribution of training examples over the classes in the dataset. In more natural settings, some classes may contain many more examples than others while a high classification performance is equally important on all classes. For instance, a gatherer may see many more examples of “green leaves” than “berries” but still requires a low error rate for both classes. We test the



impact of ACh in a modified version of MNIST in which a subset of the classes are over-represented. Here, the training dataset contains the classes “0,” “2,” “3,” “5,” and “8,” and there are 60 times more “0” and “2” (the “leaves”) than the other classes (the “berries”). To model equal importance of the classes, we take the test dataset to be uniformly distributed over the classes. For Hebbian learning, the network performs poorly on the under-represented classes as it dedicates only few neurons to these classes (Figure 10, top row). Neuromodulation significantly improves accuracy and, on these data, ACh yields gains comparable in size to those of DA. As with the standard MNIST dataset, ACh carries its effect by attributing more neurons to classes on which performance is low (those that are under-represented). DA only has minimal effects on the distribution of class preference; increases in performance derive from boosting neural selectivity.

In addition to training the network with ACh and DA separately, we combine the two neuromodulators by allowing

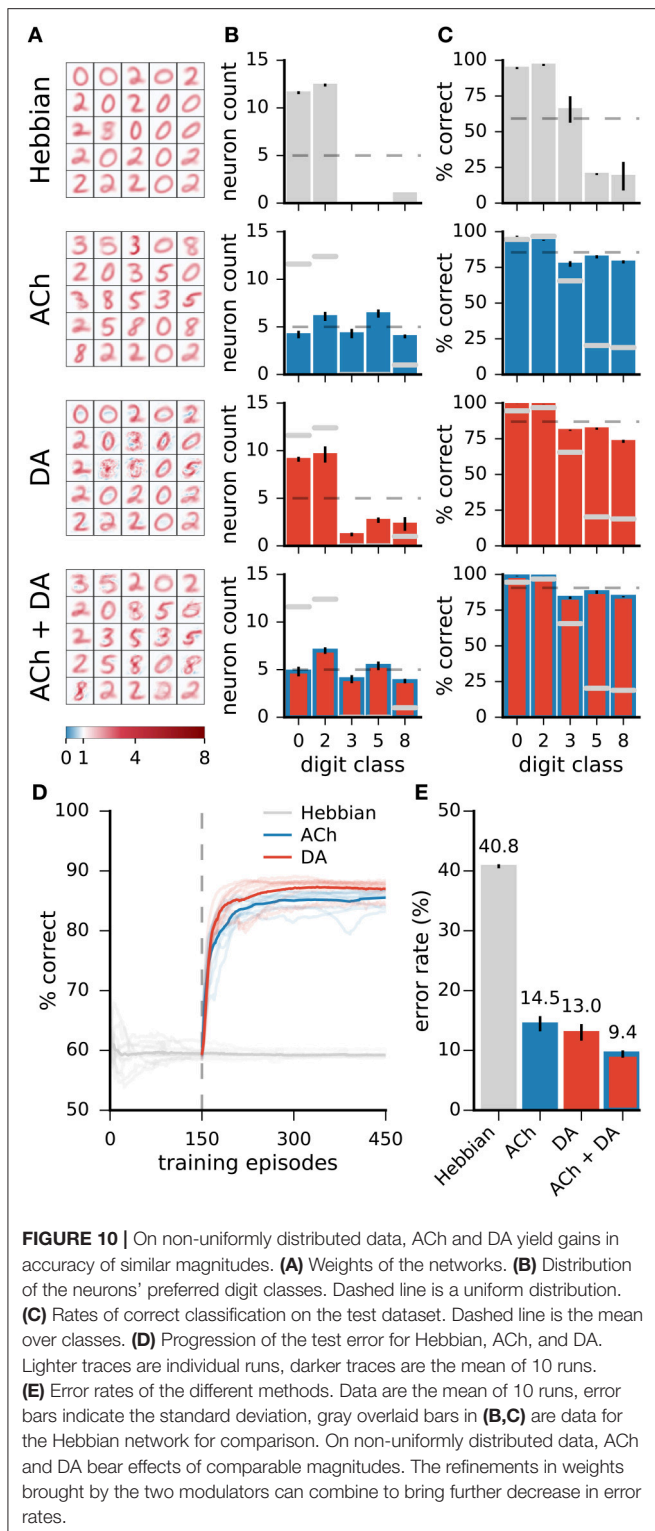


first ACh release and then DA. This procedure leads to a redistribution of the class preferences (due to ACh) followed by an enhancement in neural selectivity (due to DA). The combined activations of ACh and DA result in a further decrease in error rates compared to either modulator alone, indicating that the effects of ACh and DA can successfully combine (Figure 10).

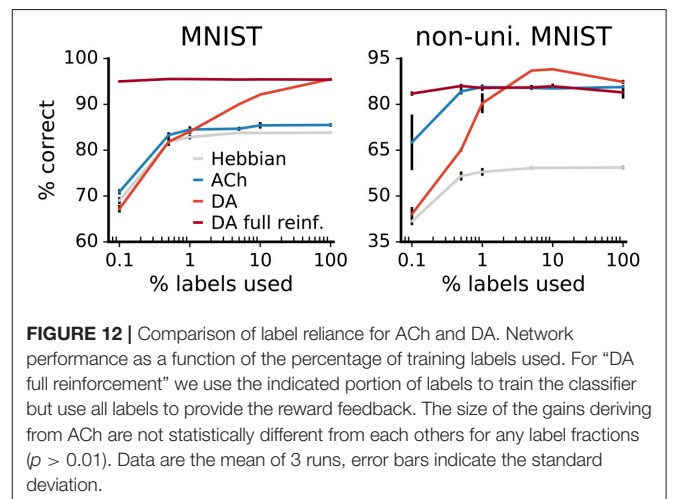
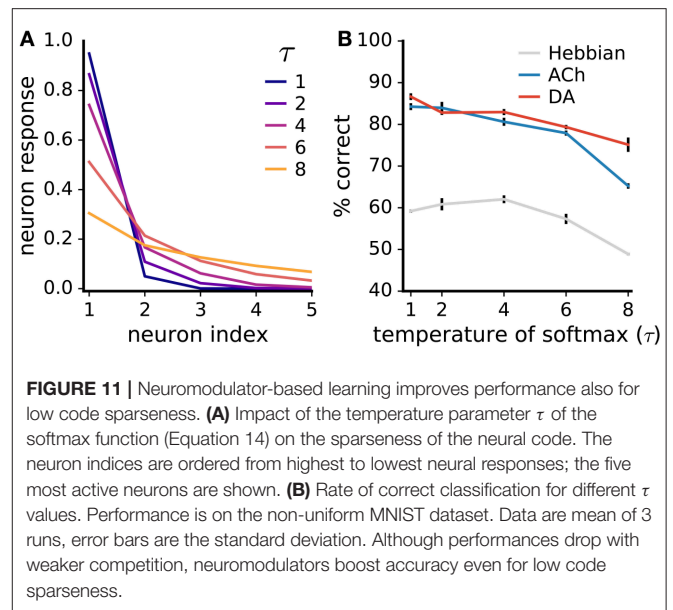
3.2.5. Impact of Code Sparseness

Lateral inhibition sparsifies the network's neural code so that inputs activate only one or a few neurons at a time (Figure 11A). Such a strong sparse code facilitates learning with neuromodulators as it avoids the credit-assignment problem. Additionally, the global neuromodulator signals are then essentially computed for a single neuron at a time. To examine the extent of the impact of the code's sparseness on learning, we introduce a temperature parameter τ to the softmax function determining the strength of the lateral competition:

$$s_c = \frac{\exp(I_c/\tau)}{\sum_{c'} \exp(I_{c'}/\tau)}. \tag{14}$$



For $\tau \rightarrow 0^+$, the softmax function gives rise to a winner-take-all competition with a single active neuron; for $\tau \rightarrow \infty$, neural responses are uniformly distributed. We train networks with different τ values on the non-uniform MNIST dataset (we use the non-uniform dataset to better discern the effects on



ACh-based learning). We find that the networks' performance drops as code sparseness decreases (**Figure 11B**). However, the neuromodulators give rise to large and statistically significant improvements even for low code sparseness, indicating that strong competition is not required for effective neuromodulator-based learning.

3.2.6. Impact of Label Availability

We examine the impact of label availability on learning by training networks with a varying fraction of labels, from 100% down to 0.1%. The accuracies of the networks decrease with label scarcity, both for learning with Hebb's rule and with neuromodulators (**Figure 12**). For the Hebbian network, labels only affect the classification layer; the decay in performance therefore derives exclusively from lower classifier accuracy.

For the neuromodulators, while label scarcity affects them both, the consequences are more substantial for DA. In particular,

when less than 1% of labels is used, the benefits of DA drop below those of ACh, this for both versions of the MNIST dataset. In error-based learning, labels are necessary to determine the correctness of an output. Reducing the ratio of labeled data consequently substantially hinders DA learning. On the other hand, the ACh signal yields gains in performance that are not statistically significantly different for all label fractions ($p > 0.01$). The constant improvements over declining label availability suggest that ACh learning relies effectively only minimally on labels, making ACh signaling beneficial even for scarcely labeled data.

DA-based learning does not require labels per se but only indications of whether outputs are right or wrong. We train an additional network using a fraction of the labels for the classifier but all labels for the reward feedback. The results show that performance remains high even for small label fractions, indicating that DA performs well in scenarios where true labels are in short supply but reinforcement feedback is available.

3.2.7. Performance Benchmark

In order to benchmark the functional performance of our algorithm, we compare it to MLPs trained with error back-propagation. We use the same architecture for our network and the MLPs (in this case, 784 input, 300 hidden, and 10 output neurons) and report the test error rate on the MNIST data. We train the MLPs using two state-of-the-art optimisation methods, the L-BFGS (Zhu et al., 1997) and Adam algorithms (Kingma and Ba, 2014) (see Appendix). In the original publication of benchmark results on the MNIST dataset, LeCun et al. (1998a) report a test error of 4.7% for an MLP of the architecture described above. Our biology-inspired algorithm yields a mean error rate of $2.88 \pm 0.05\%$, outperforming this original result. The MLPs with the L-BFGS and Adam optimisers yield an error rate of $2.15 \pm 0.04\%$ and $1.88 \pm 0.02\%$, respectively (Figure 5). In comparison, spiking neural networks intended for neuromorphic systems reach error rates of 5.0% (6,400 hidden spiking neurons, Diehl and Cook, 2015) and 4.4% (500 hidden spiking neurons, Neftci et al., 2015).

4. DISCUSSION

4.1. Learning Mechanisms

We study the effects of two modulatory signals on the representation and classification performance of a neural network. In our model, both signals act identically on synaptic plasticity but follow different release schedules, putatively those of ACh and DA. We find that these two signals give rise to distinct modifications in neural representations that both improve classification performance. Our model allows us to formulate hypotheses regarding the functional roles of ACh and DA in cortical representation learning. These roles can be explained as follows.

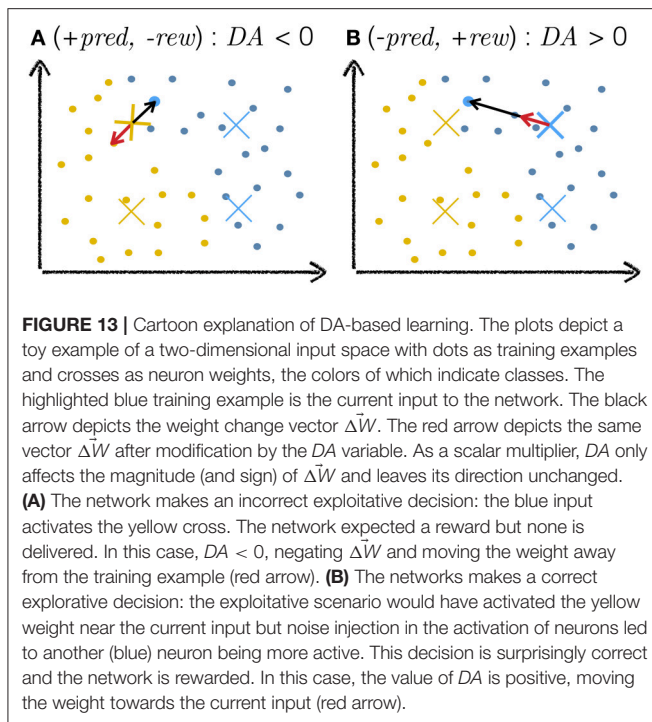
Consider the input $\vec{y}^{(n)}$ and the weights \vec{W}_c as vectors in a high-dimensional space. The activation of a neuron s_c is computed as the dot product between an input and the weight vectors. Lateral inhibition introduces a soft winner-take-all competition resulting in a few neurons having strong responses

and other neurons being silent. Hebbian learning then induces weight modifications $\Delta W_c = \epsilon \cdot s_c(\vec{y} - \vec{W}_c)$ (Equation 5). We note that, for each weight, ΔW_c points from the weight towards the current input. Both the variables ACh and DA modulate the magnitude of ΔW_c , $\|\Delta \vec{W}_c\|$ (Equations 5a and 5b).

Hebbian learning in the network performs density estimation: the distribution of the weights is determined by the density of data points in the input space. Modulating the learning rate of the network is similar to modifying data point density in that presenting a training image twice is comparable to presenting this image once but with a twice larger learning rate. For ACh-based learning, input images that are more challenging will trigger greater ACh activation, or have a larger learning rate. A cluster of data points associated with greater ACh activation is thus similar to having more data points in this cluster, inducing more neurons to represent the cluster. Or in other words, data points with $ACh > 1$ will have $\|\Delta \vec{W}_c\|$ of a greater magnitude, thereby exerting an increased “pull” on the weights.

For DA-based learning, the variable DA takes a value $\delta_{i,j}$ specified by the current RPE scenario. According to the parameter search, for correct reward predictions ($+pred, +rew$ or $-pred, -rew$), the optimal $\delta_{+/+}$ and $\delta_{-/-}$ are of approximately zero. In both cases, $\|\Delta \vec{W}_c\| \approx 0$; all the network’s weights remain unchanged. When the network takes an exploitative decision that turns out to be wrong ($+pred, -rew$), the optimal $\delta_{+/-}$ is inferior to zero. The vector $\Delta \vec{W}_c$ is negated so that it points away from the current input (Figure 13A). Active neurons will have their weights move away from the current input and are then less likely to win the softmax competition at future presentations of this input. When the network takes an explorative decision that is surprisingly correct ($-pred, +rew$), the optimal $\delta_{-/+}$ is positive. The weights of active neurons move towards the input (Figure 13B). The explorative decision (expected incorrect) turned out to be right; this decision should be taken again on future presentation of the same stimulus. DA-based learning can be understood as reinforcement learning at the level of sensory representations.

These learning mechanisms are related to several known machine learning algorithms. In the purely Hebbian case, the network is akin to a Kohonen map (Kohonen, 1982) in that learning proceeds iteratively through neural competition and weight adaptation (without however the cooperation aspect which confers the topological organization to Kohonen maps). The ACh learning mechanism is reminiscent of boosting methods, for instance AdaBoost (Freund et al., 1999), which attribute greater weights to misclassified training examples. The DA learning mechanism is closely related to algorithms such as REINFORCE (Williams, 1992) which make use of a reinforcement signal acting on the learning rate of a neural network’s weight update rule. It is interesting to note that, despite this close correspondence, the decision to model DA as a modulation of the network’s learning rate was made not to match those rules but rather to mirror biology. Indeed, our model of DA (and ACh) emulates the observation that stimuli coinciding with release of the neuromodulators are over-represented in animal sensory cortices (Figure 4). The close similarity between our model of DA and REINFORCE’s learning rule can thus



be taken as further support for the biological realism of the latter.

4.2. Acetylcholine

Activation of the cholinergic system in mammals appears to follow attentional efforts. Sarter et al. (2006) review evidence suggesting that deteriorating performances, as indicated by a rise in error rates and a decline in reward rates, trigger effortful cognitive control to prevent erroneous behavior. Attentional efforts are paralleled by a heightened activation of cholinergic neurons in the basal forebrain (Himmelheber et al., 2000; Passetti et al., 2000; Dalley et al., 2001; Arnold et al., 2002; McGaughy et al., 2002; Kozak et al., 2006) which in turn broadcast this signal to the cortical mantle (Hasselmo and Sarter, 2011). For instance, engaging in a demanding motor (Conner et al., 2010) or tactile (Butt et al., 1997) task enhances ACh release in the motor and somatosensory cortices, respectively.

There is broad evidence that ACh acts as a permissive plasticity agent at its projection sites (Buchanan et al., 2010; Giessel and Sabatini, 2010), for instance promoting alterations of neural representations in sensory cortices (Greuel et al., 1988; Bröcher et al., 1992; Kilgard and Merzenich, 1998a,b; Ji et al., 2001; Ma and Suga, 2005; Suga, 2012; Chun et al., 2013). The scientific literature contains several hypotheses regarding the functional role of the modifications elicited by ACh. Froemke et al. (2007) suggest that shifts in neural tunings toward a stimulus paired with ACh activation serves as a long-term enhancement of attention to this stimulus. Others postulate that this modification stores the behavioral relevance of the stimulus (Kilgard and Merzenich, 1998a; Weinberger, 2003) or generally improves signal processing (Gu, 2003; Froemke et al., 2013).

Here, we show that a signal modulating synaptic plasticity as a function of task difficulty improves the quality of a neural representation with respect to a classification task. The gains in performance result from assigning more neurons to challenging stimulus classes. Our model suggests that ACh serves this role in mammalian cortices.

Experimental evidence offer support for this hypothesis. For instance, motor skill acquisition and the accompanying enlargement of relevant representations in the motor cortex require ACh activation (Conner et al., 2003, 2010). Conversely, discrimination abilities rise for a tone whose representation is expanded as a result of repeated pairing with ACh activation (Reed et al., 2011). More generally, ACh antagonists or lesion of the cholinergic system impairs perceptual (Butt and Hodge, 1995; Fletcher and Wilson, 2002; Wilson et al., 2004; Leach et al., 2013) and motor skill learning (Conner et al., 2003). These results indicate that the cholinergic system is crucial for forms of learning involving modifications in sensory maps, especially those affecting the relative extent of cortical representations, as suggested in this work.

Our model of ACh is in line with a previous simulation study by Weinberger and Bakin (1998). The authors make use of a modified version of Hebb's rule and simulate the action of ACh as an amplification in the post-synaptic activation of target neurons. An *in vivo* micro-stimulation study validates this model. For the Hebbian rule used in this work, the two models of ACh are mathematically equivalent; this previous work thus offers support to the simulation employed here.

4.3. Dopamine

Dopaminergic neurons of the midbrain encode various features of rewards (Satoh et al., 2003; Tobler et al., 2005) and, in particular, strongly respond to the difference between predicted and received rewards (Schultz et al., 1997; Schultz, 2010). Midbrain neurons project to the entire cortex (Haber and Knutson, 2010) and the reward signals they carry modulate neural activity in most cortical areas (Vickery et al., 2011) including primary sensory cortices (Pleger et al., 2009; Brosch et al., 2011; Arsenault et al., 2013).

DA affects plasticity at the sites where it is released, as measured both at the level of synapses (Otani et al., 1998; Centonze et al., 1999; Blond et al., 2002; Bissière et al., 2003; Li et al., 2003; Sun et al., 2005; Matsuda et al., 2006; Calabresi et al., 2007; Navakkode et al., 2007) and behaviorally (Brembs et al., 2002; Wise, 2004; Graybiel, 2005; Kudoh and Shibuki, 2006; Klein et al., 2007; Luft and Schwarz, 2009; Molina-Luna et al., 2009; Hosp et al., 2011; Schicknick et al., 2012; Ott et al., 2014). In sensory cortices, DA efflux, triggered either by electric stimulation of the midbrain or by reward delivery, elicits plastic changes in the responses of primary sensory neurons (Bao et al., 2001, 2003; Beitel et al., 2003; Frankó et al., 2010; Poort et al., 2015).

The role of the plastic modifications induced by DA are usually understood in terms of reinforcement learning, for instance to learn the appetitive value of stimuli (Brembs et al., 2002; Wise, 2004; Frankó et al., 2010) or to learn reward-directed behaviors (Watkins and Dayan, 1992; Dayan and Balleine,

2002; Wise, 2004; Schicknick et al., 2012; Ott et al., 2014). In sensory representations, the changes brought forth by DA were previously hypothesized to enhance the saliency of stimuli predictive of rewards (Bao et al., 2001) and to adapt cortical representations to task requirements (Brosch et al., 2011).

Here, we show that a signal modulating plasticity as a function of RPEs adapts synaptic weights to the reward contingencies of a task, thereby improving performance on the task. Specifically, in our model, the responses of neurons become matched to the boundaries in conditions for reward delivery. In the digit classification task, this results in neurons being better tuned to the distinct digit classes, in this way improving classification performance. We suggest that, in mammals, dopamine carries this role of adapting sensory representations to the reward contingencies of a task.

After training monkeys on a visual discrimination task, neural responses become matched to the stimulus features that discriminate between the reward conditions of the task (Sigala and Logothetis, 2002). This process is comparable to the effect of DA in our model. We thus postulate that DA orchestrates these changes and predict that lesioning the dopaminergic system would prevent this form of learning. Animal experiments show that interfering with DA signaling impairs sensory discrimination learning (Kudoh and Shibuki, 2006; Schicknick et al., 2012), supporting this prediction.

The optimal values of the δ_j constants we find through parameter exploration are in close qualitative agreement with the release properties of DA observed in primates (Schultz et al., 1997; Tobler et al., 2005) (Figure 6). Both in animals and in the present model, unpredicted rewards lead to a rise in dopaminergic activation while the absence of predicted rewards lead to a reduction in activation. Correctly predicted rewards leave dopaminergic activation essentially unchanged. The release values in the model were selected to maximize performance on a discrimination task. It is conceivable that the dopaminergic activation schedule in animals was similarly selected through evolutionary pressures to maximize perceptual abilities.

We tested the effect of explorative decision-making while training with DA and found that exploration yields an additional relative reduction of 18% in error rates. Studies show that human subjects actively engage in exploratory behavior when making decisions (Daw et al., 2006). Explorative decision-making is usually understood as a method to sample available choices with the prospect of discovering an option richer than the current optimum. Our model suggests that, in perceptual decision making, such explorative behavior may additionally serve the purpose of refining cortical sensory representations.

4.4. Comparing Acetylcholine and Dopamine

On the non-uniform dataset, ACh gives rise to improvements comparable in size to those of DA. This result highlights the relevance of ACh in scenarios where training examples are largely non-uniformly distributed over the classes, as is often the case in natural conditions. Furthermore, in contrast to DA, the ACh signal yields gains in accuracy of constant magnitude over

decreasing label availability. This finding points to a particularly beneficial role for ACh when environmental feedback is scarce.

On the non-uniform dataset, the combined effects of the two neuromodulators are greater than either one separately. This result indicates that the weight modifications brought by ACh and DA are distinct and complimentary, and that they can successfully combine.

4.5. Functional Performances and Outlook

The learning mechanisms presented in this work yield error rates close to that of state-of-the-art optimisation methods used to train MLPs for comparable network architectures. Since evolutionary pressures must have favored well performing learning mechanisms in the brain, any candidate model of cortical learning must offer strong functional performances. Our model meets this criteria, making it a suitable model for learning in biological neural structures.

In line with recent studies of biologically-plausible learning (Keck et al., 2012; Nessler et al., 2013; Schmuker et al., 2014; Diehl and Cook, 2015; Neftci et al., 2015), we used correct classification as a measure of performance. This measure facilitates the study of the functional roles of neuromodulators and the comparison with previous work. Our neuromodulator-based learning method can be extended to tasks beyond classification, for instance by generalizing the softmax competition to k-winner-take-all (O'Reilly, 2001) or soft-k-winner-take-all (Lücke, 2009) competition.

Even in the sole context of classification, however, our approach offers several interesting advantages. For instance, compared to the traditional approach of gradient descent on a classification error, neuromodulator-based learning requires a weaker supervision signal, making use of binary rewards instead of explicit labels. Additionally, our model learns even in the absence of environmental feedback through Hebbian learning. Finally, weight modifications are based on synaptically-local information and on two signals broadcasted identically to all neurons, which matches capabilities of biological neural networks.

On the functional side, learning with DA and ACh has been shown to decisively improve classification performance in our model system. Although it was not the main focus of this study, we note that very high classification performances even for relatively small networks (compare sizes in Diehl and Cook, 2015; Neftci et al., 2015) could be achieved using neuromodulation. The use of neuromodulation in spiking neural systems for neuromorphic chips (Diehl and Cook, 2015; Neftci et al., 2015) is therefore likely to result in performance gains. Similarly, neuromodulation is expected to further improve performance of novel hierarchical networks with Hebbian learning (Forster et al., 2016) which have a functional focus on learning from data with very few labels.

It is interesting to note that, since the initial publication of the MNIST dataset, advances in gradient-based learning resulted in continuous and substantial decreases in error rates. The biologically-inspired method presented in this work is at a relatively early stage and we may expect similar improvements from future research.

AUTHOR CONTRIBUTIONS

RH carried out the simulations, analyzed the data, and designed the study with contributions of JL and KO. JL provided theoretical background and support for the neural network model. JL and KO helped revising the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported by Quebec's National Fund for Research in Nature and Technology (181120), the German National Academic Foundation, the German Research Foundation (GRK 1589 and LU 1196/5-1), and the Cluster of Excellence EXC 1077/1 Hearing4all.

REFERENCES

- Allard, T., Clark, S., Jenkins, W., and Merzenich, M. (1991). Reorganization of somatosensory area 3b representations in adult owl monkeys after digital syndactyly. *J. Neurophysiol.* 66, 1048–1058. doi: 10.1016/S0306-4522(02)00292-0
- Arnold, H., Burk, J., Hodgson, E., Sarter, M., and Bruno, J. (2002). Differential cortical acetylcholine release in rats performing a sustained attention task versus behavioral control tasks that do not explicitly tax attention. *Neuroscience* 114, 451–460. doi: 10.1016/S0306-4522(02)00292-0
- Arsenault, J. T., Nelissen, K., Jarraya, B., and Vanduffel, W. (2013). Dopaminergic reward signals selectively decrease fmri activity in primate visual cortex. *Neuron* 77, 1174–1186. doi: 10.1016/j.neuron.2013.01.008
- Assisi, C., Stopfer, M., Laurent, G., and Bazhenov, M. (2007). Adaptive regulation of sparseness by feedforward inhibition. *Nat. Neurosci.* 10, 1176–1184. doi: 10.1038/nn1947
- Bao, S., Chan, V. T., and Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412, 79–83. doi: 10.1038/35083586
- Bao, S., Chan, V. T., Zhang, L. L., and Merzenich, M. M. (2003). Suppression of cortical representation through backward conditioning. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1405–1408. doi: 10.1073/pnas.0337527100
- Bao, S., Chang, E. F., Woods, J., and Merzenich, M. M. (2004). Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat. Neurosci.* 7, 974–981. doi: 10.1038/nn1293
- Barkat, T. R., Polley, D. B., and Hensch, T. K. (2011). A critical period for auditory thalamocortical connectivity. *Nat. Neurosci.* 14, 1189–1194. doi: 10.1038/nn.2882
- Beitel, R. E., Schreiner, C. E., Cheung, S. W., Wang, X., and Merzenich, M. M. (2003). Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11070–11075. doi: 10.1073/pnas.1334187100
- Bissière, S., Humeau, Y., and Lüthi, A. (2003). Dopamine gates ltp induction in lateral amygdala by suppressing feedforward inhibition. *Nat. Neurosci.* 6, 587–592. doi: 10.1038/nn1058
- Blond, O., Crépel, F., and Otani, S. (2002). Long-term potentiation in rat prefrontal slices facilitated by phased application of dopamine. *Eur. J. Pharmacol.* 438, 115–116. doi: 10.1016/S0014-2999(02)01291-8
- Brembs, B., Lorenzetti, F. D., Reyes, F. D., Baxter, D. A., and Byrne, J. H. (2002). Operant reward learning in aplysia: neuronal correlates and mechanisms. *Science* 296, 1706–1709. doi: 10.1126/science.1069434
- Bröcher, S., Artola, A., and Singer, W. (1992). Agonists of cholinergic and noradrenergic receptors facilitate synergistically the induction of long-term potentiation in slices of rat visual cortex. *Brain Res.* 573, 27–36. doi: 10.1016/0006-8993(92)90110-U
- Brosch, M., Selezneva, E., and Scheich, H. (2011). Representation of reward feedback in primate auditory cortex. *Front. Syst. Neurosci.* 5:5. doi: 10.3389/fnsys.2011.00005
- Buchanan, K. A., Petrovic, M. M., Chamberlain, S. E. L., Marrion, N. V., and Mellor, J. R. (2010). Facilitation of long-term potentiation by muscarinic m1 receptors is mediated by inhibition of sk channels. *Neuron* 68, 948–963. doi: 10.1016/j.neuron.2010.11.018
- Butt, A. E., and Hodge, G. K. (1995). Acquisition, retention, and extinction of operant discriminations in rats with nucleus basalis magnocellularis lesions. *Behav. Neurosci.* 109:699. doi: 10.1037/0735-7044.109.4.699
- Butt, A. E., Testylier, G., and Dykes, R. W. (1997). Acetylcholine release in rat frontal and somatosensory cortex is enhanced during tactile discrimination learning. *Psychobiology* 25, 18–33.
- Calabresi, P., Picconi, B., Tozzi, A., and Di Filippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends Neurosci.* 30, 211–219. doi: 10.1016/j.tins.2007.03.001
- Calford, M. B., and Tweedale, R. (1988). Immediate and chronic changes in responses of somatosensory cortex in adult flying-fox after digit amputation. *Nature* 332, 446–448. doi: 10.1038/332446a0
- Centonze, D., Gubellini, P., Picconi, B., Calabresi, P., Giacomini, P., and Bernardi, G. (1999). Unilateral dopamine denervation blocks corticostriatal ltp. *J. Neurophysiol.* 82, 3575–3579.
- Chun, S., Bayazitov, I. T., Blundon, J. A., and Zakharenko, S. S. (2013). Thalamocortical long-term potentiation becomes gated after the early critical period in the auditory cortex. *J. Neurosci.* 33, 7345–7357. doi: 10.1523/JNEUROSCI.4500-12.2013
- Conner, J., Kulczycki, M., and Tuszynski, M. (2010). Unique contributions of distinct cholinergic projections to motor cortical plasticity and learning. *Cereb. Cortex* 20, 2739–2748. doi: 10.1093/cercor/bhq022
- Conner, J. M., Culberson, A., Packowski, C., Chiba, A. A., and Tuszynski, M. H. (2003). Lesions of the basal forebrain cholinergic system impair task acquisition and abolish cortical plasticity associated with motor skill learning. *Neuron* 38, 819–829. doi: 10.1016/S0896-6273(03)00288-5
- Dalley, J. W., McGaughy, J., O'Connell, M. T., Cardinal, R. N., Levita, L., and Robbins, T. W. (2001). Distinct changes in cortical acetylcholine and noradrenaline efflux during contingent and noncontingent performance of a visual attentional task. *J. Neurosci.* 21, 4908–4914.
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879. doi: 10.1038/nature04766
- Dayan, P., and Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron* 36, 285–298. doi: 10.1016/S0896-6273(02)00963-7
- de Villers-Sidani, E., Chang, E. F., Bao, S., and Merzenich, M. M. (2007). Critical period window for spectral tuning defined in the primary auditory cortex (a1) in the rat. *J. Neurosci.* 27, 180–189. doi: 10.1523/JNEUROSCI.3227-06.2007
- Diehl, P. U., and Cook, M. (2015). Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Front. Comput. Neurosci.* 9:99. doi: 10.3389/fncom.2015.00099
- Dinse, H. R., Ragert, P., Pleger, B., Schwenkreis, P., and Tegenthoff, M. (2003). Pharmacological modulation of perceptual learning and associated cortical reorganization. *Sci. Signal.* 301, 91. doi: 10.1126/science.1085423
- Eggermont, J. J., and Roberts, L. E. (2004). The neuroscience of tinnitus. *Trends Neurosci.* 27, 676–682. doi: 10.1016/j.tins.2004.08.010
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305.
- Fletcher, M. L., and Wilson, D. A. (2002). Experience modifies olfactory acuity: acetylcholine-dependent learning decreases behavioral generalization between similar odorants. *J. Neurosci.* 22:RC201.
- Flor, H., Nikolajsen, L., and Jensen, T. S. (2006). Phantom limb pain: a case of maladaptive cns plasticity? *Nat. Rev. Neurosci.* 7, 873–881. doi: 10.1038/nrn1991
- Forster, D., and Lücke, J. (2017). Truncated variational em for semi-supervised neural simpletrons. *arXiv preprint arXiv:1702.01997*.
- Forster, D., Sheikh, A.-S., and Lücke, J. (2016). Neural simpletrons—minimalistic directed generative networks for learning with few labels. *Stat* 1050:23.

- Frankó, E., Seitz, A. R., and Vogels, R. (2010). Dissociable neural effects of long-term stimulus-reward pairing in macaque visual cortex. *J. Cogn. Neurosci.* 22, 1425–1439. doi: 10.1162/jocn.2009.21288
- Freund, Y., Schapire, R., and Abe, N. (1999). A short introduction to boosting. *J. Jpn. Soc. Artif. Intell.* 14, 1612.
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223. doi: 10.1038/nn1141
- Froemke, R. C., Carcea, I., Barker, A. J., Yuan, K., Seybold, B. A., Martins, A. R. O., et al. (2013). Long-term modification of cortical synapses improves sensory perception. *Nat. Neurosci.* 16, 79–88. doi: 10.1038/nn.3274
- Froemke, R. C., Merzenich, M. M., and Schreiner, C. E. (2007). A synaptic memory trace for cortical receptive field plasticity. *Nature* 450, 425–429. doi: 10.1038/nature06289
- Gambino, F., and Holtmaat, A. (2012). Spike-timing-dependent potentiation of sensory surround in the somatosensory cortex is facilitated by deprivation-mediated disinhibition. *Neuron* 75, 490–502. doi: 10.1016/j.neuron.2012.05.020
- Giessel, A. J., and Sabatini, B. L. (2010). M1 muscarinic receptors boost synaptic potentials and calcium influx in dendritic spines by inhibiting postsynaptic sk channels. *Neuron* 68, 936–947. doi: 10.1016/j.neuron.2010.09.004
- Godde, B., Leonhardt, R., Cords, S. M., and Dinse, H. R. (2002). Plasticity of orientation preference maps in the visual cortex of adult cats. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6352–6357. doi: 10.1073/pnas.082407499
- Graybiel, A. M. (2005). The basal ganglia: learning new tricks and loving it. *Curr. Opin. Neurobiol.* 15, 638–644. doi: 10.1016/j.conb.2005.10.006
- Greuel, J. M., Luhmann, H. J., and Singer, W. (1988). Pharmacological induction of use-dependent receptive field modifications in the visual cortex. *Science* 242, 74–77. doi: 10.1126/science.2902687
- Gu, Q. (2003). Contribution of acetylcholine to visual cortex plasticity. *Neurobiol. Learn. Mem.* 80, 291–301. doi: 10.1016/S1074-7427(03)00073-X
- Haber, S. N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* 35, 4–26. doi: 10.1038/npp.2009.129
- Halligan, P. W., Marshall, J. C., Wade, D. T., Davey, J., and Morrison, D. (1993). Thumb in cheek? sensory reorganization and perceptual plasticity after limb amputation. *Neuroreport* 4, 233–236. doi: 10.1097/00001756-199303000-00001
- Han, Y. K., Köver, H., Insanally, M. N., Semerdjian, J. H., and Bao, S. (2007). Early experience impairs perceptual discrimination. *Nat. Neurosci.* 10, 1191–1197. doi: 10.1038/nn1941
- Harris, J. A., Harris, I. M., and Diamond, M. E. (2001). The topography of tactile learning in humans. *J. Neurosci.* 21, 1056–1061.
- Hasselmo, M. E., and Sarter, M. (2011). Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology* 36, 52–73. doi: 10.1038/npp.2010.104
- Himmelheber, A. M., Sarter, M., and Bruno, J. P. (2000). Increases in cortical acetylcholine release during sustained attention performance in rats. *Cogn. Brain Res.* 9, 313–325. doi: 10.1016/S0926-6410(00)00012-4
- Hosp, J. A., Pektanovic, A., Rioult-Pedotti, M. S., and Luft, A. R. (2011). Dopaminergic projections from midbrain to primary motor cortex mediate motor skill learning. *J. Neurosci.* 31, 2481–2487. doi: 10.1523/JNEUROSCI.5411-10.2011
- Isaacson, J. S., and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron* 72, 231–243. doi: 10.1016/j.neuron.2011.09.027
- Ji, W., Gao, E., and Suga, N. (2001). Effects of acetylcholine and atropine on plasticity of central auditory neurons caused by conditioning in bats. *J. Neurophysiol.* 86, 211–225.
- Keck, C., Savin, C., and Lücke, J. (2012). Feedforward inhibition and synaptic scaling—two sides of the same coin? *PLoS Comput. Biol.* 8:e1002432. doi: 10.1371/journal.pcbi.1002432
- Kilgard, M. P., and Merzenich, M. M. (1998a). Cortical map reorganization enabled by nucleus basalis activity. *Science* 279, 1714–1718. doi: 10.1126/science.279.5357.1714
- Kilgard, M. P., and Merzenich, M. M. (1998b). Plasticity of temporal information processing in the primary auditory cortex. *Nat. Neurosci.* 1, 727–731. doi: 10.1038/3729
- Kim, H., and Bao, S. (2009). Selective increase in representations of sounds repeated at an ethological rate. *J. Neurosci.* 29, 5163–5169. doi: 10.1523/JNEUROSCI.0365-09.2009
- Kingma, D., and Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, abs/1412.6980.
- Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., and Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science* 318, 1642–1645. doi: 10.1126/science.1145044
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43, 59–69. doi: 10.1007/BF00337288
- Kozak, R., Bruno, J. P., and Sarter, M. (2006). Augmented prefrontal acetylcholine release during challenged attentional performance. *Cereb. Cortex* 16, 9–17. doi: 10.1093/cercor/bhi079
- Kudoh, M., and Shibuki, K. (2006). Sound sequence discrimination learning motivated by reward requires dopaminergic d2 receptor activation in the rat auditory cortex. *Learn. Mem.* 13, 690–698. doi: 10.1101/lm.390506
- Law, C.-T., and Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual decision task. *Nat. Neurosci.* 12:655. doi: 10.1038/nn.2304
- Leach, N. D., Nodal, F. R., Cordery, P. M., King, A. J., and Bajo, V. M. (2013). Cortical cholinergic input is required for normal auditory perception and experience-dependent plasticity in adult ferrets. *J. Neurosci.* 33, 6659–6671. doi: 10.1523/JNEUROSCI.5039-12.2013
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998a). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791
- LeCun, Y., Cortes, C., and Burges, C. J. (1998b). *The Mnist Database of Handwritten Digits*.
- Li, G., and Cleland, T. A. (2013). A two-layer biophysical model of cholinergic neuromodulation in olfactory bulb. *J. Neurosci.* 33, 3037–3058. doi: 10.1523/JNEUROSCI.2831-12.2013
- Li, S., Cullen, W. K., Anwyl, R., and Rowan, M. J. (2003). Dopamine-dependent facilitation of ltp induction in hippocampal ca1 by exposure to spatial novelty. *Nat. Neurosci.* 6, 526–531. doi: 10.1038/nn1049
- Liu, J., Lu, Z.-L., and Doshier, B. A. (2010). Augmented hebbian reweighting: interactions between feedback and training accuracy in perceptual learning. *J. Vis.* 10:29. doi: 10.1167/10.10.29
- Lücke, J. (2009). Receptive field self-organization in a model of the fine-structure in V1 cortical columns. *Neural Comput.* 21, 2805–2845. doi: 10.1162/neco.2009.07-07-584
- Luft, A. R., and Schwarz, S. (2009). Dopaminergic signals in primary motor cortex. *Int. J. Dev. Neurosci.* 27, 415–421. doi: 10.1016/j.ijdevneu.2009.05.004
- Ma, X., and Suga, N. (2005). Long-term cortical plasticity evoked by electric stimulation and acetylcholine applied to the auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9335–9340. doi: 10.1073/pnas.0503851102
- Maaten, L. V. D., and Hinton, G. (2008). Visualizing data using t-sne. *J. Mach. Learn. Res.* 9, 2579–2605.
- Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., and Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nat. Neurosci.* 8, 1690–1697. doi: 10.1038/nn1556
- Matsuda, Y., Marzo, A., and Otani, S. (2006). The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex. *J. Neurosci.* 26, 4803–4810. doi: 10.1523/JNEUROSCI.5312-05.2006
- McGaughy, J., Dalley, J., Morrison, C., Everitt, B., and Robbins, T. (2002). Selective behavioral and neurochemical effects of cholinergic lesions produced by intrabasal infusions of 192 igg-saporin on attentional performance in a five-choice serial reaction time task. *J. Neurosci.* 22, 1905–1913.
- Mittmann, W., Koch, U., and Häusser, M. (2005). Feed-forward inhibition shapes the spike output of cerebellar purkinje cells. *J. Physiol.* 563, 369–378. doi: 10.1113/jphysiol.2004.075028
- Molina-Luna, K., Pektanovic, A., Rohrich, S., Hertler, B., Schubring-Giese, M., Rioult-Pedotti, M.-S., et al. (2009). Dopamine in motor cortex is necessary for skill learning and synaptic plasticity. *PLoS ONE* 4:e7082. doi: 10.1371/journal.pone.0007082
- Navakkode, S., Sajikumar, S., and Frey, J. U. (2007). Synergistic requirements for the induction of dopaminergic d1/d5-receptor-mediated ltp in

- hippocampal slices of rat ca1 in vitro. *Neuropharmacology* 52, 1547–1554. doi: 10.1016/j.neuropharm.2007.02.010
- Neftci, E. O., Pedroni, B. U., Joshi, S., Al-Shedivat, M., and Cauwenberghs, G. (2015). Unsupervised learning in synaptic sampling machines. *arXiv preprint arXiv:1511.04484*.
- Nessler, B., Pfeiffer, M., Buesing, L., and Maass, W. (2013). Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Comput. Biol.* 9:e1003037. doi: 10.1371/journal.pcbi.1003037
- Olsen, S. R., and Wilson, R. I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature* 452, 956–960. doi: 10.1038/nature06864
- O'Reilly, R. C. (2001). Generalization in interactive networks: The benefits of inhibitory competition and hebbian learning. *Neural Comput.* 13, 1199–1241. doi: 10.1162/08997660152002834
- Otani, S., Blond, O., Desce, J.-M., and Crepel, F. (1998). Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex. *Neuroscience* 85, 669–676. doi: 10.1016/S0306-4522(97)00677-5
- Ott, T., Jacob, S. N., and Nieder, A. (2014). Dopamine receptors differentially enhance rule coding in primate prefrontal cortex neurons. *Neuron* 84, 1317–1328. doi: 10.1016/j.neuron.2014.11.012
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature* 392, 811–814. doi: 10.1038/33918
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., and Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 12, 169–174. doi: 10.1097/00001756-200101220-00041
- Parikh, V., Kozak, R., Martinez, V., and Sarter, M. (2007). Prefrontal acetylcholine release controls cue detection on multiple timescales. *Neuron* 56, 141–154. doi: 10.1016/j.neuron.2007.08.025
- Passetti, F., Dalley, J., O'Connell, M., Everitt, B., and Robbins, T. (2000). Increased acetylcholine release in the rat medial prefrontal cortex during performance of a visual attentional task. *Eur. J. Neurosci.* 12, 3051–3058. doi: 10.1046/j.1460-9568.2000.00183.x
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Pleger, B., Ruff, C. C., Blankenburg, F., Klöppel, S., Driver, J., and Dolan, R. J. (2009). Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biol.* 7:e1000164. doi: 10.1371/journal.pbio.1000164
- Polley, D. B., Steinberg, E. E., and Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *J. Neurosci.* 26, 4970–4982. doi: 10.1523/JNEUROSCI.3771-05.2006
- Poort, J., Khan, A. G., Pachitariu, M., Nemri, A., Orsolich, I., Krupic, J., et al. (2015). Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* 86, 1478–1490. doi: 10.1016/j.neuron.2015.05.037
- Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B. V., and Scanziani, M. (2009). Input normalization by global feedforward inhibition expands cortical dynamic range. *Nat. Neurosci.* 12, 1577–1585. doi: 10.1038/nn.2441
- Pouille, F., and Scanziani, M. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science* 293, 1159–1163. doi: 10.1126/science.1060342
- Ramachandran, V. S., Stewart, M., and Rogers-Ramachandran, D. (1992). Perceptual correlates of massive cortical reorganization. *Neuroreport* 3, 583–586. doi: 10.1097/00001756-199207000-00009
- Recanzone, G. A., Schreiner, C., and Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103.
- Recanzone, G. H., Merzenich, M. M., Jenkins, W. M., Grajski, K. A., and Dinse, H. R. (1992). Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *J. Neurophysiol.* 67, 1031–1056.
- Reed, A., Riley, J., Carraway, R., Carrasco, A., Perez, C., Jakkamsetti, V., et al. (2011). Cortical map plasticity improves learning but is not necessary for improved performance. *Neuron* 70, 121–131. doi: 10.1016/j.neuron.2011.02.038
- Roelfsema, P. R., and Ooyen, A. v. (2005). Attention-gated reinforcement learning of internal representations for classification. *Neural Comput.* 17, 2176–2214. doi: 10.1162/0899766054615699
- Roelfsema, P. R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends Cogn. Sci.* 14:64. doi: 10.1016/j.tics.2009.11.005
- Rombouts, J., Roelfsema, P., and Bohte, S. M. (2012). “Neurally plausible reinforcement learning of working memory tasks” in *Advances in Neural Information Processing Systems*, eds F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Lake Tahoe, NV: NIPS), 1871–1879.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1985). *Learning Internal Representations by Error Propagation*. Technical Report, DTIC Document.
- Sarter, M., Gehring, W. J., and Kozak, R. (2006). More attention must be paid: the neurobiology of attentional effort. *Brain Res. Rev.* 51, 145–160. doi: 10.1016/j.brainresrev.2005.11.002
- Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23, 9913–9923.
- Schicknick, H., Reichenbach, N., Smalla, K.-H., Scheich, H., Gundelfinger, E. D., and Tischmeyer, W. (2012). Dopamine modulates memory consolidation of discrimination learning in the auditory cortex. *Eur. J. Neurosci.* 35, 763–774. doi: 10.1111/j.1460-9568.2012.07994.x
- Schmuker, M., Pfeil, T., and Nawrot, M. P. (2014). A neuromorphic network for generic multivariate data classification. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2081–2086. doi: 10.1073/pnas.1303053111
- Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in v1 neurons. *Nature* 412, 549–553. doi: 10.1038/35087601
- Schultz, W. (2007). Behavioral dopamine signals. *Trends Neurosci.* 30, 203–210. doi: 10.1016/j.tins.2007.03.007
- Schultz, W. (2010). Review dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6:24. doi: 10.1186/1744-9081-6-24
- 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
- Sclar, G., Maunsell, J. H., and Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vis. Res.* 30, 1–10. doi: 10.1016/0042-6989(90)90123-3
- Sengpiel, F., Stawinski, P., and Bonhoeffer, T. (1999). Influence of experience on orientation maps in cat visual cortex. *Nat. Neurosci.* 2, 727–732. doi: 10.1038/11192
- Sigala, N., and Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415, 318–320. doi: 10.1038/4151318a
- Stopfer, M., Jayaraman, V., and Laurent, G. (2003). Intensity versus identity coding in an olfactory system. *Neuron* 39, 991–1004. doi: 10.1016/j.neuron.2003.08.011
- Suga, N. (2012). Tuning shifts of the auditory system by corticocortical and corticofugal projections and conditioning. *Neurosci. Biobehav. Rev.* 36, 969–988. doi: 10.1016/j.neubiorev.2011.11.006
- Sun, X., Zhao, Y., and Wolf, M. E. (2005). Dopamine receptor stimulation modulates ampa receptor synaptic insertion in prefrontal cortex neurons. *J. Neurosci.* 25, 7342–7351. doi: 10.1523/JNEUROSCI.4603-04.2005
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*, Vol. 135. Cambridge, MA: MIT press Cambridge.
- Swadlow, H. A. (2003). Fast-spike interneurons and feedforward inhibition in awake sensory neocortex. *Cereb. Cortex* 13, 25–32. doi: 10.1093/cercor/13.1.25
- Tegenthoff, M., Ragert, P., Pleger, B., Schwenkreis, P., Förster, A.-F., Nicolas, V., et al. (2005). Improvement of tactile discrimination performance and enlargement of cortical somatosensory maps after 5 Hz rTMS. *PLoS Biol.* 3:e362. doi: 10.1371/journal.pbio.0030362
- Tobler, P. N., Fiorillo, C. D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645. doi: 10.1126/science.1105370
- Vickery, T. J., Chun, M. M., and Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron* 72, 166–177. doi: 10.1016/j.neuron.2011.08.011
- Wang, Y., Sereno, J. A., Jongman, A., and Hirsch, J. (2003). fMRI evidence for cortical modification during learning of mandarin lexical tone. *J. Cogn. Neurosci.* 15, 1019–1027. doi: 10.1162/089892903770007407

- Watkins, C. J., and Dayan, P. (1992). Q-learning. *Mach. Learn.* 8, 279–292. doi: 10.1007/BF00992698
- Wehr, M., and Zador, A. M. (2005). Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron* 47, 437–445. doi: 10.1016/j.neuron.2005.06.009
- Weinberger, N. M. (2003). The nucleus basalis and memory codes: auditory cortical plasticity and the induction of specific, associative behavioral memory. *Neurobiol. Learn. Mem.* 80, 268–284. doi: 10.1016/S1074-7427(03)00072-8
- Weinberger, N. M., and Bakin, J. S. (1998). Learning-induced physiological memory in adult primary auditory cortex: receptive field plasticity, model, and mechanisms. *Audiol. Neurotol.* 3, 145–167. doi: 10.1159/000013787
- Werbos, P. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral sciences*. PhD Thesis, Harvard University.
- Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* 8, 229–256. doi: 10.1007/BF00992696
- Wilson, D. A., Fletcher, M. L., and Sullivan, R. M. (2004). Acetylcholine and olfactory perceptual learning. *Learn. Mem.* 11, 28–34. doi: 10.1101/lm.66404
- Wise, R. A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–494. doi: 10.1038/nrn1406
- Yang, T., and Maunsell, J. H. (2004). The effect of perceptual learning on neuronal responses in monkey visual area v4. *J. Neurosci.* 24, 1617–1626. doi: 10.1523/JNEUROSCI.4442-03.2004
- Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transact. Math. Soft.* 23, 550–560. doi: 10.1145/279232.279236

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Holca-Lamarre, Lücke and Obermayer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A. APPENDIX

A.1. Code and data

The model was written in the Python programming language and was run on a computer cluster. The code for the neural network is available at <https://github.com/raphaelholca/hebbianRL>. The original MNIST dataset is available at <http://yann.lecun.com/exdb/mnist/>. The dataset is randomly split into training and testing sets; the network's performance is reported on the testing images not seen during training. The network with 300 hidden units used for performance comparison with other work was trained with the full (unbalanced) dataset. For all other results, the datasets were balanced so that they contain the same number of examples for each digit class. This balancing has negligible effects on the results.

A.2. Weight initialisation

We pre-compute the activation of input neurons \bar{y} through Equation 1 for the whole training dataset. Learning proceeds through full iterations over the dataset during which \bar{y} are presented in a random order to the network. Weights of representation neurons are initialised using the statistics of the input images. Specifically, we initialise the weights with the mean activation of input neurons taken over the whole dataset, with the addition of noise to break symmetry:

$$W_{cd} = \mu(y_d) - \sigma^2(y_d) \cdot \eta_{\text{init}}, \quad (\text{A1})$$

where $\mu(\cdot)$ and $\sigma^2(\cdot)$ are the mean and variance taken over all training images N , respectively, and η_{init} is noise drawn from a uniform distribution in the interval $[0,0, 2.0)$. Activations propagate through the network as a succession of Equations 2, 4, 5. Values for all hyper-parameters were found through grid search (see **Table A1**).

A.3. Batch learning

To speed up computation, we train the network using mini-batches; weight updates are computed over batches of 50 training examples. Using mini-batches only negligibly affects representation learning and the network's performance.

In the case of DA-based learning, negative learning rates (for absent expected rewards, $+pred -rew$) could potentially result in negative weights. For biological realism and computational stability, we prevent this by excluding weight updates for a representation neuron c if any weight W_{cd} would become negative after the weight update. For the parameter set presented in **Table A1** in Supplementary Methods, this rule only rarely prevents learning ($\sim 0.1\%$ of all batch updates). However,

TABLE A1 | Hyper-parameters used in training the network. Values were determined through parameter exploration.

Parameter	Description	Value
α	Amplitude of the sigmoid function for ACh release	2.0
β	Slope of the sigmoid function for ACh release	20.0
$\delta_{+/+}$	DA activation for correctly predicted reward	0.01
$\delta_{+/-}$	DA activation for incorrectly predicted reward	-1.0
$\delta_{-/+}$	DA activation for unexpected reward	4.0
$\delta_{-/-}$	DA activation for correctly predicted absence of reward	-0.25
A	Normalization constant for feedforward inhibition	1.0×10^3
ϵ	Learning rate	5.0×10^{-3}
ν	Variance of the normal distribution of noise η	0.3

TABLE A2 | Hyper-parameters for the benchmarking algorithms, as implemented in the *Scikit-learn* module.

Parameter	Description	Value
hidden_layer_sizes	Number of neurons in the hidden layer	300
activation	Activation function for the hidden layer	'relu'
algorithm	Algorithm for weight optimisation	'adam' or 'l-bfgs'
alpha	Regularisation term (L2 penalty)	1e-06
batch_size	Size of mini-batches for stochastic optimisation	200
learning_rate_init	Initial learning rate (Adam)	0.001
beta_1	Exp. decay for estimates of 1st moment (Adam)	0.8
beta_2	Exp. decay for estimates of 2nd moment (Adam)	0.9
epsilon	Value for numerical stability (Adam)	1e-08

when performing parameter exploration of the $\delta_{j/}$ variables, some parameter sets lead to rapid decay to negative weight values, and this rule is then necessary to ensure computational stability.

A.4. Comparison benchmarks

The MLP algorithm was obtained from the *Scikit-learn* module (Pedregosa et al., 2011) (version 18.dev0, downloaded on 04/29/16). We used 3-fold cross-validation and grid search to determine the values of the hyper-parameters (see **Table A2**). The two optimisation methods used to train the MLP were the Adam and L-BFGS algorithms. Adam is a first-order stochastic optimisation method that uses individual adaptive learning rates for the different parameters. L-BFGS (Limited-memory Broyden-Fletcher-Goldfarb-Shanno) is a quasi-Newton method.