JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Behavioral/Cognitive

On the role of cortex-basal ganglia interactions for category learning: A neuro-computational approach

Francesc Villagrasa[1], Javier Baladron[1], Julien Vitay[1], Henning Schroll[1], Evan G. Antzoulatos[2], Earl K. Miller[3] and Fred H. Hamker[1]

[1]*Chemnitz University of Technology, Department of Computer Science, 09107 Chemnitz, Germany*
[2]*UC Davis Center for Neuroscience and Department of Neurobiology, Physiology and Behavior, Davis, CA 95616, United States*
[3]*The Picower Institute for Learning and Memory and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, United States*

DOI: 10.1523/JNEUROSCI.0874-18.2018

Received: 5 April 2018

Revised: 7 August 2018

Accepted: 28 August 2018

Published: 18 September 2018

Author contributions: F.V., J.V., E.G.A., and F.H.H. performed research; F.V., J.B., J.V., H.S., E.G.A., and E.K.M. analyzed data; F.V. wrote the first draft of the paper; J.B. and F.H.H. designed research; J.B., J.V., H.S., E.G.A., E.K.M., and F.H.H. edited the paper; F.H.H. wrote the paper.

Conflict of Interest: The authors declare no competing financial interests.

This work has been supported by the German Research Foundation (DFG, grant agreements no. HA2630/4-2 and HA2630/8-1), the European Social Fund and the Free State of Saxony (ESF, grant agreement no. ESF-100269974), the NIMH R01MH065252, and the MIT Picower Institute Innovation Fund.

Corresponding author: Fred H. Hamker, fred.hamker@informatik.tu-chemnitz.de, 09107 Chemnitz, Germany

Cite as: J. Neurosci ; 10.1523/JNEUROSCI.0874-18.2018

On the role of cortex-basal ganglia interactions for category learning: A neuro-computational approach

Abbreviated title: Category learning in cortex and basal ganglia

Authors: Francesc Villagrasa[1], Javier Baladron[1], Julien Vitay[1], Henning Schroll[1], Evan G.Antzoulatos[2], Earl K. Miller[3], and Fred H. Hamker[1]

Affiliations:

[1] Chemnitz University of Technology, Department of Computer Science, 09107 Chemnitz, Germany

[2] UC Davis Center for Neuroscience and Department of Neurobiology, Physiology and Behavior, Davis, CA 95616, United States

[3] The Picower Institute for Learning and Memory and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

Corresponding author: Fred H. Hamker, fred.hamker@informatik.tu-chemnitz.de, 09107 Chemnitz, Germany

Number of pages: 32

Number of figures: 14

Number of words: 213 (abstract), 591 (introduction), 977 (discussion)

Conflict of Interest: The authors declare no competing financial interests

Acknowledgements: This work has been supported by the German Research Foundation (DFG, grant agreements no. HA2630/4-2 and HA2630/8-1), the European Social Fund and the Free State of Saxony (ESF, grant agreement no. ESF-100269974), the NIMH R01MH065252, and the MIT Picower Institute Innovation Fund.

Abstract

213 words out of 250 maximum.

In addition to the prefrontal cortex (PFC), the basal ganglia (BG) have been increasingly often reported to play a fundamental role in category learning, but the systems-level circuits of how both interact remain to be explored. We developed a novel neuro-computational model of category learning that particularly addresses the BG-PFC interplay. We propose that the BG bias PFC activity by removing the inhibition of cortico-thalamo-cortical loop and thereby provide a teaching signal to guide the acquisition of category representations in the cortico-cortical associations to the PFC. Our model replicates key behavioral and physiological data of macaque monkey learning a prototype distortion task from Antzoulatos and Miller (2011). Our simulations allowed us to gain a deeper insight into the observed drop of category selectivity in striatal neurons seen in the experimental data and in the model. The simulation results and a new analysis of the experimental data, based on the model's predictions, show that the drop in category selectivity of the striatum emerges as the variability of responses in the striatum rises when confronting the BG with an increasingly larger number of stimuli to be classified. The neuro-computational model therefore provides new testable insights of systems-level brain circuits involved in category learning which may also be generalized to better understand other cortico-basal ganglia-cortical loops.

Significance Statement

119 words out of 120 maximum.

Inspired by the idea that basal ganglia (BG) teach the prefrontal cortex (PFC) to acquire category representations, we developed a novel neuro-computational model and tested it on a task that was recently applied in monkey experiments. As an advantage over previous models of category learning, our model allows to compare simulation data with single cell recordings in PFC and BG. We not only derived model predictions, but already verified a prediction to explain the observed drop in striatal category selectivity. When testing our model with a simple real-world face categorization task, we observed that the fast striatal learning with a performance of 85% correct responses can teach slower PFC learning to push the model performance up to almost 100%.

Introduction

591 words out of 650

The world is composed of an overwhelming number of different objects and variants of those objects. Category formation is the ability to extract commonalities among these diverse objects, allowing us to group experiences by concepts or categories, and therefore imbuing our world with meaning. Furthermore, we can generalize, and hence classify, stimuli we have never seen before into a category, a property also fundamental for the emergence of language.

At least two brain areas are involved in category learning: the basal ganglia (BG) and the prefrontal cortex (PFC) (Seger and Miller, 2010). The BG have been shown to participate in a wide range of categorization tasks, particularly those that require implicit learning via trial and error (Merchant et al., 1997; Poldrack et al., 1999, 2001; Seger and Cincotta, 2005; Nomura et al., 2007; Cincotta and Seger, 2007; Zeithamova et al., 2008). The PFC, in contrast, appears to hold category knowledge. Freedman et al. (2001, 2002, 2003) found PFC neurons that became preferably activated by stimuli of a particular category. Also, PFC cells are known to represent abstract rule-based categories (Wallis et al., 2001; Wallis and Miller, 2003; Muhammad et al., 2006; Antzoulatos and Miller, 2016).

Some studies have suggested that the BG may train the PFC to slowly learn categories (Pasupathy and Miller, 2005; Miller and Buschman, 2008; Seger and Miller, 2010; Antzoulatos and Miller, 2011; Hélie et

al., 2015). Antzoulatos and Miller (2011) carried out an experiment in which monkeys were trained to classify a large number of different abstract stimuli composed of several dots into two possible categories. While monkeys learned this task, neurons from the PFC and the striatum were recorded. Early in this experiment, when there were just a few stimuli to classify, category selectivity was strong in the striatum, but weak in the PFC. As the task advanced, the number of possible stimuli to classify increased and the category selectivity became weak in the striatum and strong in the PFC (Antzoulatos and Miller, 2011).

The fact that the striatum predicted categories better in the beginning of the experiment and the PFC later led Antzoulatos and Miller (2011) to suggest that the BG teach the PFC to encode categories. However, there is no obvious explanation for the observed decrease in striatal category selectivity. Further, the exact relationship between BG and PFC during category formation, e.g. the systems-level circuits that allow the BG to teach the PFC are not yet fully worked out.

To study these open questions, we here developed a neuro-computational model and had it learn the experiment devised by Antzoulatos and Miller (2011). Our simulations suggest that although the striatal cells decrease on average their category selectivity, they typically remain selective enough to contribute to the final category decision: the knowledge acquired by the striatal cells can be very specific but also associated to several stimuli of the same category. Furthermore, our simulations predict that the striatal category selectivity decrease is due to an increase in the variability in the striatum cells' category response, i.e. the striatal cells only respond to a subset of stimuli of one category as well as to some stimuli of the other category. We supported this prediction by re-analyzing the original experimental data of Antzoulatos and Miller (2011).

In addition to the task used by Antzoulatos and Miller (2011), the model was tested on a task in which real-world face images had to be classified. This study revealed that even an imperfect teacher (the BG) can still train the PFC to push the model's classification performance up to almost 100%.

Methodology

Model Description

Overview

Our model comprises an open cortico-basalganglio-thalamic (CBGT) loop that interacts with a cortico-cortical-thalamo-cortical pathway to acquire category information and to produce category decisions. The two cortical areas involved are the Inferior Temporal cortex (IT) and the PFC (see Figure 1); the IT encodes stimulus information and the PFC learns to encode category knowledge. The BG bias activity in PFC such that Hebbian learning of the IT - PFC connectivity is sufficient to develop category selective cells in PFC.

In this rate coding model, the membrane potential of all simulated neurons and the learning rules that determine synaptic plasticity between neighboring neurons are controlled by differential equations.

The BG

Our BG model is based on previous work (Schroll et al., 2014, 2015) , and contains three basal ganglia pathways (Schroll and Hamker, 2013): the direct (striatum $\rightarrow$ substantia nigra pars reticulata), hyperdirect (subthalamic nucleus $\rightarrow$ substantia nigra pars reticulata), and short indirect pathway (striatum $\rightarrow$ external globus pallidus $\rightarrow$ substantia nigra pars reticulata). Each of these three BG pathways obtains the input information from the IT and converges in the substantia nigra pars reticulata (SNr), a BG nucleus that tonically inhibits the ventral anterior nucleus (VA) of the thalamus.

The function of each BG pathway emerges as a learning process, implemented via a three factor learning rule which considers the pre-synaptic activity, the post-synaptic activity and a dopamine (DA) signal. In our model, this DA signal estimates a reward prediction error based on the striatal activity at the time of reward delivery.

In the direct pathway, learning occurs in the projections between the IT and the striatal D1 cells and between the striatal D1 cells and the SNr. Associations between neurons in these connections become strengthened with dopamine burts and weakened with dopamine dips as motivated by experimental data (Shen et al., 2008; Fisher et al., 2017). Consequently, this pathway learns to select a patch of VA neurons that are linked with the correct category decision, in agreement with the well-known GO-function of this

BG pathway (Nambu et al., 2002; O'Reilly and Frank, 2006; Braak and Del Tredici, 2008; Schroll and Hamker, 2013).

In the hyperdirect pathway, learning occurs in the connections between the IT and the STN and between the STN and the SNr. Associations between neurons in these connections are also strengthened with dopamine peaks and weakened with dopamine dips (Kreiss et al., 1996; Schroll et al., 2012). Particularly, this pathway learns to suppress VA cells that encode currently unrewarded responses. Thus, both the direct and hyperdirect pathways work together to facilitate the selection of the correct category decision, in agreement with the well-known center-surround structure (Nambu et al., 2002).

In the indirect pathway, learning takes place in the projections between the IT and the striatal D2 cells and between the striatal D2 cells and GPe. In contrast to the other two BG pathways, but consistent with biological evidence (Surmeier et al., 2007; Shen et al., 2008; Fisher et al., 2017), associations between cells of this pathway become strengthened with dopamine dips and weakened with dopamine peaks. Therefore, this pathway learns to suppress VA cells linked to an incorrect category decision, in accordance with the well-documented NO-GO-function of this BG pathway (Apicella et al., 1992; Mink, 1996). This pathway is particularly relevant if changes in the stimulus-response associations occur.

A specific connectivity pattern is not forced on any of these plastic projections, providing our model with high flexibility. Connections are initialized in an all to all configuration with random low weights. The connectivity pattern is then automatically shaped through plasticity. On many previous modeling approaches of the BG a connectivity pattern with parallel channels (one for each action or here category) was enforced, without any clear account on how this arrangement could develop. Plasticity was therefore required only on early stages of the different pathways. An interesting feature of having plasticity in the late stages is that the knowledge acquired in the early stages of the pathways can be kept when learning a new task, allowing relearning to be faster than the initial learning as shown by Schroll et al. (2012).

The basal ganglia-cortex interaction

Our model includes a cortico-thalamo-cortical pathway which allows the BG to teach category knowledge to the cortico-cortical pathway from IT to PFC by biasing thalamic and thus, PFC activity. Once the category knowledge in the PFC is established, the PFC can also contribute to the final category decision by means of the cortico-thalamo-cortical pathway. Thus, the thalamus plays a key role in integrating the category decisions produced by both the BG and the PFC.

Category information is learned in the cortico-cortical connections between the IT and the PFC by an unsupervised Hebbian learning rule (please refer to the discussion regarding the assumption of unsupervised learning). As the BG disinhibits the thalamus, BG will bias PFC activation, which in turn guides (dopamine-free) Hebbian learning in the IT-PFC connections. The PFC cells in our model slowly learn over a large number of stimuli to extract category representations, in agreement with ideas suggesting that slow learning in the cortex is required to develop category representations in the PFC (Seger and Miller, 2010). Evidence has been found for the existence of Hebbian plasticity in cortico-cortical long-range connections (Sjöström et al., 2001; Koch et al., 2013).

Experimental design and statistical analysis

Prototype distortion task

In the experiment carried out by Antzoulatos and Miller (2011), two female monkeys performed a prototype distortion task in which they learned to classify stimuli into one of two different categories. We here re-analysed data from this previous experiment as explained later.

We tested our model with a very similar version of the original experiment as follows. Each stimulus was composed of 7 white small squares (7 x 7 pixels each) drawn on black background within an image of 140x140 pixels. Each stimulus belonged either to category A or B and was generated from the underlying category's prototype by shifting the seven squares from the prototype's coordinates randomly into nearby locations (Figure 2a). To mimic early visual processing up to area IT, we preprocessed the images using Gaussian receptive fields (RFs) with a standard deviation of 10 pixels (cut-off at 3.5 standard deviations which equals a diameter of 35 pixels), and a sampling distance between RF centers of 15 pixels (1.5 standard deviations of RF size).

The set of stimuli used in each experimental run consisted of 170 stimuli per category (each generated from its category's prototype image) and was distributed into 8 blocks, where the stimulus set increased in size with each block: in each block n the set size was 2^n, equally balanced for each category. In the first block, therefore, only two different stimuli were presented. In the second block two more stimuli were added to the set, reaching a total of 4. In subsequent blocks, only the stimuli added in the last block were kept and new stimuli were incorporated until a total of 2^n was reached. Figure 2b illustrates the exact procedure. Each new block began only when 16 out of the last 20 trials were successfully performed,

identical to the original experiment.

Because we aimed to focus on category learning only and did not model any eye-movement or working memory components involved in the original animal task, we simplified the trial design by omitting the delay period and the oculomotor response. At the beginning of each trial, a stimulus was randomly drawn from the set of the current block and presented to the model for 550 ms. After 50 ms we determined the model's decision using a softmax rule on a set of output neurons:

$$P_i = \frac{r_i + \theta}{(\sum_{j=1}^{N} r_j) + \theta} \tag{1}$$

where P_i is the probability of choosing category i, r_i is the rate of the output neuron associated to category i, N is the number of categories, and $\theta = 10^{-7}$ prevents from dividing by zero. The output neurons read our model's decision from the thalamic activity. Although 50 ms is a short time period, it is large enough for the model to reach a stable response to the presented stimulus. Data show that monkeys can make a decision between 25-50 ms if visual and motor latency are not considered (Stanford et al., 2010).

In the case of a correct response, dopaminergic SNc cells were excited for 500 ms, simulating the delivery of reward (reward period). To meaningfully compare our model's results with data from monkeys, we ran a very large number of experimental runs (100000) each with different initial synaptic weights and with slightly different values of 64 model parameters (see the mathematical model description). For each experimental run, a different set of stimuli was chosen among 100 possible sets of stimuli (each generated from two different category prototypes).

Model susceptibility to parameter variation

To study the susceptibility of our model to modest changes in model value parameters, we computed the correlation between the model performance and each of the 64 parameters modified in the 100000 experimental runs. Each of these correlations was computed with the Pearson correlation coefficient (PCC), employing the *corrcoef* numpy function, and considering 100000 data pairs, each made up of the model performance and the parameter value (or the absolute value of the distance between the parameter value and the mean parameter value, for a second version of the PCC) from a different experimental run. The model performance at each experimental run was evaluated by computing the average of correct

trials in the last 16 trials of the experimental run.

Category selectivity

In order to compare our model with the neurophysiological findings reported by Antzoulatos and Miller (2011), category selectivity was measured from model neurons' activity during the display of novel stimuli in correct trials, as previously done by the authors of the physiological experiment. As with the experimental data, category selectivity was computed within a trial-time window (size of 10 trials and 7 ms) moving in trial and time space (trial step size of 1 trial and a time step size of 3 ms). Trial-time windows with less than two trials associated to one category were discarded. The d' sensitivity index

$$d'(\mu_A, \mu_B, \sigma_A, \sigma_B) = \frac{|\mu_A - \mu_B|}{\sqrt{\frac{\sigma_A^2 \cdot (na-1) + \sigma_B^2 \cdot (nb-1)}{na+nb+2}}} \tag{2}$$

was computed for each cell within each window, where μ_A and σ_A are the mean and standard deviation, respectively, of the cell's firing rates recorded in trials where stimuli of category A were presented, μ_B and σ_B are the mean and standard deviation of the firing rates recorded in trials where stimuli of category B were presented, and na and nb are the number of trials in the corresponding window that relate to stimuli of category A and B, respectively. In the striatum, we only considered cells of the direct (Go-) pathway as these cells are mainly responsible for selection while cells in the indirect (No-Go-) pathway are responsible for suppression (Schroll et al., 2014).

Stimulus selectivity and category selectivity per cell

To study if cells in PFC and STR become stimulus selective rather than category selective, we applied the following procedure. At the end of each block, learning was frozen and each stimulus seen so far in the experiment was presented once to the model for 50 ms, followed by a period of 100 ms without a stimulus. The response of a cell to each presented stimulus was computed by averaging the cell's activity over 50 ms presentation time and normalized by its maximum response to all stimuli within a block.

We defined a stimulus selectivity index ($\mathrm{SI}_{\mathrm{stim}}$) which measures if a cell is particularly tuned to a single stimulus compared to the rest of the stimuli belonging to the same category:

$$\mathrm{SI}_{\mathrm{stim}} = \max(\forall s \in S.(R_s - \overline{R}_s)) \tag{3}$$

where s is a presented stimulus, S is the set of presented stimuli, R_s is the cell's response to s, and $\overline{R}_s$ is the mean cell's response to those stimuli in S that are different from s and belong to the same category as s.

Category selectivity ($\mathrm{SI}_{\mathrm{cat}}$) was measured by computing the absolute value of the difference between the cell's mean response to stimuli of one category and the cell's mean response to stimuli of the other category (i.e. the numerator of the d' sensitivity index).

$$\mathrm{SI}_{\mathrm{cat}} = |\overline{R}_A - \overline{R}_B| \tag{4}$$

where $\overline{R}_A$ is the mean cell's response to the stimuli that belong to category A, $\overline{R}_B$ is the mean cell's response to the stimuli that belong to category B.

We did not compute the full d' sensitivity index because we wanted the stimulus selectivity and the category selectivity to be plotted in the same scale. Therefore, the category selectivity was normalized in the same way as the stimulus selectivity (via normalizing the responses of each cell).

Only experiments that learned to criterion (in each block 16 out of 20 consecutive trials have to be correctly classified before the maximum number of trials determined in each block is reached) were considered for the analysis.

Face categorization task

To test the model's performance in a real-world classification scenario, we created an additional face categorization task. Face pictures of George W. Bush and Bill Clinton were extracted from videos and presented to the model for classification purposes.

All videos were taken from the *YouTube Faces Database* (Wolf et al., 2011), which consists of 3425 videos of 1595 different people, downloaded from Youtube and manually annotated. The shortest clip duration was 48 frames, the longest clip consisted of 6,070 frames, and the average length of a video clip was 181.3 frames. For Bill Clinton, we obtained 4 videos with a total of 851 frames and for George W. Bush, we

obtained 5 videos with a total of 820 frames.

For each frame, the face region was detected using a Viola-Jones filter (Viola and Jones, 2004), allowing to extract and resize each face to a 100x100 grayscale image. Figure 3 shows a few examples of the resulting face images.

To obtain high-level facial features that mimic the computation in visual areas, we trained a neural network using the *keras* library (https://keras.io/) and the *Theano* backend (http://deeplearning.net/software/theano/). The training set consisted of all the images obtained from the *YouTube Faces Database* except those of George W. Bush and Bill Clinton, providing us with 619455 input images of 1593 people (labels).

The neural network starts with a single convolutional layer, extracting 16 filters of size 6x6 and using a rectified linear transfer function (ReLU). It is followed by a max-pooling layer over 2x2 units and a dropout layer with $p = 0.5$. This layer feeds a fully connected layer (100 neurons, ReLU transfer function and dropout 0.5) which itself feeds a softmax layer with 1593 neurons (one per label).

The network was trained by minimizing the categorical cross-entropy between the true labels and the predictions using the Stochastic Gradient Descent (SGD) method, with mini-batches of 100 samples, an initial learning rate of 0.01 decaying by 10^{-6} in each epoch, and a Nesterov momentum of 0.9. After 100 epochs, the network obtained an accuracy of 99.2% on a test set composed of 61945 randomly selected samples (10% of the whole data, not used for training). Finally, the high-level facial features for category learning of Bush and Clinton images were extracted by taking the neural activation prior to the last softmax layer.

Mathematical model description

The neuro-computational model was implemented using the ANNarchy neural simulator (Vitay et al., 2015) version 3.0. The forward Euler method had been used to numerically solve these differential equations with a time step of 1 ms. Figure 4 shows our model' architecture with more detail than in Figure 1 by illustrating the number of cells in each neural population, all the connections in the model, and the type of these connections.

The IT

IT is composed of 100 neurons whose membrane potentials are computed by:

$$\tau_m \cdot \frac{dm_j^{IT}(t)}{dt} + m_j^{IT}(t) = S_j \tag{5}$$

where m_j^{IT} is the membrane potential of the neuron j; $\tau_m = 10$ ms is the time constant; and S_j is the part of the preprocessed image that this neuron receives. The firing rate $r_j^{IT}(t)$ is calculated by applying $()^+$ to the membrane potential, where $()^+$ is a function that takes the positive part of its argument (all negative arguments are transformed to 0)

The BG

The BG model is based on the one by Schroll et al. (2014). We again briefly describe the model and highlight the changes we implemented.

The membrane potential of all cells in the BG is defined by a leaky-integration equation:

$$\tau_m \cdot \frac{dm_j(t)}{dt} + m_j(t) = \sum_{pre \in N_e} I_j^{pre}(t) - \sum_{pre \in N_y} I_j^{pre}(t) + B_j + \epsilon_j(t) \tag{6}$$

where m_j is the membrane potential of neuron j, $\tau_m = 10$ ms the time constant, B_j the baseline of the cell's membrane potential (2.4 for the SNr, 1.0 for the GPe and 0.4 for the other nuclei), $\epsilon_j(t)$ is random noise sampled from a uniform distribution in the interval [-1.0 , 1.0] for the GPe and SNr, and [-0.1 , 0.1] for the other nuclei; I^{pre} the input from the presynaptic neural population to neuron j, N_y the set of presynaptic neural populations with inhibitory synapses to neuron j, and N_e the set of presynaptic neural populations with excitatory synapses to neuron j.

The inputs are computed as:

$$I_j^{Pre}(t) = \sum_{i \in Pre} w_{i,j} \cdot r_i^{Pre} \tag{7}$$

where $w_{i,j}(t)$ is the weight of the synapse between the presynaptic neuron i and the postsynaptic neuron j, and $r_i^{Pre}(t)$ is the firing rate of the presynaptic cell i.

Equation 7 is used to compute the impact of all the connections in our model except for the case of the

SNr lateral connections. The model includes plasticity in connections from the striatum and the STN to the SNr. Although uncommon, this approach gives the model a high level of flexibility as it doesn't force a particular connectivity pattern but lets the network develop it by itself. Unfortunately, except of the striatum, little is known about neural plasticity in the BG. As reviewed by Schroll and Hamker (2013) not only the striatum, but other nuclei are innervated by axons of dopamine neurons. Further, administration of the dopamine precursor levodopa has been shown to affect synaptic plasticity in SNr (Prescott et al., 2008). Dopamine dependent plasticity in our model SNr avoids that striatal cells are hard wired to one category in the SNr — an approach necessary for previous models of BG, which are hard wired from striatum to thalamus. Further, learning requires competition between cells, otherwise all neurons would learn similar features. To implement competition in the SNr, the impact of the SNr laterals is computed by multiplying the synaptic weights by a reversal factor $(1 - r_i^{Pre}(t))^+$:

$$I_j^{Pre}(t) = \sum_{i \in Pre} w_{i,j} \cdot (1 - r_i^{Pre}(t))^+ \cdot r_i^{Pre} \tag{8}$$

where the synaptic weights of the lateral connections in the SNr are excitatory and fixed to 1. There is no direct evidence for our assumed SNr circuitry, mainly due to a lack of studies, but our assumption agrees with data showing that activations of the direct pathway cells in the striatum can elicit both excitation and inhibition of SNr neurons (Freeze et al., 2013; Hikosaka et al., 1993). Lateral connections in the striatum D1 (StrD1), striatum D2 (StrD2) and STN are inhibitory and set to 0.3.

SNc follows an equation that produces the dopamine signal and is the only part of our network that is not governed by equation 6:

$$\tau_m \cdot \frac{dm_j^{SNc}(t)}{dt} + m_j^{SNc}(t) = (1 - R) \cdot (-10 \cdot I_j^{StrD1}(t)) + R \cdot (1 - B_{DA} - I_j^{StrD1}(t))^+ + B_{DA} \tag{9}$$

where $B_{DA} = 0.1$ is the baseline of the cell's membrane potential. $I_j^{StrD1}(t)$ is the impact from the connections of all StrD1 cells to the SNc which learn to represent the reward prediction at the time of the reward delivery. R is a term that changes depending on whether reward is delivered (set to 1) or omitted (set to 0). The dopamine signal is only computed during the reward presentation period and it encodes a reward prediction error at the time of the reward delivery using D1 striatal neurons activity for the prediction, as these cells have been reported to be part of the pathway that project to the dopaminergic neurons, see Vitay and Hamker (2014) for a more detailed model of the reward prediction

error computation.

The firing rate of all cells in our model is calculated by applying $()^+$ to the membrane potential. The learning rule to update the synaptic weights from the IT cells to the 16 StrD1 cells, 16 StrD2 cells and 16 STN cells is:

$$\tau_w \cdot \frac{dw_{i,j}^{IT-POST}(t)}{dt} = f_{DA}(DA(t) - B_{DA}) \cdot C - \alpha_j^{POST}(t) \cdot ((r_j^{POST}(t) - \overline{r}^{POST}(t))^+)^2 \tag{10}$$

with C being the covariance term:

$$C = (r_i^{IT}(t) - \overline{r}_t^{IT}(t) - \gamma_{pre}) \cdot (r_j^{POST}(t) - \overline{r}^{POST}(t))^+ \tag{11}$$

$f_{DA}(x)$ a function that determines how dopamine influences learning (where $T_d = 1$ for cells in the direct and hyperdirect pathway and $T_d = -1$ for cells in the indirect pathway):

$$f_{DA}(x) = \begin{cases} (T_d \cdot 2 \cdot x) & if\ (T_d \cdot x) > 0 \\ (T_d \cdot 0.8 \cdot x) & if\ (T_d \cdot x) < 0 \cup (T_c \cdot C) > 0 \\ 0 & else \end{cases} \tag{12}$$

and α_j the adaptive normalization variable ($T_c = 1$ for excitatory connections and $T_c = -1$ for inhibitory connections):

$$\frac{d\alpha_j^{POST}(t)}{dt} + \alpha_j^{POST}(t) = (T_c \cdot m_j(t) - m^{MAX})^+ \tag{13}$$

Where $\tau_w = 75$ ms is the time constant. Synapses are randomly initialized with a uniform distribution in the interval [0.0 , 0.3].

With dopamine peaks, very active StrD1 and STN cells will strengthen their connections with the active IT cells and weaken their connections with the rest of IT cells. With dopamine dips, the connections between very active StrD1 and STN cells and active IT cells weaken. The dopamine learning effect is reversed in the projections from IT to StrD2 cells. Thus, with dopamine dips, the most active StrD2 cells will strengthen their connections with the active IT cells and weaken their connections with the rest of IT cells. With dopamine peaks, the connections between very active StrD2 cells and active IT cells

weaken.

The covariance term C depends on the following parameters and variables: the firing rate of the postsynaptic cell $r_j^{POST}(t)$; the mean of the firing rates in the postsynaptic layer $\overline{r}^{POST}(t)$; a threshold $\gamma_{pre} = 0.15$; the firing rate of the IT neuron $r_i^{IT}(t)$; and the mean firing rate in the IT layer $\overline{r}^{IT}(t)$. $f_{DA}(x)$ depends on the dopamine level $DA(t)$ and the dopamine baseline $B_{DA} = 0.1$.

The subtractive term of the right hand side of equation 10 serves to saturate the synaptic weights of a cell so that the cell's firing rate is also bound. Equation 13 shows that α_j^{POST} depends on the membrane potential of the postsynaptic cell ($m_j(t)$); and a threshold ($m^{MAX} = 1$).

The learning rule for changing the connection from the StrD1 to the SNr, from StrD2 to GPe cells, and from STN to SNr cells is:

$$\tau_w \cdot \frac{dw_{i,j}^{PRE-POST}(t)}{dt} = f_{DA}(DA(t) - B_{DA}) \cdot (-C) - \alpha_j^{POST}(t) \cdot (-C)^+ \tag{14}$$

with the covariance term:

$$C = T_c \cdot (r_i^{PRE}(t) - \overline{r}_t^{PRE}(t))^+ \cdot (-r_j^{POST}(t) + \overline{r}^{POST}(t) - \gamma_{post}) \tag{15}$$

$f_{DA}(x)$: the variable that determines how dopamine influences learning via Equation 12; and α_j: the adaptive normalization variable computed via Equation 13.

where $\tau_w = 50$ ms is the time constant. Synapses are randomly initialized by values taken from a uniform distribution in the interval [0.0 , 0.05].

The additive term on the left side of equation 14 ensures that during peaks of dopamine, the most active StrD1 cells will strengthen their connections with the less active SNr cell and weaken their connections with the other SNr cell; and the most active STN cells will strengthen their connections with the most active SNr cell and weaken their connections with the other SNr cell. With dopamine dips, the most active StrD1 cells will weaken their connections with the less active SNr cell and the most active STN cells will strengthen their connections with the less active SNr cell.

In the case of the StrD2-GPe projections, the dopamine learning effect is the opposite of the dopamine effect in the StrD1-SNr projections. Then with dopamine dips, the most active StrD2 cells will strengthen their connections with the less active GPe cell and weaken their connections with the other GPe cell. With dopamine peaks, most active StrD2 cells will weaken their connections with the less active GPe cell.

C depends on the following parameters and variables: a threshold $\gamma_{post} = 0.15$, the firing rate of the presynaptic neuron $r_j^{PRE}(t)$, the mean of the firing rates in the presynaptic layer $\overline{r}^{PRE}(t)$, the firing rate of the postsynaptic cell $r_j^{POST}(t)$, and the mean of the firing rates in the postsynaptic layer $\overline{r}^{POST}(t)$.

The threshold of α_j^{POST} is $m^{MAX} = 1$ for the StrD1-SNr connections, $m^{MAX} = 2.6$ for the STN-SNr connections, and $m^{MAX} = 2$ for the StrD2-GPe connections. The SNr and GPe also receive thalamic feedback which is provided by direct connections from the VA to a sub-population of striatal cells (StrThal in Figure 4), that in turn project to both the GPe and the SNr. These projections help to stabilize the BG decision by enhancing the inhibition of the selected category in the SNr. This stabilization allows to reliably notify the BG pathways which category decision should be reinforced when a dopamine peak is generated (Brown et al., 2004). The connections from the StrThal to the SNr and GPe are set to 0.3, from the VA to the StrThal to 1, and the lateral connections in StrThal to 0.3.

The connections from the StrD1 cells to the SNc cell are updated by equations 16 and 17:

$$\tau_w \cdot \frac{dw_{i,j}^{StrD1-SNc}(t)}{dt} = g_{DA} \cdot (DA(t) - B_{DA}) \cdot (r_i^{StrD1}(t) - \overline{r}_t^{StrD1}(t))^+ \tag{16}$$

with

$$g_{DA} = \begin{cases} 1 & if\ reward \\ 3 & if\ no\ reward \end{cases} \tag{17}$$

where $\tau_w = 100000$ ms is the time constant; $r_j^{StrD1}(t)$ is the firing rate of the neuron j in the StrD1 layer; $\overline{r}^{StrD1}(t)$ is the mean of the firing rates in the StrD1 layer; g_{DA} is a parameter that scales the effect of dopamine dips and peaks in learning; $DA(t)$ is the dopamine level; and B_{DA} is the baseline of the dopamine level. Consequently, peaks in dopamine will strengthen the connections between the SNc cell

and the most active StrD1 cells and dips will weaken these connections.

Finally, we summarize the difference between our model and the one by Schroll et al. (2014). As our examples required only two categories to learn, our SNr and GPe are composed of two cells instead of four. As a result, the synaptic values of the SNr lateral connections are fixed to 1.0 instead of being plastic. The synaptic values of the plastic connections in the model are randomly initialized from a uniform distribution; instead, Schroll et al. (2014) initialized these synaptic weights to zero. The projections from the IT to the BG input nuclei are only excitatory in our model. In contrast, Schroll et al. (2014) allowed these synaptic weights to switch their character between excitatory and inhibitory during learning. Learning in the present model does not rely on calcium traces implemented in the previous model as they are not required for the purpose of this study. The learning rules from the IT to the BG input nuclei have been slightly changed (in the subtractive term of the learning rules). Finally, the time constant of the IT membrane potential, the m^{MAX} for the STN-SNr connections and the fixed weights of w^{SNr-VA} were modified.

The cortico-thalamic architecture

The membrane potential m_j of the 2 VA and the 16 PFC cells is computed by the equations 6 and 7, with a time constant of 10 ms; the random noise is generated from a uniform distribution in the interval [-0.05 , 0.05] for the PFC and in [-0.0001 , 0.0001] for the VA; and the baseline is 0 for both populations. The firing rate is calculated by applying $()^+$ to the membrane potential.

The connectivity between PFC and VA is fixed and ensures that a PFC cell can only obtain its input from a single VA cell to avoid any overlapp. The number of PFC cells connected to a VA cell is balanced equally. The weight values are defined as follows: w^{VA-PFC}, w^{PFC-VA} and $w^{PFC-PFC}$ are fixed with values 0.35, 0.15, and 0.1 respectively. w^{IT-PFC} are randomly initialized with a uniform distribution in the interval [0.2 , 0.4] and modified by the following learning rule:

$$\tau_w \cdot \frac{dw_{i,j}^{IT-PFC}(t)}{dt} = (r_i^{IT}(t) - \overline{r}_t^{IT}(t) - \gamma_{pre}) \cdot (r_j^{PFC}(t) - \overline{r}^{PFC}(t))^+ \\ -\alpha_j^{PFC}(t) \cdot ((r_j^{PFC}(t) - \overline{r}^{PFC}(t))^+)^2 \cdot w_{i,j}^{IT-PFC}(t) \quad (18)$$

where $\tau_w = 15000$ ms; $\gamma_{pre} = 0.15$, and $\alpha_j^{PFC}(t)$ is the variable that contributes to the dynamic synaptic

saturation (eq. 13), with threshold $m^{MAX} = 3.5$.

In the most active PFC cells, the synapse will be strengthened if the presynaptic cell's firing rate is above the population mean and will be weakened otherwise. The subtractive term on the right side of equation 18 ensures dynamic synaptic saturation as in the Oja's learning rule (Oja, 1982).

Variation of 64 model parameters

Each of the 100000 simulations was performed with different values for 64 model parameters. The value of each of these parameters was randomly selected from an uniform distribution in the interval between plus/minus 10% of the parameter's value previously specified in the mathematical model description. The 64 model parameters are: the membrane potential's baseline for the different neural populations, the membrane potential's noise for the different populations, the time constant of the different learning rules, the m^{MAX} of each learning rule, the γ_{pre} of each learning rule, the scaling factor for dopamine peaks and the scaling factor for dopamine dips in the $f_{DA}(x)$ of each projection, the scaling factor for the reward prediction signal when reward is not delivered, the value of g_{DA} when reward is not delivered, and the synaptic weights of the different fixed connections.

Results

Simulation Results

To meaningfully compare our model's results with physiological and biological data and, at the same time, test the robustness of our model, we ran 100000 category learning experiments each with randomly generated initial synaptic weights and with randomly generated values for 64 model parameters. Each of these parameters' values was randomly determined from a uniform distribution in an interval between plus/minus 10% of its value specified in the method's section (base value). With a total of 100000 of these experimental runs, we consider a large number of variations for the 64 model value parameters. Further, we used some variability in the learning task by choosing for each experimental run a different set of stimuli among 100 possible sets of stimuli, each generated from two different category prototypes.

An experiment was considered successful when, within 65 trials per block, 16 out of 20 consecutive trials were correct in each block. The model successfully executed 82639 out of 100000 experiments

(82.639%), a proportion slightly better than that of monkeys (Antzoulatos and Miller, 2011): 19 out of 24 (79.166%). Further, the model showed a similar learning performance across the paradigm than that of monkeys (Figure 5): initially, the model randomly selected a category (50% correct performance); the performance gradually improved from the first block to the fourth block; and from the fifth block on, the performance saturated at around 96.5%.

The Pearson correlation coefficient (PCC) between each of the 64 parameters and the model's performance (computed for each experimental run as the percentage of correct trials in the 16 last trials) is very low: between -0.035 and 0.041, indicating that the model tolerates modest changes in any of the specified model value parameters. When the PCC considered the absolute values of the perturbations produced in each parameter base value instead of the values of each parameter, correlations are even smaller: between -0.01 and 0.01.

Importantly, our model reproduced the key neurophysiological findings of Antzoulatos and Miller (2011): at the beginning of the paradigm, striatal cells were strongly category selective and PFC cells were weakly category selective, while later on, PFC cells became highly category selective and striatal cells weakly category selective (Figure 6).

In the following, we use the model as a tool to better understand this key finding. When we analyse each cell's category and stimulus selectivity over 100 simulations (with fixed model parameters) we see that PFC and striatal cells show a different selectivity profile (Figure 7). Throughout the paradigm, there were striatal cells that were stimulus selective and striatal cells that were category selective, indicating that striatal cells encode both, specific and abstract knowledge. Importantly, this result shows that, although the striatum d' sensitivity is reduced late in the experiment, there are striatal cells involved in category learning throughout the whole experiment. PFC cells, in contrast, increased their category selectivity across blocks while their stimulus selectivity remained low throughout the paradigm, supporting that these cells encode generalized, categorical knowledge.

Three example striatal cells illustrate different response characteristics to stimuli of both categories (Figure 8). The first cell exclusively responds to stimuli of one category throughout the experiment, but from block IV onwards, it does not respond to all stimuli of its preferred category. Thus, its category responses become more variable within the set of stimuli of the preferred category. The second cell switches its category selectivity. Furthermore, the variability of this cell's category response is higher in the last blocks than in the first blocks. A third cell responds to stimuli of one category in the first blocks, but

becomes selective to stimuli of the other category in the later blocks as well. Therefore, this cell loses its category selectivity and appears to become selective to input patterns common to both categories.

When we analyse the response characteristics across all cells, we observe that the distance between the mean response to the preferred category μ_P and the mean response to the non-preferred category μ_N reduces after the first phase, due to a reduction in μ_P and a small increase in μ_N (Figure 9a). However, the mean response to the preferred category stays much higher than the one to the non-preferred category in all three different phases of the experiment, showing that striatal cells have on average a clearly preferred category throughout the experiment. Thus, a cell responding to both categories (see Figure 8) is not the typical case.

The increase in the standard deviation of the response to the preferred category σ_P and the standard deviation of the response to the non-preferred category σ_N (Figure 9b) confirms our observations of the example cells in that the striatal response to category information becomes more variable after the first phase of the experiment.

As these results suggest that the decrease in $\mu_P - \mu_N$ is due to an increase in the variability of the category response, our model predicts that the decrease of the d' sensitivity index is primarily the result of an increase in the variability of the category response.

We next explored why the decrease in striatal category selectivity and the accompanying increase in variability occurred. As a first hypothesis, we reasoned that - as PFC category selectivity increased with learning - striatal category selectivity became less required for successful task performance and was therefore unlearned as the neural activity in the striatum may not be the cause of the final decision. To test this hypothesis, we ran 100 additional simulations with our model, but we now blocked learning in the PFC so that the BG were performing the experiment alone. However, the striatal d' sensitivity index abruptly decreased after the first phase and stayed at a low level in the next two phases, qualitatively very similar to the full model, therefore, ruling out that the decrease in striatal category selectivity occurs due to a PFC dominance in later blocks.

As another hypothesis, we tested whether, as task performance increased, dopamine peaks (i.e., positive reward prediction errors) in the model stopped appearing - which would have impaired further learning in the striatum. However, dopamine peaks were only reduced to 43% on average, enough to still produce large synaptic changes in the striatum.

Next, we tested whether the increase of the variability in the striatal category response and therefore the decrease of the striatal category selectivity is produced by the learning of a large diversity of stimuli. To test this idea, we ran 100 simulations with the full model performing a new prototype distortion task in which the diversity of exposed stimuli is large and constant from the beginning of the task. Rather than subdividing the prototype distortion task into blocks with increasing numbers of stimuli across blocks, any stimulus from the whole repertoire of stimuli available per experiment could be presented in each trial. We now observe a low striatal category selectivity from the beginning of the experiment (Figure 10) and no drop in the selectivity index. Since PFC category selectivity rises to high values, the BG still teach PFC cells to develop category representations, indicating that the BG are involved in the categorization task. Consequently, this result supports that the decrease in the d' sensitivity index is due to the fast learning of a large diversity of exposed stimuli, which makes it impossible for the striatal cells to acquire complete category representations and to respond to all stimuli of the preferred category. Thus, the fact that the d' sensitivity index in this revised experiment is low from the beginning, discards a PFC dominance in the category decision and an omission of dopamine peaks as reasons for the low striatal category selectivity, since both effects occur later in the experiment.

To further explore BG and PFC interactions, we compared the performance of the full model with the performance of the BG and the PFC alone in a slightly more challenging, real world category learning task. In each of the 2500 trials of this task, an image randomly selected from 1671 images of Bill Clinton and George W. Bush (extracted from Youtube videos and therefore varying in perspective and facial expressions) was presented to the model for classification purposes. In order to mimic early visual perception up to area IT, we used a convolutional network trained on other faces to transform each image into a 100 dimensional vector representing the high-level facial features of the corresponding image (see Methods). In order to adapt the model to these new inputs, two small changes were implemented. First, slower learning in the PFC ($\tau_w = 100000$ ms) was required to guarantee that the common patterns among inputs of a category were extracted. Secondly, the pre-synaptic threshold in the PFC learning rule was set to γ_{pre}=0.0 to ensure that all relevant features of the input space were learned. Surprisingly, the BG model alone achieves a much weaker performance than the full model and the PFC alone (Figure 11). The BG and the full model very soon reached 85%. While the full model slowly improved its performance, finally achieving a level of 97.4%, the BG alone lack further improvement and their performance fluctuates around 85%. The performance of the PFC alone, with a 95.6% of correct responses after the training of the full model, is only 1.8% lower than the full model's performance. Thus, with a difference in

performance between the full model and the BG alone of around 12.2% and a difference in performance between the PFC alone and the BG alone of around 10.4%, we can corroborate the relevant role of the PFC in pushing the categorization performance to high levels with complex input stimuli. We have also found that Hebbian learning alone is not enough to reach a high performance on the task.

To ensure that the slow learning in the PFC is key for pushing the categorization performance of the full model above the BG categorization performance, we compared the categorization performance of five full model configurations each with different learning speed in the PFC. Categorization performance was here evaluated across 100 runs on the category learning task of faces. Figure 12 illustrates that the full model's performance increases as the speed in the PFC learning rule decreases, confirming the principal role of the slow learning in the PFC for achieving high categorization performance.

To study if the slow learning alone can be sufficient, we ran the BG alone with slow learning in an additional region of the striatum. However, the BG with slow learning reach a significantly lower performance (around 17.4% lower) than the full model (Figure 13). Also, to study the effect of additional basal-ganglio-cortical connections seen in vivo, we added to our model the well-known connection from the PFC to the STR (Ferry et al., 2000). Figure 13 shows that this version of the model does not alter the performance previously achieved by our full model, indicating that this connection does not have a relevant effect on the simulated task.

Physiological results

Even though the striatum d' sensitivity index abruptly decreases after the first phase, our model predicts that on average, the striatal cells remain selective for a preferred category, but their response to category information becomes more variable (Figure 9). To empirically test these predictions, we went back to the original monkey data. In accordance with our simulation results, the monkey data shows that the mean response to the preferred category μ_P weakly decreases after the first phase and that the mean response to the non-preferred category μ_N slightly increases after the first phase (Figure 14a). Most importantly, μ_P is significantly higher than μ_N in the three phases, confirming that although the striatum d' sensitivity index shows a large decrease, the striatal cells show a preferred categorical response in the course of the experiment.

In agreement with our model, the monkey data shows that both the variability of the striatal response

to the preferred category σ_P and the variability of the striatal response to the non-preferred category σ_N increases after the first phase, while σ_P is higher than σ_N (Figure 14b).

Discussion

977 words out of 1500.

We introduced a new neuro-computational model of category learning to investigate interactions of basal ganglia and PFC. While it is known that the PFC receives a dopaminergic input, though less than the striatum (Seger and Miller, 2010), its phasic properties are less pronounced due to the very slow decay of DA in PFC as reviewed by Lapish et al. (2007). Lapish et al. (2007) suggested that the co-release of glutamate from DA neurons may serve as a temporally precise signal to allow PFC neurons to switch between different modes that affect local network dynamics. However, how such DA dependent states may subserve reinforcement learning of cortico-cortical connections has not been discussed. Thus, we took the conservative assumption that cortico-cortical connections follow a Hebbian learning rule. Hebbian learning however, is insufficient for categorizing complex input stimuli at high performance levels. Therefore, the PFC requires teaching signals to guide learning towards useful representations at an intermediate level between perception and action. We here explored the hypothesis that the BG modulate the cortico-thalamo-cortical loop and thus provide the PFC with the required task related information. Unlike the mesocortical DAergic signal, the BG teaching signal provides no reward related information to the PFC. It supplies the PFC with a desired response for the current input as estimated by the BG. Interestingly, the 'teacher' can display a much lower performance than the 'student' (Figure 11). Due to the slow learning of the cortico-cortical connections (here IT-PFC), occasionally wrong decisions transferred by the BG are tolerated. The BG with 85% correct performance still teach the PFC to push the model's performance up to almost 100%. Potential benefits of combined fast and slow learners have been laid out in the context of memory consolidation based on models of the cortex and hippocampus (McClelland et al., 1995; O'Reilly and Rudy, 2000), but although the main ideas are intuitive, clear model demonstrations were rare since then. Our simulation results underline these previous ideas, here with respect to basal ganglia - cortex interactions, and clearly demonstrate the additional advantage of a slow learning system which complements fast learners, required for survival.

As our model replicates key behavioral and physiological data of macaque monkeys performing a prototype

distortion task (Antzoulatos and Miller, 2011), our model provides some confidence to allow us to further delineate the potential mechanistic causes behind the observation that striatal neural activity is initially a good predictor of stimulus category, while this category selectivity declines as the number of stimuli to classify increases. Our simulations suggest that the drop in striatal category selectivity does not relate to a disengagement of the striatum in category learning: the striatal cells' response to category information exhibits on average a strong preference to one category throughout the whole experiment, indicating that striatal cells can acquire category knowledge when learning to classify a large number of stimuli. Importantly, our model predicts that the decrease in the category selectivity of striatal cells occurs due to an increase in the variability of the category response. We tested and confirmed this model prediction by re-analysing the original monkey data obtained by Antzoulatos and Miller (2011). The large number of simulations (100000) each performed with different model value parameters, assures that the results are robust to modest changes in these parameters. The model did not show significant susceptibility to changes in any of these parameters.

Our results may advance the field of computational models of category learning, which has already a long tradition. Category learning models with a more psychological focus typically tend to abstract from details of brain computation and focus mainly on a replication of behavioral data in different category learning tasks such as prototype, probabilistic, rule-based and information-integration categorization tasks (for a review see Richler and Palmeri (2014)). Most recent neuro-computational models of category learning focus on the role of the ventral visual pathway, but typically simplify at the level of category decision by relying on mechanisms of supervised learning to link feature and object information to categories (Serre, 2016). However, Cantwell et al. (2017) also emphasized the role of the basal ganglia in category learning by merging a model of visual object processing with a model of procedural-learning based on the direct pathway of the basal ganglia. While this is interesting, their model has been directed to learn correct stimulus–response associations and to match the performance of human behavioural data in category learning, it did not focus on the formation of category representations in PFC and likewise has not been used to explain electrophysiological data such as those from the study of Antzoulatos and Miller (2011). Our model demonstrates that the PFC can be trained by the BG to develop useful internal representations for completing a prototype distortion and a simple, but real-world face categorization task. As this learning is generic it may also provide the basis of other category learning tasks, although some of them may recruit additional or slightly different brain areas (Seger and Miller, 2010; Richler and Palmeri, 2014) and may therefore require more complex models.

As the BG form multiple loops with most parts of cortex (Alexander et al., 1986; Haber, 2003) our model could provide an inspiration for the organization of other loops as well. A further challenging and unanswered question is how different cortical areas connect with each other across loops. The organization of connections in early sensory areas may be approximately well explained by Hebbian learning. Cortico-cortical areas further downstream likely require error signals and fast learning BG circuits to bias cortex in a meaningful way so that brain circuits self-organize to find solutions that allow the organism to survive, reproduce and evolve. Hélie et al. (2015) already suggested that the BG are required for learning such cortico-cortical associations. Our study provides an example of how this may actually work and may offer a blueprint for the organization of other cortico-cortical associations.

References

Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357–381.

Antzoulatos EG, Miller EK (2011) Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. Neuron 71:243–249.

Antzoulatos EG, Miller EK (2016) Synchronous beta rhythms of frontoparietal networks support only behaviorally relevant representations. eLife 5:e17822.

Apicella P, Scarnati E, Ljungberg T, Schultz W (1992) Neuronal activity in monkey striatum related to the expectation of predictable environmental events. J Neurophysiol 68:945–960.

Braak H, Del Tredici K (2008) Cortico-basal ganglia-cortical circuitry in parkinson's disease reconsidered. Exp Neurol 212:226–229.

Brown JW, Bullock D, Grossberg S (2004) How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. Neural Netw 17:471–510.

Cantwell G, Riesenhuber M, Roeder JL, Ashby FG (2017) Perceptual category learning and visual processing: An exercise in computational cognitive neuroscience. Neural Netw 89:31–38.

Cincotta CM, Seger CA (2007) Dissociation between striatal regions while learning to categorize via feedback and via observation. J Cogn Neurosci 19:249–265.

626 Ferry AT, Öngür D, An X, Price JL (2000) Prefrontal cortical projections to the striatum in macaque
627 monkeys: evidence for an organization related to prefrontal networks. J Comp Neurol 425:447–470.

628 Fisher SD, Robertson PB, Black MJ, Redgrave P, Sagar MA, Abraham WC, Reynolds JN (2017) Rein-
629 forcement determines the timing dependence of corticostriatal synaptic plasticity in vivo. Nat. Com-
630 mun. 8:1–13.

631 Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2001) Categorical representation of visual stimuli
632 in the primate prefrontal cortex. Science 291:312–316.

633 Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2002) Visual categorization and the primate pre-
634 frontal cortex: neurophysiology and behavior. J Neurophysiol 88:929–941.

635 Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and
636 inferior temporal cortices during visual categorization. J Neurosci 23:5235–5246.

637 Freeze BS, Kravitz AV, Hammack N, Berke JD, Kreitzer AC (2013) Control of basal ganglia output by
638 direct and indirect pathway projection neurons. J Neurosci 33:18531–18539.

639 Haber SN (2003) The primate basal ganglia: parallel and integrative networks. J Chem Neu-
640 roanat 26:317–330.

641 Hélie S, Ell SW, Ashby FG (2015) Learning robust cortico-cortical associations with the basal ganglia:
642 An integrative review. Cortex 64:123–135.

643 Hikosaka O, Sakamoto M, Miyashita N (1993) Effects of caudate nucleus stimulation on substantia nigra
644 cell activity in monkey. Exp Brain Res 95:457–472.

645 Koch G, Ponzo V, Di Lorenzo F, Caltagirone C, Veniero D (2013) Apomorphine and dopamine D1
646 receptor agonists increase the firing rates of subthalamic nucleus neurons. J Neurosci 33:9725–9733.

647 Kreiss D, Anderson L, Walters J (1996) Apomorphine and dopamine D1 receptor agonists increase the
648 firing rates of subthalamic nucleus neurons. Neuroscience 72:863–876.

649 Lapish CC, Kroener S, Durstewitz D, Lavin A, Seamans JK (2007) The ability of the mesocortical
650 dopamine system to operate in distinct temporal modes. Psychopharmacology (Berl) 191:609–625.

651 McClelland JL, McNaughton BL, O'reilly RC (1995) Why there are complementary learning systems in
652 the hippocampus and neocortex: insights from the successes and failures of connectionist models of
653 learning and memory. Psychol Rev 102:419–457.

654 Merchant H, Zainos A, Hernández A, Salinas E, Romo R (1997) Functional properties of primate putamen
655 neurons during the categorization of tactile stimuli. J Neurophysiol 77:1132–1154.

656 Miller E, Buschman T (2007) Rules through recursion: how interactions between the frontal cortex and
657 basal ganglia may build abstract, complex rules from concrete, simple ones. In: The Neuroscience of
658 Rule-Guided Behavior (Bunge S, Wallis J eds), pp419–440. Oxford:Oxford UP.

659 Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. Prog
660 Neurobiol 50:381–425.

661 Muhammad R, Wallis JD, Miller EK (2006) A comparison of abstract rules in the prefrontal cortex,
662 premotor cortex, inferior temporal cortex, and striatum. J Cogn Neurosci 18:974–989.

663 Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal 'hy-
664 perdirect' pathway. Neurosci Res 43:111–117.

665 Nomura E, Maddox W, Filoteo J, Ing A, Gitelman D, Parrish T, Mesulam M, Reber P (2007) Neural
666 correlates of rule-based and information-integration visual category learning. Cereb Cortex 17:37–43.

667 Oja E (1982) Simplified neuron model as a principal component analyzer. J Math Biol 15:267–273.

668 O'Reilly RC, Frank MJ (2006) Making working memory work: a computational model of learning in the
669 prefrontal cortex and basal ganglia. Neural Comput 18:283–328.

670 O'Reilly RC, Rudy JW (2000) Computational principles of learning in the neocortex and hippocampus.
671 Hippocampus 10:389–397.

672 Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex
673 and striatum. Nature 433:873–876.

674 Poldrack RA, Prabhakaran V, Seger CA, Gabrieli JD (1999) Striatal activation during acquisition of a
675 cognitive skill. Neuropsychology 13:564–574.

676 Poldrack RA, Clark J, Pare-Blagoev E, Shohamy D, Moyano JC, Myers C, Gluck M (2001) Interactive
677 memory systems in the human brain. Nature 414:546–550.

678 Prescott I, Dostrovsky J, Moro E, Hodaie M, Lozano A, Hutchison W (2008) Levodopa enhances synaptic
679 plasticity in the substantia nigra pars reticulata of parkinson's disease patients. Brain 132:309–318.

680 Richler JJ, Palmeri TJ (2014) Visual category learning. Wiley Interdiscip Rev Cogn Sci 5:75–94.

Schroll H, Beste C, Hamker FH (2015) Combined lesions of direct and indirect basal ganglia pathways but not changes in dopamine levels explain learning deficits in patients with huntington's disease. Eur J Neurosci 41:1227–1244.

Schroll H, Hamker FH (2013) Computational models of basal-ganglia pathway functions: focus on functional neuroanatomy. Front Syst Neurosci 7:1–18.

Schroll H, Vitay J, Hamker FH (2012) Working memory and response selection: a computational account of interactions among cortico-basalganglio-thalamic loops. Neural Netw 26:59–74.

Schroll H, Vitay J, Hamker FH (2014) Dysfunctional and compensatory synaptic plasticity in parkinson's disease. Eur J Neurosci 39:688–702.

Seger CA, Cincotta CM (2005) The roles of the caudate nucleus in human classification learning. J Neurosci 25:2941–2951.

Seger CA, Miller EK (2010) Category learning in the brain. Annu Rev Neurosci 33:203.

Serre T (2016) Models of visual categorization. Wiley Interdiscip Rev Cogn Sci 7:197–213.

Shen W, Flajolet M, Greengard P, Surmeier DJ (2008) Dichotomous dopaminergic control of striatal synaptic plasticity. Science 321:848–851.

Sjöström, PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron 32:1149–1164.

Stanford TR, Shankar S, Massoglia DP, Costello MG, Salinas E (2010) Perceptual decision making in less than 30 milliseconds. Nat Neurosci 13:379–385.

Surmeier DJ, Ding J, Day M, Wang Z, Shen W (2007) D1 and d2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. Trends Neurosci 30:228–235.

Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vis 57:137–154.

Vitay J, Dinkelbach HÜ, Hamker FH (2015) Annarchy: a code generation approach to neural simulations on parallel hardware. Front Neuroinform 9:1–20.

Vitay J, Hamker FH (2014) Timing and expectation of reward: a neuro-computational model of the afferents to the ventral tegmental area. Front Neurorobot 8:1–25.

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953–956.

Wallis JD, Miller EK (2003) From rule to response: neuronal processes in the premotor and prefrontal cortex. J Neurophysiol 90:1790–1806.

Wolf L, Hassner T, Maoz I (2011) Face recognition in unconstrained videos with matched background similarity. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pp529–534. Piscataway,New Jersey:IEEE.

Zeithamova D, Maddox WT, Schnyer DM (2008) Dissociable prototype learning systems: evidence from brain imaging and behavior. J Neurosci 28:13194–13201.

Figure Legends

Figure 1. Outline of the components of the neuro-computional network to train the cortico-cortical, IT-PFC connection by the basal ganglia (BG). All adaptive connections are displayed in green color. While the IT-PFC connections are updated by Hebbian learning, the BG learn based on a three factor learning rule including a reward prediction error signal (DA). We propose that the BG bias the activity of PFC neurons which allow the PFC to learn a categorical representation. IT: inferior temporal cortex. PFC: prefrontal cortex. Striatum D1 and Striatum D2: striatum cells expressing D1 and D2 dopamine receptors, respectively. STN: subthalamic nucleus. SNr: substantia nigra pars reticulata. GPe: external globus pallidus. VA: ventral anterior nucleus of the thalamus.

Figure 2. Image examples and block description of the prototype distortion task. a) Dot prototype stimuli of two categories, taken from one experimental run. b) Number and type of stimuli per block. Stimuli are distinguished according to whether they are added in the previous block or in the current block. c) Stimuli of the second block, derived from the prototypes.

Figure 3. Examples of face stimuli presented to the model. The upper-row pictures correspond to the category of Bill Clinton, the lower-row pictures to the category of George W. Busch. Each picture shows a face with a particular expression and from a different perspective. Each greyscale image has a size of 100x100 pixels and was obtained by applying a Viola-Jones filter to a particular frame of a Youtube video.

Figure 4. Detailed outline of the components of the neuro-computional network. The number of cells that each neural population has is shown at the left-bottom corner of each population box. The reward prediction error signal used by the BG to learn is generated at SNc. SNc: substantia nigra pars compacta. StrThal: striatum with thalamic afferents. The rest of model nuclei are already specified in Figure 1 as follows. IT: inferior temporal cortex. PFC: prefrontal cortex. Striatum D1 and Striatum D2: striatum cells expressing D1 and D2 dopamine receptors, respectively. STN: subthalamic nucleus. SNr: substantia nigra pars reticulata. GPe: external globus pallidus. VA: ventral anterior nucleus of the thalamus.

Figure 5. Model performance across blocks for categorizing novel stimuli averaged from 82639 experiments. Since each block had a minimum of 16 trials (due to the criterion to succeed in a block), we analyzed only the first 16 trials per block. Applying a sliding three-trial window, we then measured the percentage of correct trials for each relevant trial across the all successful experiments (black line) and the corresponding standard error of the mean (SEM). The obtained SEMs are too small to be shown in this plot (smaller than 0.002) due to the large number of experiments considered in this analysis.

Figure 6. Mean d' sensitivity index for category selectivity of neurons in the model's PFC and striatum. Horizontal and vertical axes refer to time and trials, respectively. The first phase represents the first two blocks, the second phase the next two blocks, and the last phase the last four blocks. Only successful experiments (82639 experiments) and successful trials of novel stimuli were considered in this analysis. The analyzed data spread across a time interval spanning from cue onset to reward onset. Each phase includes only its first 16 trials (i.e. 7 trial windows). Different values for the 64 model parameters were set at the beginning of each experimental run.

Figure 7. Category and stimulus selectivity for each model cell in the striatum (red dots) and in the PFC (blue dots) in all successful experiments. Both selectivities were measured at the end of each block as outlined in the Methods section. The first block was omitted because there was only one stimulus per category and, therefore, stimulus selectivity could not be computed. A maximum category selectivity of 1 indicates that the corresponding cell responds maximally to all stimuli of one category and becomes inactive for the stimuli from the other category. Maximum stimulus selectivity, in contrast, indicates that the corresponding cell responds to a single stimulus with maximum activity, but that it remains inactive for the rest of the stimuli from the same category. While the category selectivity in the PFC

clearly increases with each block the category selectivity in the striatum does not and cells stay stimulus selective. The mean category and the mean stimulus selectivity of the PFC and STR cells is shown at each axis by a blue and red triangle respectively. The error bars indicate the standard deviation.

Figure 8. Firing rates of three typical striatal cells plotted across all trials of one experimental run - subdivided for presentation of stimuli from category A (left subplots) and category B (right subplots). The seven vertical lines indicate block borders. The 64 model parameters were set to their base values.

Figure 9. Model prediction with respect to the individual components of the d' sensitivity index of the striatum. The values are computed at the last trial-time bin of each phase and from the same simulation recordings used to obtain the STR d' sensitivity index in Figure 6. a) Mean response to the preferred category (dots in magenta line, μ_P) and to the non-preferred category (dots in green line, μ_N) per phase. The preferred category is the category for which each STR cell responds on average most strongly in each single trial-time window. b) Mean standard deviation of the response to the preferred category (dots in magenta line, σ_P) and to the non-preferred category (dots in green line, σ_N) per phase.

Figure 10. Averaged d' sensitivity index across trials and time for the PFC and the striatal activities recorded in a prototype distortion task without blocks. Horizontal and vertical axes refer to trials and time, respectively. The subplots on the left and right hand side belong to PFC and striatal recordings, respectively. Only successful experiments and successful trials of novel stimuli were considered in this analysis. The analyzed data spread across a time interval spanning from cue onset to reward onset. The 64 model parameters were set to their base values in all considered simulations.

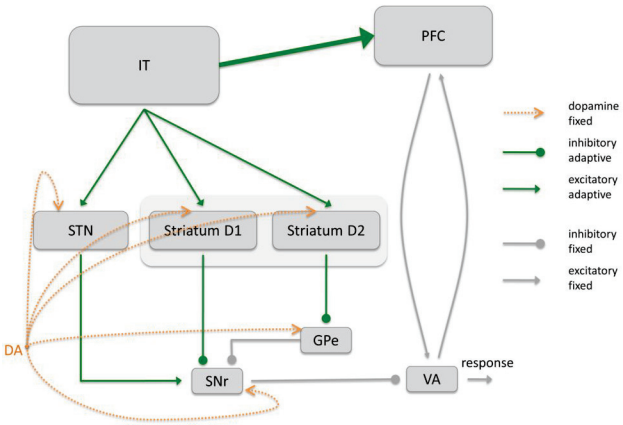
Figure 11. Across-trial performance of the full model (dark red line) and the BG-alone model (dark blue line) across the 2500 trials of the task with real-world face stimuli. For each of the 2500 trials, performance was averaged across 100 experimental runs. Moreover, standard errors of the mean (SEM) were computed (filling color around the lines). A 25-trial window was employed to smooth the plot. The black dot in the final trial represents the mean performance of the PFC alone in an extra set of 1000 trials, performed at the end of the 2500 trials, across 100 experimental runs. The corresponding SEM is 0.00065, too small for being shown in the plot as the dot's error bar. The 64 model parameters were set to their base values in all considered simulations.

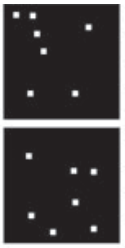
Figure 12. Effect of different PFC learning speeds on the full model's performance in the learning task

with real-world faces images. a) τ_{STR} is the time constant of the learning rule in the striatum and was equal to 75 ms. τ_{PFC} is the time constant of the learning rule in the PFC and its different values here studied were 75, 975, 33375, 66675, and 100000 ms. Each red dot shows the mean performance across 100 experimental runs at the last trial (trial 2500). The error bars show the corresponding SEMs. b) example (for each different condition) of PFC weight trajectory (red line) and STR weight trajectory (blue line) across the trials of one experimental run. The weight value was normalized via dividing by the maximum weight value of the recorded PFC and STR cells. In all considered simulations, the 64 model parameters were set to their base values except for the learning rule time constant in the PFC, which was set to the value specified by each condition.

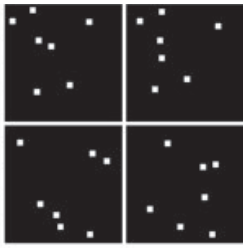
Figure 13. Across-trial performance of the full model (dark red line), the full model with an extra connection from the PFC to the STR (dark blue line), and a model with the slow learning in the STR instead of the PFC (gold line) across the 2500 trials of the task with real-world face stimuli. For each of the 2500 trials, performance was averaged across 100 experimental runs. SEMs are also shown (filling color around the lines). A 25-trial window was employed to smooth the plot. The 64 model parameters were set to their base values in all considered simulations. The small model sketch summarizes the main difference between the three models.

Figure 14. Analysis of the individual components of the striatum d' sensitivity index computed from the monkeys' recordings obtained in the cue period. For each phase, the last trial-time bin is plotted. a) Mean response to the preferred category (dots in magenta line, μ_P) and to the non-preferred category (dots in green line, μ_N) per phase. The preferred category is the category for which each STR cell responds on average most strongly in each single trial-time window. Because neural spiking activity tended to rise with time of recording, we had to correct the neural firing rate to the pre-trial baseline, hence, the negative average firing rate for the non-preferred category. b) Mean standard deviation of the response to the preferred category (dots in magenta line, σ_P) and to the non-preferred category (dots in green line, σ_N) per phase.





BLOCK \ STIMULI	PREVIOUSLY ADDED	CURRENTLY ADDED	TOTAL
I	0	2	2
II	2	2	4
III	2	6	8
IV	6	10	16
V	10	22	32
VI	22	42	64
VII	42	86	128
VIII	86	170	256



(a) (b) (c)