

RESEARCH ARTICLE

Basal ganglia role in learning rewarded actions and executing previously learned choices: Healthy and diseased states

Garrett Mulcahy¹, Brady Atwood^{2,3}, Alexey Kuznetsov^{3,4*}

1 Department of Mathematics, Purdue University, West Lafayette, Indiana, United States of America,

2 Departments of Psychiatry and Pharmacology & Toxicology, IUSM, Indianapolis, Indiana, United States of America, **3** Indiana Alcohol Research Center, IUSM, Indianapolis, Indiana, United States of America,

4 Department of Mathematical Sciences, IUPUI, Indianapolis, Indiana, United States of America

* askuznet@iupui.edu



Abstract

The basal ganglia (BG) is a collection of nuclei located deep beneath the cerebral cortex that is involved in learning and selection of rewarded actions. Here, we analyzed BG mechanisms that enable these functions. We implemented a rate model of a BG-thalamo-cortical loop and simulated its performance in a standard action selection task. We have shown that potentiation of corticostriatal synapses enables learning of a rewarded option. However, these synapses became redundant later as direct connections between prefrontal and premotor cortices (PFC-PMC) were potentiated by Hebbian learning. After we switched the reward to the previously unrewarded option (reversal), the BG was again responsible for switching to the new option. Due to the potentiated direct cortical connections, the system was biased to the previously rewarded choice, and establishing the new choice required a greater number of trials. Guided by physiological research, we then modified our model to reproduce pathological states of mild Parkinson's and Huntington's diseases. We found that in the Parkinsonian state PMC activity levels become extremely variable, which is caused by oscillations arising in the BG-thalamo-cortical loop. The model reproduced severe impairment of learning and predicted that this is caused by these oscillations as well as a reduced reward prediction signal. In the Huntington state, the potentiation of the PFC-PMC connections produced better learning, but altered BG output disrupted expression of the rewarded choices. This resulted in random switching between rewarded and unrewarded choices resembling an exploratory phase that never ended. Along with other computational studies, our results further reconcile the apparent contradiction between the critical involvement of the BG in execution of previously learned actions and yet no impairment of these actions after BG output is ablated by lesions or deep brain stimulation. We predict that the cortico-BG-thalamo-cortical loop conforms to previously learned choice in healthy conditions, but impedes those choices in disease states.

OPEN ACCESS

Citation: Mulcahy G, Atwood B, Kuznetsov A (2020) Basal ganglia role in learning rewarded actions and executing previously learned choices: Healthy and diseased states. *PLoS ONE* 15(2): e0228081. <https://doi.org/10.1371/journal.pone.0228081>

Editor: Gennady Cymbalyuk, Georgia State University, UNITED STATES

Received: May 8, 2019

Accepted: January 7, 2020

Published: February 10, 2020

Copyright: © 2020 Mulcahy et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The model is available in the ModelDB database <http://modeldb.yale.edu/261616>.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The basal ganglia (BG) is an evolutionarily conserved complex network of excitatory and inhibitory neurons located in the deep brain of vertebrates that controls action selection (see e.g. [1]). The BG is comprised of the dorsal striatum, external and internal portions of the globus pallidus (GPe, GPi), subthalamic nucleus (STN) and substantia nigra [2]. It is traditionally implicated in motor control since BG lesions are associated with movement disorders [3,4]. The BG is a shared processing center involved in a broad spectrum of motor and cognitive control [2]. A cortico-BG-thalamo-cortical neurocircuit loop is suggested to be the structure that provides this control [2,5]. However, understanding how this loop functions remains far from complete and requires more experimental and theoretical studies.

The BG is also widely recognized for its involvement in learning [6,7]. Reinforcement learning is recognized as the mechanism that establishes behavioral responses for rewards, such as food or drugs of abuse and is altered in numerous disorders and disease states including Parkinson's disease [8–10]. Reinforcement learning is based on communication between mid-brain dopamine neurons and the striatum [9], specifically ventral tegmental area (VTA) projections to ventral striatum in the mesolimbic neurocircuit and substantia nigra pars compacta (SNc) projections to dorsal striatum in the BG [11,12]. Dopamine (DA) released by dopaminergic VTA and SNc inputs to striatum signals the difference between received and expected rewards—the reward prediction error (RPE) [10,13]. RPE encoding in VTA-ventral striatal neurocircuits involves prediction of reward value which in turn feeds back to both VTA and SNc dopamine neurons [13]. Given its role in motor control, the SNc-dorsal striatum component of the BG translates RPE into action: the hypothesized critic-actor roles of these two dopaminergic projections [13,10]. If the RPE is positive, additional DA release leads to positive reinforcement of the preceding action; if the error is negative (expected more than received), a pause in DA release leads to negative reinforcement and blocks the action. As a mechanism for this control, DA modulates plasticity of synaptic projections from the cortex to striatal medium spiny neurons (MSNs) [14,15]. As a reflection of the bidirectional DA modulation, there are two types of MSNs. Those that are responsible for promoting movement are part of the BG direct pathway and express D1-type dopamine receptors (D1-MSNs) and those that inhibit movement are part of the BG indirect pathway and express D2 dopamine receptors (D2-MSNs) [16–18]. Indirect and direct BG pathways respectively inhibit or disinhibit the thalamocortical relay neurons responsible for producing particular movements [19,20]. The coordination of activity within the two types of MSNs determines action [21–23]. Within the BG loops, synaptic plasticity of corticostriatal projections is a key node in the learning of rewarded choices [6,7,24,15].

The BG is suggested to remain involved in action selection after the action-reward association is learned [5,25]. On the other hand, clinical interventions for Parkinson disease (PD) do not cause impairments in learned movements [26–28]. Specifically, GPi lesions and deep brain stimulation (DBS) in the STN, which both thought to disrupt the main output of the BG, are used to improve motor functions. This observation gave rise to a hypothesis that the BG play a critical role in learning, but not in the expression of already learned actions or choices [29,30]. These choices are suggested to instead be stored in synaptic connections within cortex. This hypothesis apparently contradicts the suggested involvement of the BG in executing actions learned previously. Therefore, it is essential to fill in this knowledge gap by further investigating the role of the BG in action learning.

Mathematical modeling have been widely used to reproduce and explain various aspects of BG electrophysiology and related behavior. A large set of these models is focused on understanding the dynamics of specific neurons in disease and control conditions irrespective of the

BG function [31–34]. Other models are constructed based on functional ideas and emulate how biophysical changes caused by a disorder violate the functions [35–39]. Both types of models contribute to understanding of the BG function at different levels [40,41]. However the picture remains far from complete, and the obvious reason for that is the complexity of the BG circuitry and physiology, as well as the diversity of its functions. The present model was designed as a simple implementation of principles suggested to underlie the learning and action selection functions of the BG in the simplest, yet most frequently used two-choice instrumental conditioning task. This simplification allows for a comprehensive implementation of mechanisms for BG functions and dysfunctions. Thus, our paper addresses the need for simplified BG model that reproduces learning and action selection in a standard behavioral task.

The goal of the present study was to design a simple model of BG function that utilized experimentally known physiological processes and replicate behavior in a classical task. Such a model would provide an opportunity for identifying gaps in knowledge to better guide additional experimentation. To this end, this paper presents a computational model of the cortico-BG-thalamo-cortical loop involved in a two-choice instrumental conditioning task [25]. This task is standard for assessing action-reward association in animals and humans. Our model design is similar to a previously published design [30,42], but focused on choice selection. We implemented two synaptic mechanisms that can mediate learning: reward-related plasticity of corticostriatal synapses [43] and activity-dependent Hebbian plasticity [44,45] of cortico-cortical synapses. To elucidate the role of the BG in Parkinson's and Huntington diseases, we calibrate the model to reflect the altered BG connectivity documented for these diseases and simulate these changes in BG activity.

Materials and methods

We adopt rate model formalism extensively used to reproduce activity and function of numerous brain structures [46]. In particular, we follow a validated model of motor control [42] and modify it for action selection.

Structure of the basal ganglia

Fig 1 presents a schematic diagram of nuclei and connections within the BG and their connections with cortices. The cortico-BG-thalamo-cortical loop is separated into channels selective for each of the two actions of the model (see below). First, the striatum, the primary input structure of the BG, receives excitatory inputs from the prefrontal cortex (PFC) and premotor cortex (PMC) in the cerebrum as well as the thalamus. From the striatum, two competing pathways are activated: a direct pathway (striatum-SNr/GPi) and an indirect pathway (striatum-GPe-STN-SNr/GPi). These two pathways converge at the BG output nuclei, the SNr and GPi, and serve to modulate their activities. In the model SNr and GPi activity are treated as one unit. SNr/GPi activity inhibits a corresponding neural group in the thalamus and PMC and blocks the corresponding action. Thalamus and PMC activity is treated as a single unit (PMC/Thal). To execute the action, SNr/GPi activity must decrease and disinhibit the PMC/Thal neurons. The actions (channels) compete with each other via reciprocal inhibition at the PMC level. Reciprocal inhibition also exists at the GPe level, but it was omitted in the model as STN-GPe network dynamics was shown to be only weakly dependent on this inhibition [47,48]. In addition, DA neurons in the SNc signal a reward prediction error (RPE), which change synaptic weights of PFC-striatum connections via DA-dependent long-term synaptic potentiation (LTP) and long-term synaptic depression (LTD) to allow for reward-based learning.

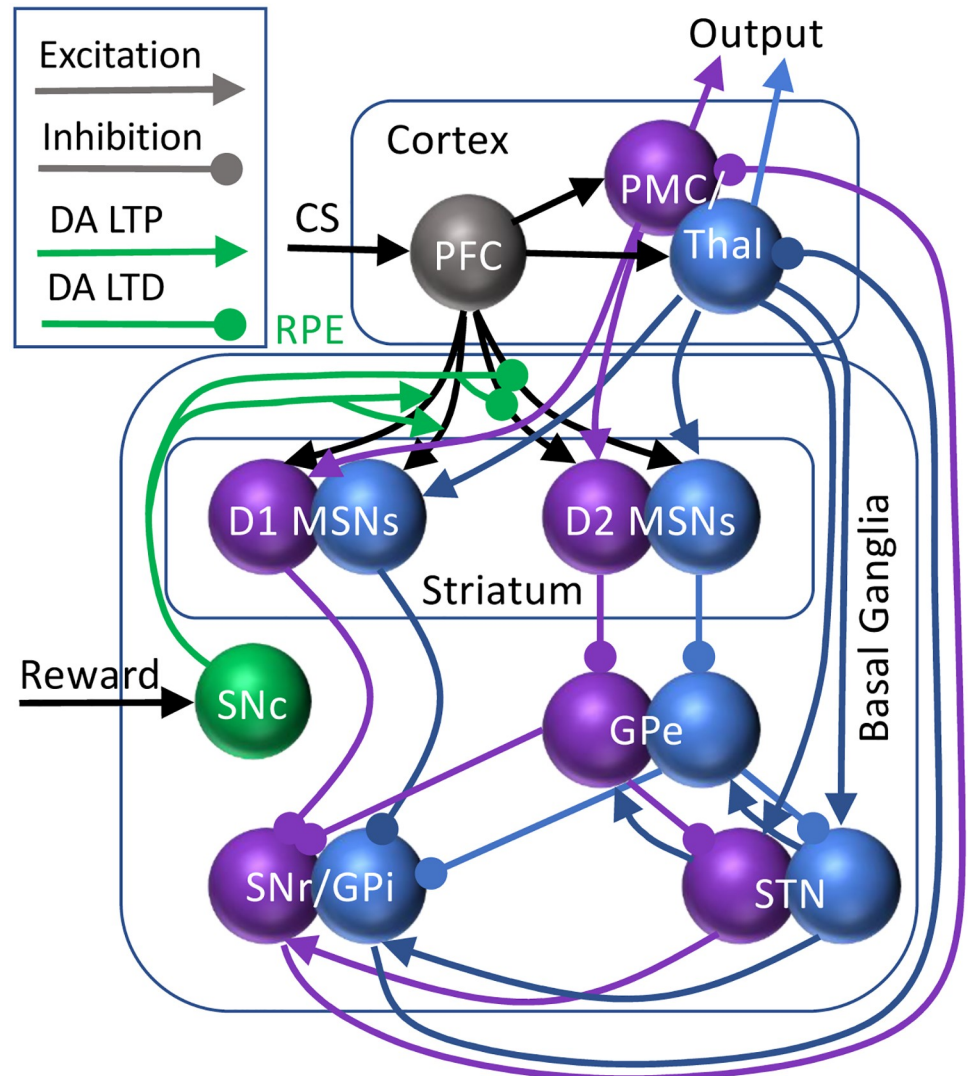


Fig 1. The structure of the cortico-basal ganglia-thalamo-cortical loop model. The BG receives inputs from the prefrontal cortex (PFC) signaling the conditioning stimulus (CS) as well as reward inputs via substantia nigra pars compacta (SNc). The SNc forms a dopamine reward prediction error (RPE) signal, which governs plasticity of the connections from the PFC (DA LTP/LTD; green). The BG input structure, striatum, contains medium spiny neurons (MSNs), which cluster in 2 subtypes: D1 and D2 dopamine receptor-containing (direct and indirect pathways respectively). The rest of the nuclei are the globus pallidus external (GPe), subthalamic nucleus (STN), and the output structures: substantia nigra pars reticulata and globus pallidus internal (SNr/GPi). The loop is completed by connections from and to premotor cortices/thalamus (PMC/Thal). The two channels of the loop are colored purple/blue.

<https://doi.org/10.1371/journal.pone.0228081.g001>

Behavioral task

Our model implements a standard design for intertemporal choice tasks [25]. The circuitry shown in Fig 1 is built to reproduce selection between two actions, one of which is rewarded. A typical task is to learn that, for instance, action 1 is rewarded if a conditioning stimulus (CS) is presented. Then, this task is “reversed”: after learning this contingency, the reward following the same CS is shifted to action 2. Thus, the cortico-BG-thalamo-cortical loop has 2 channels: for choice 1 and 2, except for the PFC that represents the CS and the SNc that represents the unexpected reward. Activation of neural groups 1 and 2 in the PMC/thalamus correspond to

execution of action 1 and 2 respectively. Thus, in the model, an action is considered selected if the activity level of the corresponding PMC neural group at the end of a simulated trial exceeds that of the other group plus the noise level (0.1; see below). The behavioral readout is if the stimulus-reward contingencies can be learned, and how many trials learning takes.

Firing rate equations

The activity of every neuron (except the dopaminergic neurons in the SNc) is governed by the following differential equation [42]:

$$\tau \frac{dA}{dt} = \sigma(I) - A + N(t) \tag{1}$$

where A is the instantaneous activity level of the neuron. $N(t)$ is a uniformly distributed noise with amplitude 0.1. Here, τ is a time constant taken to equal 12.8 msec for the STN, 20 msec for the GPe, and 15 msec for all other neural groups based on previous models and experimental studies [49]. I is the synaptic input to the neuron. The expressions for synaptic input to each neuron group, and the formula are compiled in Table 1. $\sigma(I)$ is a normalized response function defined as:

$$\sigma(I) = \begin{cases} 0, & \text{if } I \leq 0 \\ \tanh(I), & \text{if } I > 0 \end{cases} \tag{2}$$

This formalism normalizes firing rates for all nuclei to be between 0 and 1 to avoid difficulties with modeling very different firing rates observed in BG of different species and allows us to focus on the general learning mechanism. We have adapted the following notation: X_m to denote the activity (firing rate) of neural group X in the pathway for the m^{th} action. Since our model contains only two actions, the only possible values for m are 1 and 2. The index n in the formula for X_m is equal to 2 if $m = 1$, and $n = 1$ if $m = 2$, i.e. it refers to the other of the two channels and describes interaction between them. Further, w_{X-Y} denotes the synaptic weight (strength of connection) from group X to group Y and dr_X denotes a tonic drive to group X . Many of these weights are assumed constant throughout our trials, but several of them are plastic as described below.

Synaptic plasticity

The synaptic weights from PFC to PMC neurons and from PFC to MSNs are plastic, which means that they change depending on the activity of these nuclei and behavioral outcome (reward received) respectively [44,45,43]. In simulations, the synaptic weights are updated at the beginning of every trial depending on the behavior of the model in previous trials. Before

Table 1. Synaptic inputs.

Neuron	Formula for Synaptic Input
PFC	$I_{PFC} = input_pfc$
D1 MSN	$I_{D1\ MSN_m} = w_{PFC-D1} PFC + w_{PMC-D1} PMC_m$
D2 MSN	$I_{D2\ MSN_m} = w_{PFC-D2} PFC + w_{PMC-D2} PMC_m$
GPe	$I_{GPe_m} = dr_{GPe} - w_{D2-GPe} D2\ MSN_m + w_{STN-GPe} STN_m$
STN	$I_{STN_m} = dr_{STN} - w_{GPe-STN} GPe_m + w_{HD} PMC_m$
GPI	$I_{GPI_m} = dr_{GPI} - w_{D1-GPI} D1\ MSN_m + w_{STN-GPI} STN_m$
PMC	$I_{PMC_m} = dr_{PMC} + w_{PFC-PMC_m} PFC - w_{GPI-PMC} GPI_m - w_{PMC_n-PMC_m} PMC_m$

<https://doi.org/10.1371/journal.pone.0228081.t001>

we discuss the specific mechanisms by which we updated these plastic synaptic weights, we will first discuss how we calculated the activity of the dopaminergic neurons in the SNc, which essentially mediate reward-based learning.

The activity of the SNc neurons is associated with a reward prediction error (RPE) [50]. Following previous models (e.g. [42]), we assume that the activity of the SNc neural group reflects the difference between the expected reward and the actual reward:

$$SNc = R - R_j^e \tag{3}$$

where R is the actual reward given based on the action selected, and R_j^e is the expected reward at the j^{th} trial. The animals are pre-trained on single choice task, and, therefore, they expect a reward. The expected reward on the first trial, R_1^e , is equal to 1 and is then subsequently updated according to the following scheme [42]:

$$R_{j+1}^e = \alpha R_j + (1 - \alpha)R_j^e \tag{4}$$

where α is a constant (set equal to 0.15) and R_j denotes the actual reward received by the model on the j^{th} trial. Note that the RPE given by Eq (3) may be positive (actual greater than expected) or negative (actual less than expected).

The actual reward received in simulations, R , is determined by the following:

$$R = \begin{cases} 1, & \text{if rewarded action performed} \\ 0, & \text{if rewarded action not performed} \end{cases}$$

where we determined which action is selected by comparing the activities of the PMC neurons at the end of each trial as described above.

Altogether, after each trial, the PFC-striatal synaptic connections are updated by adding the following increments to the previous values:

$$\Delta w_{PFC-D1m} = \lambda_{D1} * SNc * PFC * D1_m - d * w_{PFC-D1m} \tag{5}$$

$$\Delta w_{PFC-D2m} = -\lambda_{D2} * SNc * PFC * D2_m - d * w_{PFC-D2m} \tag{6}$$

where PFC , $D1_m$, and $D2_m$ denote the activity of the respective neural group at the end of the trial ($m = 1,2$). Here, λ_{D1} and λ_{D2} are learning rate constants and d is the decay rate constant. The value for these constants are adopted from the previous literature [30,42] with a modification that takes into account the differences for synapses contacting D1 and D2 MSNs: Gurney et al. [51] has shown using experiments and modeling that plasticity of the synapses on D2 MSNs is weaker approximately by a factor of two than on the D1 MSNs. Note that the formalism does not allow for direct modeling of the eligibility traces necessary for stimulus-reward association [52–54], and we account for that by PFC activity that persists for the duration of the trial.

Lastly, we describe the mechanism by which we updated the connections between the PFC and PMC neurons. Here, we let $w_{PFC-PMCm}$ denote the synaptic weight of the connection between the PFC neural group and the m^{th} PMC neural group. After each trial, the synaptic weights are updated according to the following Hebbian Learning Rule:

$$\Delta w_{PFC-PMCm} = \lambda_{CM} * PFC * PMC_m - d_{CM} * w_{PFC-PMCm} \tag{7}$$

where λ_{CM} is the learning rate and d_{CM} is the decay rate of the cortical connections. Here, PFC and PMC_m denote the activity of the PFC neurons and m^{th} PMC neuron group at the end of the trial.

Now, we will outline our methodologies for calibrating our three different BG model states: healthy, Parkinsonian, and Huntington’s disease.

Healthy BG state

We target to reproduce rodent behavior in instrumental conditioning (IC) tasks [29,25]. Thus, an animal will learn contingencies between a conditioning signal and a rewarded action—pressing one of two levers. We reduce the model by [30,42] and focus our model on the interaction of the thalamocortical and BG networks (Fig 1) and reproduce the function of the cortico-BG-thalamo-cortical loop in the above two-choice task. The parameter values are shown in Table 2. The values were taken from previous studies [30,42] with a few minor modifications that allow for both robust instrumental conditioning as well as reversal learning.

Specifically, equations for the D1 and D2 MSN neurons reproduce their balanced excitation by cortical inputs in vivo [21,55]. The balance is supported by a number of complex mechanisms, from differential effects of DA on excitability of the D1 and D2 MSNs [56] to their lateral inhibition and contribution of striatal fast spiking interneurons [32]. These mechanisms are very hard to implement in a rate model, and we calibrate the D1 and D2 MSN equations identically to reflect the balance. The balance is perturbed in the PD DA depleted state (see below).

Parkinsonian BG state

The neuropathology of Parkinson’s Disease (PD) is well-understood: it begins with the destruction of the dopaminergic neurons in the SNc [57,58]. Further, the disease is accompanied by a decreased firing rate of the D1 MSNs [59,55], GPe [60–62], and PMC [63] as well as increased firing rates in the D2 MSNs [59,55], STN [64,65], and GPi [66,60,67]. We induced an in silico mild Parkinsonian state in our model by suppressing SNc output by 70% and changing synaptic weights along with tonic drives [32,58,68–75] as outlined in Table 3. There

Table 2. Parameters of the healthy BG model state.

Parameter	Value used in this model
$input_{pfc}$	3.0
w_{PFC-D1_m} & w_{PFC-D2_m}	Randomly set between 0 and 0.001, updated after each trial
w_{PMC-D1}	2.0
w_{PMC-D2}	2.0
dr_{GPe}	1.6
w_{D2-GPe}	2.0
dr_{STN}	0.8
$w_{GPe-STN}$	1.0
dr_{GPi}	0.2
w_{D1-GPi}	1.4
$w_{STN-GPi}$	1.6
dr_{PMC}	1.3
$w_{PFC-PMC_m}$	Initial 0; varies with trials
$w_{GPi-PMC}$	1.8
$w_{PMC_m-PMC_n}$	1.6
$w_{STN-GPe}$	0.4
w_{HD}	0.3
λ	0.0005
λ_{CM}	0.0005

<https://doi.org/10.1371/journal.pone.0228081.t002>

Table 3. Changes in the parameters of the model that reproduce Parkinsonian BG state.

Parameter	Value in Healthy state	Value in mild Parkinsonian state	Justifying literature
w_{PMC-D1}	2.0	1.25	[57,32,53]
w_{PMC-D2}	2.0	2.75	[57,32,53]
w_{D2-GPe}	2.0	2.4	[67–69]
$w_{GPe-STN}$	1.0	1.2	[70–72]
dr_{STN}	0.8	1.0	[62,63]
dr_{GPe}	0.2	0.25	[58,64,65]
w_{D1-GPi}	1.4	1.1	[58,64,65]
$w_{STN-GPi}$	1.6	2.0	[58,64,65]
$w_{STN-GPe}$	0.4	0.5	[73]

<https://doi.org/10.1371/journal.pone.0228081.t003>

are multiple mechanisms that break the activation balance of D1 and D2 MSNs in the DA depleted state [56]. All of them lead to the net increase in the activation of the D2 MSNs and decreased activation of D1 MSNs [32]. Thus, we model these changes by decreased synaptic excitation of D1 MSNs and increased synaptic excitation of D2 MSNs (Table 3).

Huntington’s BG state

The pathology of Huntington’s Disease (HD) is less well-understood; however, it is clear that there is a progression of the disease from chorea (involuntary, jerky movement) at its onset to akinesia (loss of the power of voluntary movement) at its conclusion [76]. We modeled the chorea phase (Grade 2 HD) by weakening the D2 MSN-GPe connection by 75%, weakening the D1 MSN-GPi connection by 35%, and decreasing the PFC and PMC inputs to account for destruction of the cortices [76,77]. These percentages are gathered from the physiological observations of Reiner et al. [76]. The resulting parameters are shown in Table 4.

Numerical simulations

Our model was coded in MATLAB. We considered a trial to last 750 msec, and at the end we register the activity of each neuron in the circuit. We chose to cutoff trials at this point because it was sufficient to guarantee that the neural activity converges to a steady state. An exception is a case when neural activity does not approach a steady state and remains oscillatory, which we also found in this study. We update strengths for the plastic synapses after each trial. Finally, we reset the initial activity of the neurons to be at randomized levels at the beginning of each subsequent trial. We ran simulations consisting of 500 such trials. The code is available in ModelDB database <http://modeldb.yale.edu/261616>.

Results

We simulated the same standard two-choice IC and reversal task in three conditions: Healthy, Parkinsonian, and Huntington’s BG. Fig 1 presents a schematic diagram of nuclei and

Table 4. Changes in the parameters of the model that reproduce Huntington disease state.

Parameter	Value in Healthy state	Value in Grade 2 HD State	Justifying literature
$input_{pfc}$	3.0	0.8	(74,75)
$w_{PMC-D1,2}$	2.0	1.5	(74,75)
w_{D1-GPi}	1.4	0.9	(74,75)
w_{D2-GPe}	2.0	0.5	(74,75)

<https://doi.org/10.1371/journal.pone.0228081.t004>

connections within the BG and their connections with cortices. The model is described in detail in Materials and Methods. The models received a stimulus (CS) that activates prefrontal cortical (PFC) neurons for all 500 trials. We say that the network chooses action 1 if the activity of the premotor cortical (PMC) neural group 1 exceeds the activity of the PMC group 2 by 0.1. The comparison of the activity levels is made at the end of each trial. For reversal training, after action 1 is rewarded in trials 1 through 199, for trials 200 through 500, action 2 was rewarded instead. We analyze and compare the learning and reversal performance in the three model states below.

Healthy BG facilitates learning of rewarded choices

Fig 2A shows choices made in the simulations: a higher activity of the PMC1 manifests choice 1 and vice versa. The graph shows the activity at the end of each trial, which is taken to be 750 msec long. On early trials, the choice is made randomly due to random initial conditions in the PMC network and mutual inhibition of PMC1 and PMC2. This reproduces the exploration phase, where the information about reward is collected [78,79]. The modeled animal receives an unexpected reward every time it chooses action 1 (PMC1 on top). Within several trials, the system starts to consistently choose the rewarded action, although a few exploratory deviations may be made after that. This fast initial learning replicates experiments and is thought to occur so fast because animals are pretrained first on a single choice task (e.g. to press a single lever

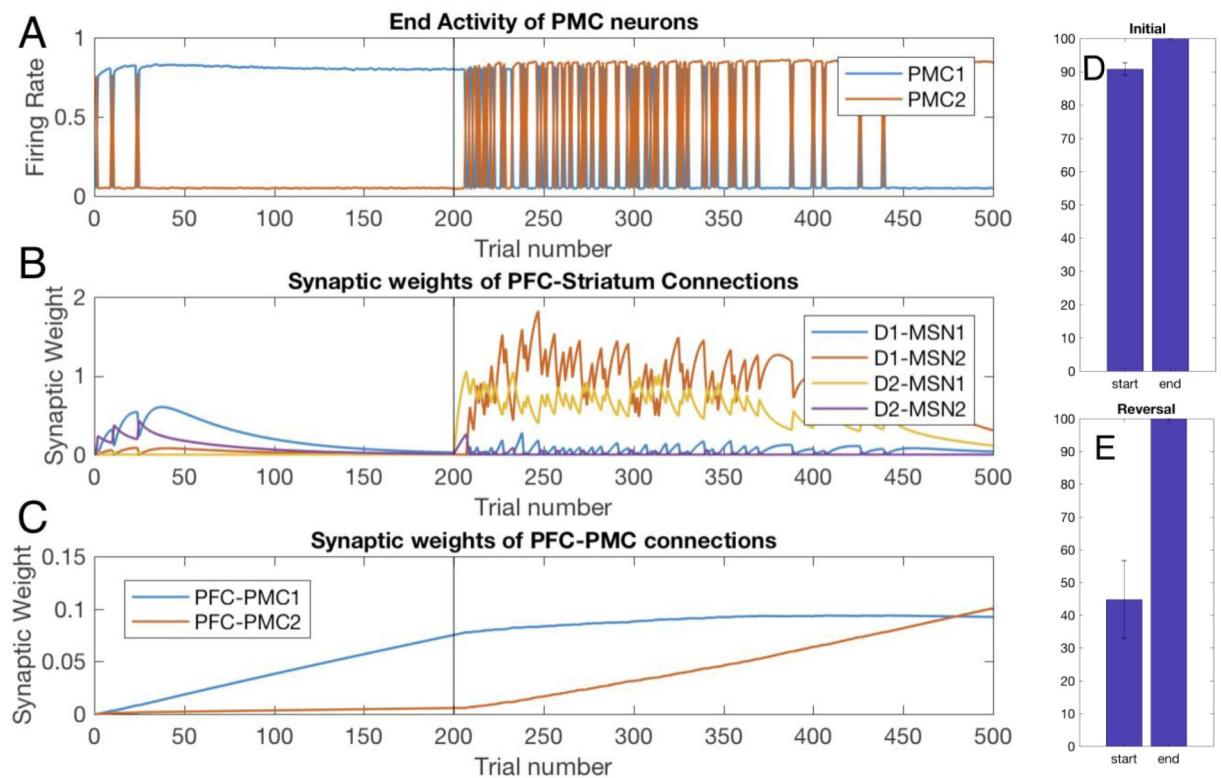


Fig 2. Healthy BG facilitates learning of the initial task and reversal. Trial-by-trial dynamics of the PFC activity and underlying modulation of synaptic weights in the Healthy BG model. Trials 1–199: initial learning; trials 200–500: reversal (A) A higher activity of PMC1 (blue) manifests choice 1, whereas higher activity of PMC2 manifests choice 2. (B) Synaptic weights of the PFC to striatum connections. (C) Synaptic weights of the PFC to PMC connections. (D) and (E) Percentage correct trials over 25 trials at the start and the end of the initial learning and reversal learning respectively. The scores were averaged over 10 simulated animals.

<https://doi.org/10.1371/journal.pone.0228081.g002>

for reward). On trial 200, we switch the simulated task to reversal: action 2 is rewarded instead. This quickly leads to reestablished exploratory behavior, and then locks the system to the rewarded choice, with occasional exploratory returns to choice 1. Fig 2E shows performance improvement over reversal learning that matches experiments [29]. As explained below, our model allows for detailed analysis of the mechanism of this learning.

Two mechanisms facilitate learning of the rewarded choice—one fast and one slow. The first mechanism is the potentiation of the PFC-to-striatum synaptic connections (Fig 2B). Since the animals are pre-trained on single choice task, they expect a reward, and a reward omission creates a negative RPE (unexpected punishment; Fig 3) encoded by SNc DA signaling that potentiates PFC connections to all D2-MSNs (Fig 2B). Importantly, whereas the DA signal itself is not selective for MSNs specific to the rewarded action, DA-mediated potentiation of PFC-MSN synapses is selective. What makes potentiation selective is the level of activation of the corresponding MSN: in the initial trials the reward is omitted if choice 2 is selected, that is when PMC2 activity is greater, and, consequently, MSNs selective to choice 2 are activated more (due to static synaptic connections from PMC to MSNs specific for each choice). Since synaptic plasticity explicitly depends on the activity of the postsynaptic neuron, PFC-to-D2-MSN2 connections are potentiated much more strongly than D2-MSN1 connections (Fig 2B purple vs. yellow). Then, every choice that is not followed by the expected reward activates the corresponding indirect pathway (i.e. D2-MSN2), which excites the downstream GPi2 neurons, and consequently inhibits the PMC2 activity. This blocks the nonrewarded action and helps to lock the choice to the rewarded action.

Simultaneously, reward omission reduces expected reward, and the next rewarded trial results in positive RPE and leads to potentiation of the connections to D1-MSNs (Fig 2B blue). This further selectively activates D1-MSNs responsible for action 1. The mechanism for this selectivity is the same: the reward is granted only if choice 1 is selected, that is when PMC1 activity is greater, and, consequently, the MSN corresponding choice 1 is activated more (Fig 4). The increased activity level of D1-MSN1s selectively inhibits downstream GPi1 neurons

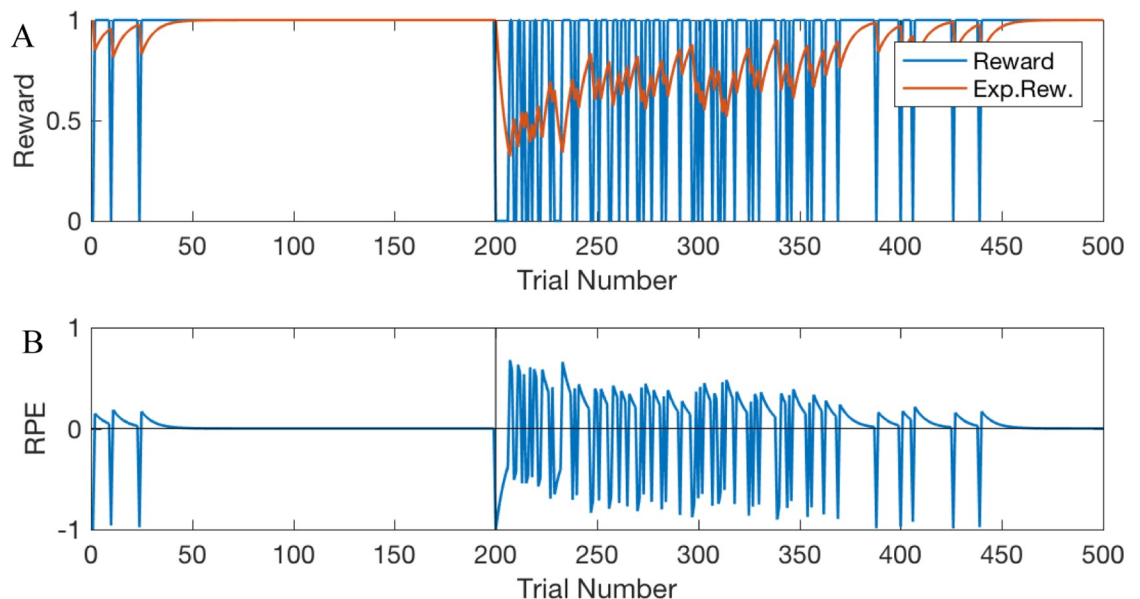


Fig 3. Reward, expected reward (A), and the RPE (B) during initial learning and reversal trials in the model with healthy BG. As before, reversal starts at trial 200 (vertical black line).

<https://doi.org/10.1371/journal.pone.0228081.g003>

and, consequently, disinhibits the PMC1 neural group (Figs 2A and 4). Thus, due to direct excitation from the PFC associated with the stimulus, activation of D2-MSNs associated with choice 2 and D1-MSNs associated with choice 1 increase. Co-activation of the two mechanisms is sufficient to lock the choice to the rewarded action. Note that the complex pattern of co-activation of D1 and D2 MSN populations is in agreement with the recent literature [21,80].

During subsequent repetitions of the same trial, the PFC-MSN connection strength starts to decrease and approaches zero (Fig 2B trials 40 to 200). However, the persistence of the rewarded choice remains intact (Fig 2A). The mechanism for this is the growth of direct PFC-PMC1 connections (Fig 2C) via classical reward-independent Hebbian synaptic plasticity: the two neural groups are co-active most of the time. This transition from PFC-MSN to PFC-PMC connections as a robust supporting mechanism for the rewarded choice occurs after the number of repetitions exceeds approximately a hundred (Fig 2). In these later trials, the PFC-MSN connection strengths are decreased, but the choice remains locked to the rewarded action. Therefore, the model predicts that direct cortico-cortical connections are responsible for the choice of the rewarded action after long training.

We next analyzed the behavior of the model when we began rewarding a choice different from the choice the model had been previously conditioned to make; this learning task is called reversal learning [81]. Beginning at trial 200, we rewarded the model for selecting the other action (choice 2). Thus, starting at trial 200 the model mimics omission of a reward (unexpected punishment) for selecting action 1. This punishment potentiates synaptic connections from the PFC to D2-MSNs associated with action 1 (D2-MSN1, Fig 2B yellow), and, slightly later, to D1-MSNs associated with action 2 (D1-MSN2, Fig 2B red). This engagement of both direct and indirect pathways offsets the model bias for action 1 and quickly sends the model into another exploratory phase. As Fig 2A demonstrates, between trials 200 and 300 the model is randomly choosing between the two actions. It is important to note that, in accordance with others' findings [82,83], this second exploratory phase lasts longer than the initial exploratory phase. During reversal, the new potentiation of PFC-MSN connections is not enough to

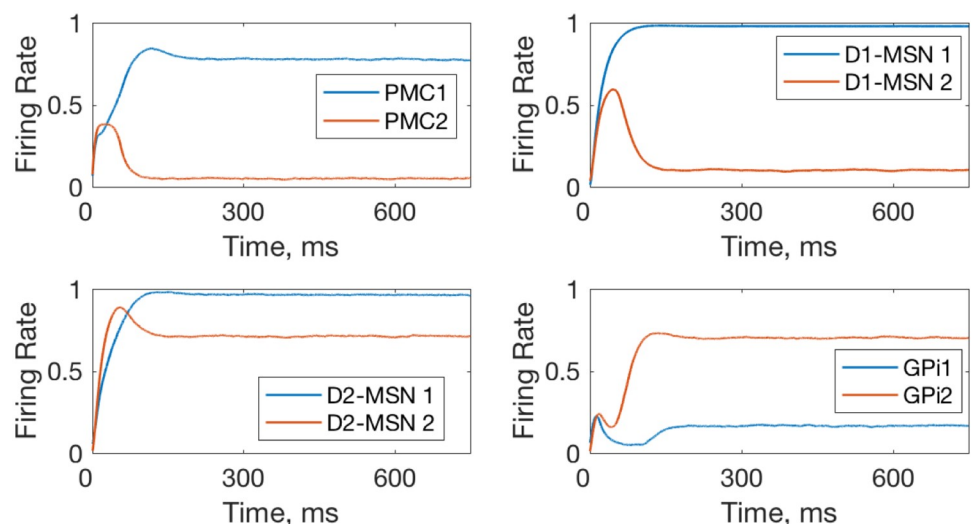


Fig 4. Within-trial dynamics of neural activity in the model with healthy BG. The network is biased towards option 1 as the PFC-D1-MSN1 and PFC-D2MSN2 connection weights are both set at 0.7, which corresponds to a trial in late initial learning phase (~100). Activation of the D1-MSN1 group inhibits GPi1 neurons, and thus disinhibits PMC1. GPi2 neurons remain excited and inhibit PMC2.

<https://doi.org/10.1371/journal.pone.0228081.g004>

effectively overcome the bias for the initially learned choice and ensure choosing the newly rewarded option. The reversal exploratory phase ends only when the PFC-PMC2 connections become as strong as PFC-PMC1 and remove the bias (Fig 2). Thus, the longer exploratory phase during reversal occurs because the model must first overcome its bias for the previously learned choice and then develop a new stimulus-choice 2 association. The lengths of the exploratory phases matches experimental results [82,83].

After the onset of reversal learning, the system continues choosing option 1, even though it's not rewarded, due to the potentiated PFC-PMC1 connection. This generates a negative reward prediction error (Fig 3) and potentiates PFC connections to the D2-type neurons associated with action 1 (D2-MSN1; Fig 2B yellow). The connection of PFC to D1-MSN2 lags by several trials (Fig 2B, red), during which the exploratory phase begins and allows finding the new rewarded option. Both the initial and the reversal learning engage direct pathways at a greater strength than the indirect (Fig 2B, yellow and red) because the model reflects a greater plasticity rate for the cortical connections to D1 than to D2 MSNs.

Mild parkinsonian BG: Impeded learning and spontaneous oscillations

In our Parkinsonian BG mode, the indirect pathway is strengthened by parameter changes reflecting physiological data. Our simulations (Fig 5) show drastic difference in dynamics of the PMC neurons during initial learning and reversal in the model with mild-parkinsonian BG. During both phases, learning is severely impaired. First, the choice remains random for approximately the first 50 trials. Second, the model does not reliably choose the rewarded

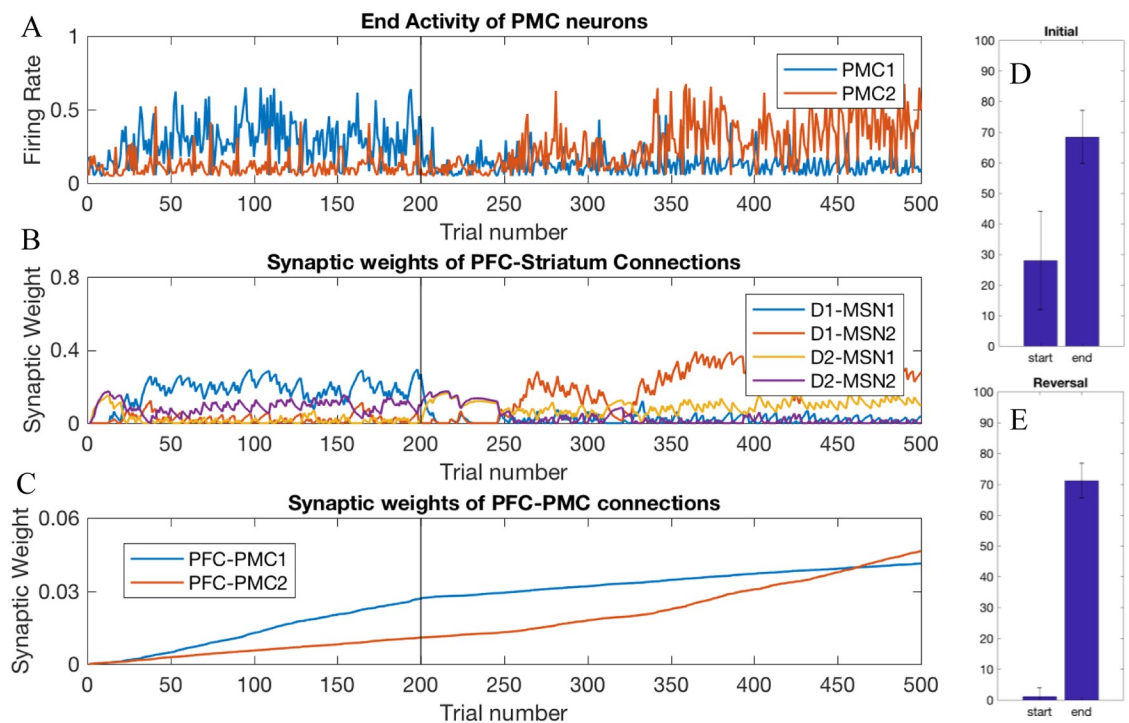


Fig 5. Decreased learning performance and increased variability of PMC activity in the model with mild-parkinsonian BG. Trial-by-trial dynamics of PMC activity (A) and underlying modulation of synaptic weights (B,C) in the model with mild-parkinsonian BG state. Notation is the same as in Fig 2. Note the difference in scale in panels (B) and (C) compared to Fig 2. (D) and (E) present performance at the beginning and the end of the initial learning and reversal respectively. The scores were averaged over 10 simulated animals.

<https://doi.org/10.1371/journal.pone.0228081.g005>

option even after this period, although the rewarded option is chosen on a much greater number of trials (Fig 5A blue above red in the initial learning and vice versa in the reversal). Third, the activity of the PMC neurons is overall reduced compared to that in the model with healthy BG, and the trial-to-trial variations of this activity are drastically increased, even when only trials with the same choice are considered.

The underlying dynamic of the synaptic weights is also significantly altered. During both initial learning and reversal, the activation levels for both direct and indirect pathways (Fig 5B) is much lower than in the model with healthy BG (Fig 2B). The latter follows directly from the reduced SNC signaling (by 70%), which decreases the RPE and, thus, impedes potentiation of PFC-MSN connections. Since both PMC neural groups are active at a similar level, both connections from PFC are potentiated (Fig 5C), and the system does not develop a preference for the rewarded choice. After trial 50, the rewarded choice starts to prevail as the PFC-PMC connections reflect the preference for choice 1. However, the PFC-PMC1 connection does not achieve the level reached in the model with healthy BG (Fig 2C) within the 200 trials designated for initial learning. Hence, exploration between the choices persists for all 200 trials, and the prevalence of the rewarded choice requires the persistent activation of PFC-MSN connections. Therefore, the model with mild parkinsonian BG is capable of learning the choices, but the effective learning rate is much lower.

Reversal learning has been shown impaired in PD conditions [84–86]. In the model, the low levels of PFC-PMC connections persist into the reversal phase and never reach the levels shown by the model with healthy BG even though plasticity rules of the PFC-PMC connections remain the same in both models. Therefore, our modeling predicts that the mild-parkinsonian BG does not allow for the proper potentiation of the PFC-PMC connections, and this leads to impaired learning. Learning based on cortical synaptic potentiation simply reflects the choice frequency because the PMC group responsible for the choice fires together with the PFC. One reason for the lack of proper potentiation is that the models with parkinsonian BG cannot maintain the rewarded choice. Experimentally, the inability to maintain the choice was observed in 6-OHDA-lesioned rats [86]. The model also reproduces perseveration of the previously correct choice as shown experimentally [86] to contribute to the low performance at the very beginning of reversal (trials 1–5). Interestingly, the reversal phase starts with activation of both indirect pathways simultaneously (Fig 5B, purple and yellow). This suppresses the activity of both PMC neural groups, blocks any choice (Fig 5E) and blocks changes in the PFC-PMC synaptic weights. Only after some 50 trials, the blocking signal for choice 2 is removed (Fig 5 purple). The abrupt drop in the connection weight to the D2-MSN2 group is caused by positive RPE following the choice of a rewarded option. The choice is made by chance, and most of the previous trials were not rewarded because PMC activity was suppressed, and the probability of one PMC group to significantly exceed the activity of the other was low. Thus, the model with mild-parkinsonian BG predicts that the exploratory phase at the beginning of the reversal learning is replaced by blockade of any choice, and this further impedes learning.

Perhaps the most interesting change in the model with parkinsonian BG is the drastic increase in the trial-to-trial variability of the PMC neurons (Fig 5A). To explain the mechanism of this variability, we considered within-trial dynamics of activity for all neural groups in the model. Fig 6 shows these dynamics for the PMC, GPe and STN neural groups in the healthy vs. parkinsonian BG models. In the healthy case activity levels come to an equilibrium, while in the parkinsonian case, they engage in persistent oscillations. The oscillations arise from the negative feedback loop that the BG, and in particular its indirect pathway, provides for the activity of each PMC neural group. Indeed, the static PMC to D2 MSN connections, which constitute this negative feedback, are stronger in the parkinsonian case (w_{PMC-D2} , in

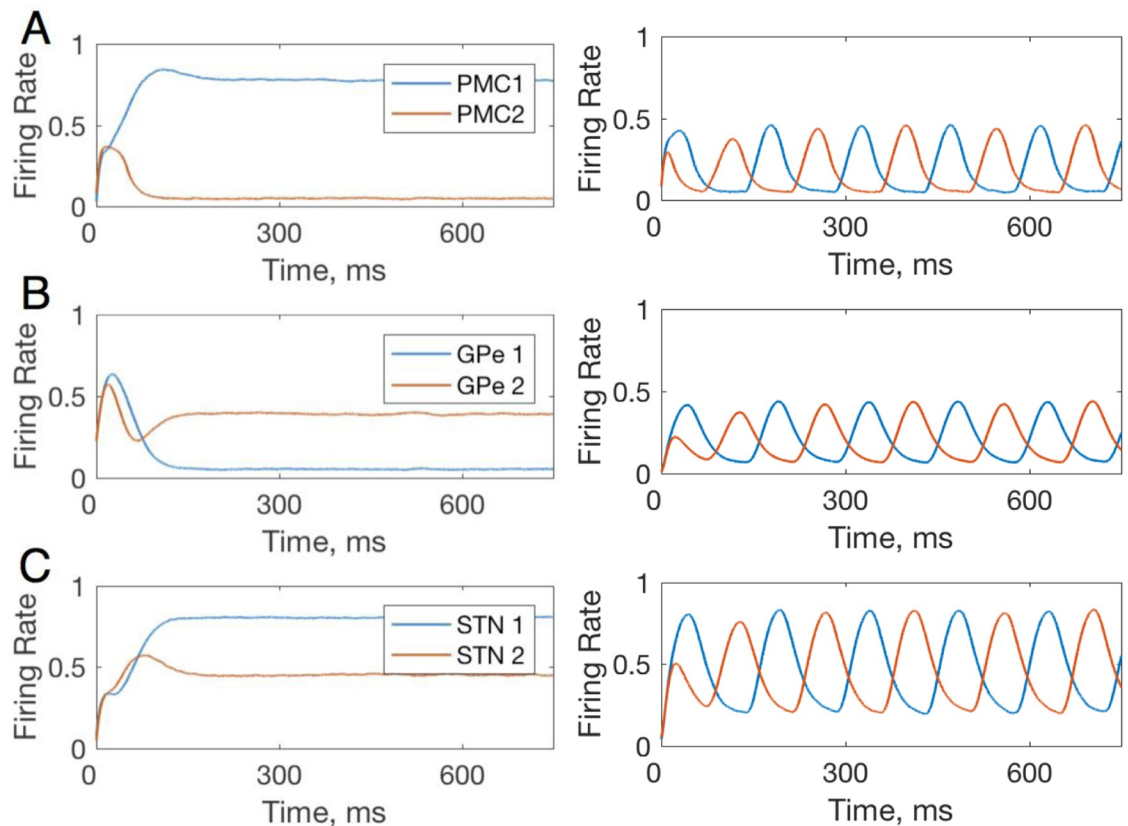


Fig 6. Within-trial dynamics of neural activity in the model with healthy (left) and parkinsonian (right) BG. Panels A, B, and C show firing rates for PMC,D1 MSNs and D2 MSNs respectively. In the healthy case, the firing rates equilibrate within 500 ms. In the parkinsonian case, oscillations in the firing rate emerge and persist. The anti-phase for the oscillations in the neural groups corresponding to the choice 1 and 2 is due to mutual competition (inhibition) between PMC1 and PMC2 groups. This synchrony may not persist for more complex tasks and setups. All plastic synaptic connections are set to zero to simulate the state of no bias towards any choice. Note different BG populations than in Fig 4 shown to focus on the contribution of the indirect pathway responsible for oscillations.

<https://doi.org/10.1371/journal.pone.0228081.g006>

Table 3). The period of these oscillations is approximately 150 ms, which is 6.7 Hz. No potentiation in the PFC-PMC and PFC-MSN connections within the ranges in Fig 5B and 5C suppress the oscillations. Therefore, the simulations predict that the trial-to-trial variability of the PMC neurons in the model with parkinsonian BG is caused by robust within-trial oscillations in the activity of all neuron groups in the model.

In order to model the impact of surgical interventions on performance and learning in PD, we performed additional simulations of the PD model in which the BG signal to PMC was ablated from trial 150 till the end (Fig 7). This directly models GPi lesions, which was the first standard surgical treatment for PD, and also mimics DBS treatment, which is suggested to reduce GPi output (see Discussion). In this period, the variability of the PMC activity vanishes completely. Furthermore, the PFC-striatal connections no longer exert any influence on the choices, but the PFC-PMC connections are strong enough to lock the choice to the rewarded option, and the cortical connections increase further at a greater rate. After the reversal on trial 200, however, the changed values of the choices remain unnoticed by the system, the choice remains locked on the now unrewarded option, and the cortical connections supporting this choice keep rising. In this state, behavior improves, but learning is impaired.

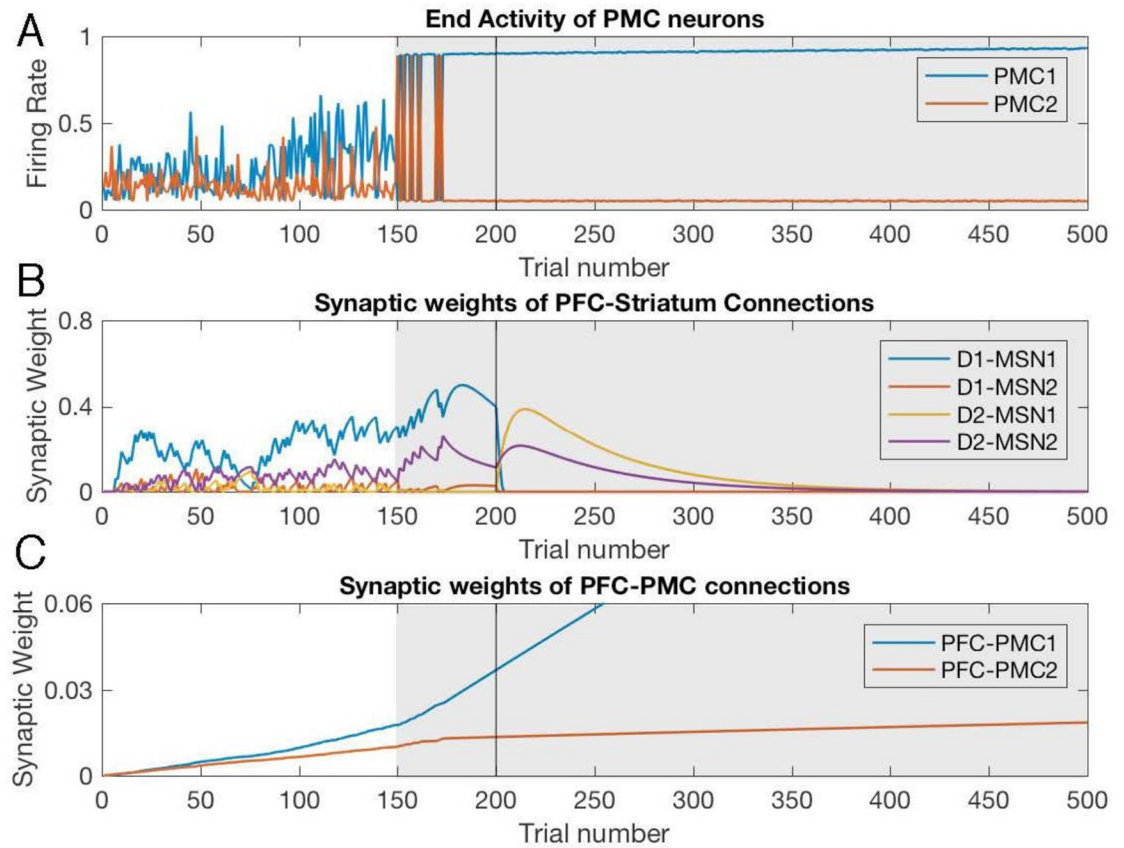


Fig 7. In the PD state model, the variability of PMC activity and switching between choice 1 and 2 cease shortly after the treatment onset. Trial-by-trial dynamics of the PMC activity and underlying modulation of synaptic weights in the PD BG model with simulated treatment starting at trial 150. Same notation as in Fig 2. (A) The levels of PMC1 and PMC2 activity (choice 1 vs. 2) at the end of each trial (B) Synaptic weights of the PFC to striatum connections reflect rewarded choices. (C) Synaptic weight of the PFC to PMC1 connection keep growing after treatment onset, and during reversal.

<https://doi.org/10.1371/journal.pone.0228081.g007>

Grade 2 Huntington's disease BG state: Persistent exploratory behavior

If the above case of Parkinson's disease is associated with strengthening the indirect pathway, in the case of Huntington's disease the connections in the indirect pathway become weaker (Table 4). The major difference with the healthy BG model is that the trial-to-trial dynamics of the PMC neural groups looks like the exploratory phase never ends (Fig 8A). At the same time, we see from the synaptic weights (Fig 8B and 8C) that choice-reward contingencies are learned almost as effectively as in the healthy case (notice similarity with Fig 2), although the synaptic weights are somewhat lower. The synaptic weight dynamics is qualitatively similar to the healthy case because the plasticity rules stay the same. The differences are the persistence of the potentiated PFC-MSN connections for the duration of initial/reversal training similar to the parkinsonian case and the activation of the indirect pathway for choice 2 lingering at the beginning of the reversal phase (Fig 8B purple). The former, however, is not a cause but rather a consequences of the continuous exploratory choices that bring no reward. Therefore, despite the efficacious learning (Fig 8C), choice behavior is impaired relative to control (Fig 8A).

The cause for the persistent exploratory phase is the positive PMC-BG feedback loop through D1 MSNs, which is not balanced by the D2 MSN pathway. Thus, activation of the D2

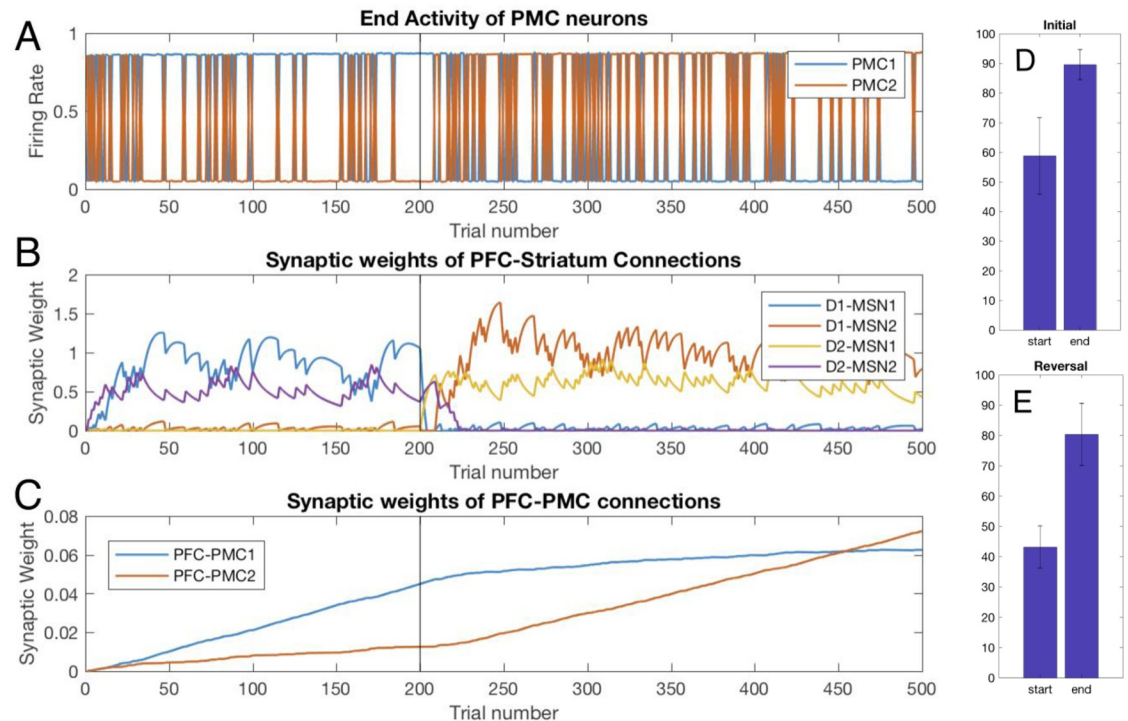


Fig 8. Random switches between rewarded and unrewarded options persist in the model with Huntington state BG. Trial-to-trial dynamics of PFC neural activity (A) and underlying dynamics of synaptic weights (B,C). The notation is the same as in Fig 2. (D) and (E) present performance at the beginning and the end of the initial learning and reversal respectively. The performance scores were averaged over 10 simulated animals.

<https://doi.org/10.1371/journal.pone.0228081.g008>

MSN pathway cannot robustly stop the unwanted action. Indeed, an occasional increase in the activity of the PMC2 neural group, which represents a non-rewarded action, excites the corresponding D1 MSN group, and through inhibition of GPi2 activity, further increases the PMC2 activity (Fig 9). The reduced connectivity in the D2 MSN pathway makes the GPi neural activity the same for choices 1 and 2 (Fig 9) and excludes the BG from the competition between the choices. This leads to occasional choices of the non-rewarded option, and our simulations show that this behavior is robust with respect to growing PFC-PMC and PFC-MSN connections (Fig 8). Therefore, the lack of balance between direct and indirect pathways in the model of Huntington's disease causes persistent random switching from rewarded to non-rewarded choice after both initial learning and reversal.

In order to model the impact of BG DBS or surgical interventions on performance and learning in HD, we also performed additional simulations of the HD model in which the BG signal to PMC was ablated from trial 100 till the end (Fig 10). The random switches between the choices cease shortly after, but not at the onset of the treatment. The response to the treatment is very similar to that in the PD case (Fig 7). In this period, the PFC-striatal connections no longer exert any influence on the choices, but the PFC-PMC connections are strong enough to lock the choice to the rewarded option. After the reversal on trial 200, however, the changed values of the choices remain unnoticed by the system, the choice remains locked on the now unrewarded option, and the cortical connections supporting this choice keep rising. Therefore, during DBS, or after surgical interventions ablating BG output, behavior improves, but learning is impaired in HD as well as in the PD state.

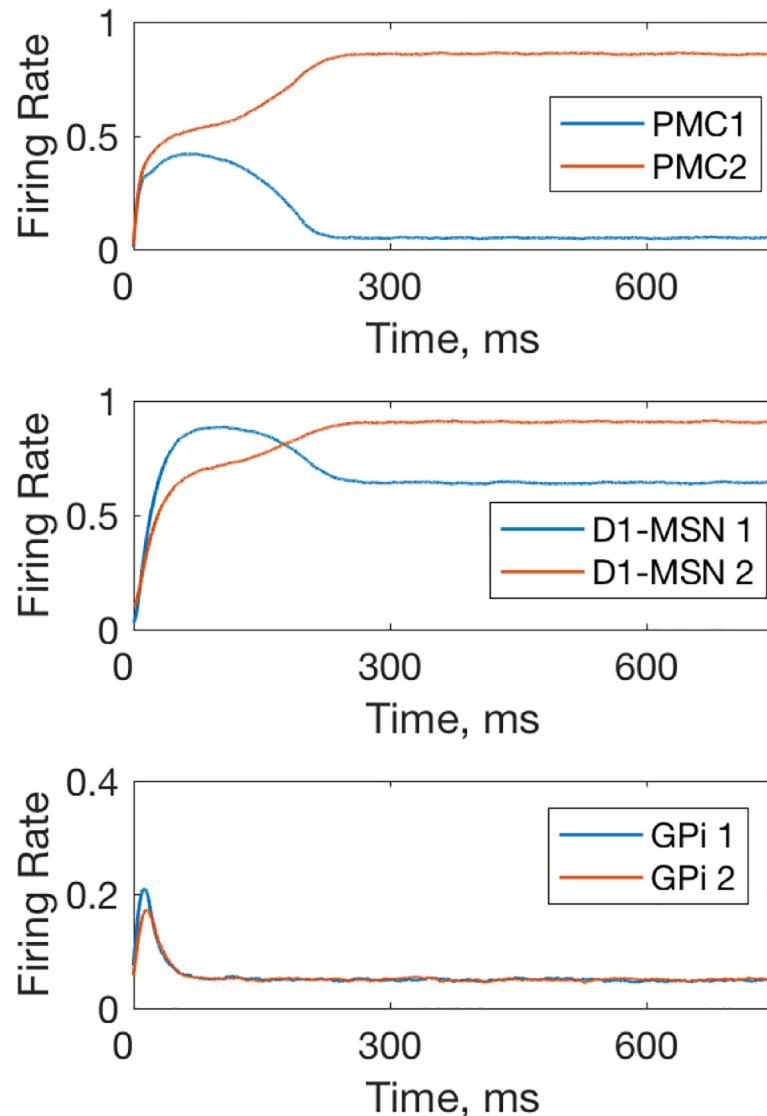


Fig 9. Occasional choice of the nonrewarded option made in the model with Huntington state BG. Within-trial dynamics of PMC, D1 MSN, and GPi neural activity is shown. The greater activity of PMC2 groups signifies that the action 2 is chosen, even though choice 1 is made preferable in the model by potentiating PFC-PMC1, PFC-D1 MSN1 and PFC-D2 MSN2 connections: $W_{PFC1-PMC1} = 0.04$, $W_{PFC1-D1MSN1} = 1$, $W_{PFC1-D2MSN2} = 1$. The rest of the connections remain at zero.

<https://doi.org/10.1371/journal.pone.0228081.g009>

Discussion

Our model implements the cortico-BG-thalamo-cortical loop function in a standard 2-choice instrumental conditioning task. We have shown that potentiation of cortico-striatal synapses enables learning of rewarded options. However, later these synapses become redundant as direct connections between prefrontal and premotor cortices (PFC-PMC) potentiate by Hebbian learning. The model shows that disease-related imbalances of the direct and indirect pathways in the BG impairs learning and suggests that these imbalances may also impede choices that have been learned previously, in spite of BG redundancy for those choices.

Our model of the parkinsonian state reproduces several major behavioral and electrophysiological features documented experimentally: First, the overall PMC activity is diminished in

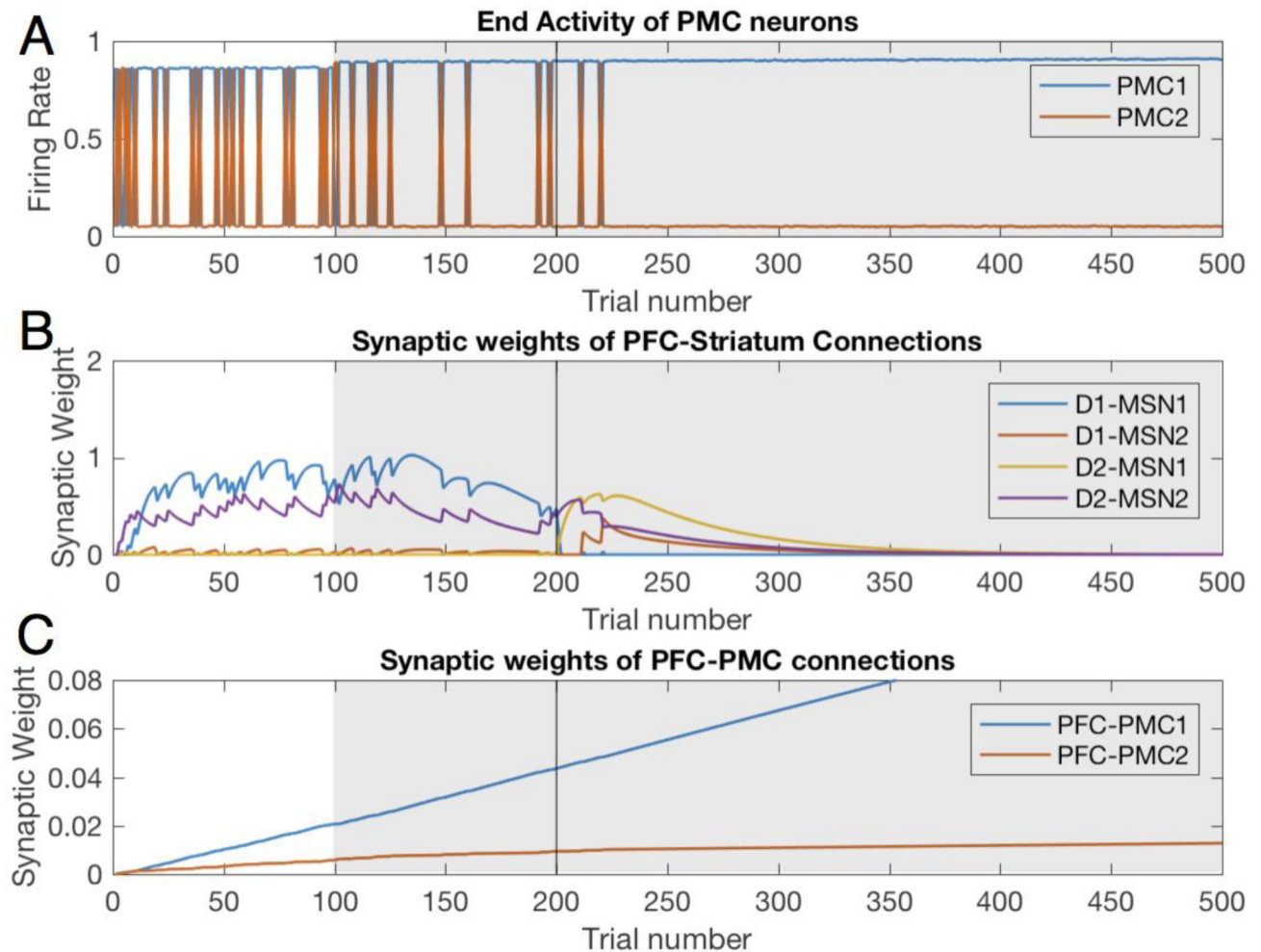


Fig 10. In the HD state, the random switches between choice 1 and 2 gradually cease after the treatment onset. Trial-by-trial dynamics of the PMC activity and underlying modulation of synaptic weights in the Huntington BG model with simulated treatment starting at trial 100. Same notation as in Fig 2. (A) The levels of PMC1 and PMC2 activity (choice 1 vs. 2) at the end of each trial (B) Synaptic weights of the PFC to striatum connections reflect rewarded choices. (C) Synaptic weight of the PFC to PMC1 connection keep growing after treatment onset, and during reversal.

<https://doi.org/10.1371/journal.pone.0228081.g010>

the PD state, consistent with PD studies [68]. Further, the model predicts that this activity is lowest at the beginning of the reversal due to aberrant engagement of the indirect pathway, which can be displayed as stronger bradykinesia and very low task performance scores. Reversal learning has been shown impaired in PD conditions [84–86]. Perseveration of the previously correct choice and impairment in maintaining of the new choice has been shown experimentally to contribute to the low performance [86], and our model reproduces these components as well. We have not found experimental evidence for the prediction of stronger bradykinesia. However, PD physiology is diverse and such bradykinesia may be evident at its more advanced stages. We tested the model for a range of parameters, and the duration of the choice blockade increases gradually as the model transitions from the healthy to the Parkinsonian state. Additionally, the block may be interrupted due to fluctuations in neuronal activity, and such perturbations as changes in the environment or forced choice trials would end the blockade phase. Second, the model shows robust oscillations in the activity of the cortico-BG-thalamo-cortical loop in the PD state. The frequency of these oscillations is about 6 Hz, which

is in the theta band. An increase in the EEG theta band is a marker of PD-related cognitive decline [87,88]. The oscillations are generated by a negative feedback branch of the loop through the indirect pathway as suggested before [40,89]. The hyperdirect pathway also contributes to this negative feedback and may support oscillations. Our simulations show that the oscillations cause multiple choice errors and, consequently, impede task performance and learning. Parkinsonian-state oscillations in the BG, although in the beta band and caused by a different mechanism, has been suggested to affect decision making by another model [90].

In the HD state, our model displays persistent randomly occurring choices of the unrewarded option, especially frequent after the reversal. This would register as impaired learning in behavioral tests, which is consistent with experimental results for cognitive [91,92] and motor tasks [93,94] in HD patients in the early stages of the disease. Furthermore, the model suggests that performance for previously learned tasks is also reduced by approximately 20%.

Therefore, our model reproduces impairments of the previously learned actions documented in BG-affecting diseases like PD and HD as well as after certain BG lesions [5,25,95]. However, surgical and DBS interventions in PD and HD patients do not impair, but rather restore motor function [26–28,96]. This raises the question: how can these two lines of evidence therefore be reconciled?

Learning in the model consists of two phases: BG-based and cortex-based. In a faster BG-based phase, the connections from PFC to MSNs are potentiated according to the RPE signal. The BG output inhibits choices with negative RPE and disinhibits those with positive RPE. Once the behavior is learned, the RPE vanishes, and the PFC-MSN connections decay to zero. The future choices are supported by the slower cortex-based learning phase: The connections from PFC directly to PMC are potentiated based on the Hebbian mechanism. Our simulations show that, even after the cortico-cortical connections increase to the levels ensuring robust choice of the rewarded option in the healthy state, both of the disease models are unable to make robust choices. Thus, behaviors that no longer need the BG are impaired. In accord with this result, 6-OHDA-lesioned rodents cannot maintain the correct choice, especially after reversal [86]. The model shows that it is an abnormal BG output that impairs the choices. Indeed, the BG output to the PMC does not vanish even when the behavior is learned and the BG no longer receives any RPE signal. In this case, due to the inputs from the PMC, the healthy BG disinhibits the previously learned choice, i.e. it conforms with the PFC-PMC associations. This disinhibitory function is impaired in both PD and HD, as well as after striatal lesions [5,25,95]. According to this prediction, disruption of the BG output would improve performance on previously learned tasks. Indeed, our model of a lesion of BG output demonstrates strengthening of performance on previously learned choices. Thus, GPi lesions were predominantly used in early surgical treatments of the PD, and sometimes are used now [97]. Additionally, DBS was successfully used in PD [26–28] and tested in HD patients [96]. The mechanism for DBS is not fully known, but thought to functionally lesion the excitatory input to GPi from STN and reduce GPi activity either by synaptic depletion or plasticity [40,98]. Therefore our model reconciles how specific GPi lesions that abolish BG output, or DBS that reduces the impact of this output, restore previously learned behaviors that were lost due to disrupted BG function, however this comes at the expense of decreased cognitive flexibility. A similar solution was suggested in an extensive computational study by Scholl and colleagues [36,41,99]. However, our model also combines the functional alterations with aberrant neural oscillations in PD.

The combination of the two learning mechanisms has been proposed and explored previously both experimentally and computationally [29,30,42]. Such combinations have been shown to be essential for cortical sensorimotor control, explained how reinforcement learning can shape cortical plasticity, and been used in brain-machine interface [100–102]. Here, we

demonstrated how cortical learning can be indirectly disrupted in PD and HD conditions. The three types of dynamics, healthy-, PD- and HD-like behavior persist in wide ranges of parameters in the model, whereas specific quantitative features, such as performance scores, show gradual parameter dependence. We tested multiple model manipulations, such as ablation of the hyperdirect pathway, or the STN-GPe feedback pathway to prove model robustness and mechanisms supporting dynamical properties of the model (data not shown). We showed that, in pathological states, ablation of the BG output may reveal hidden cortical learning and drastically improve performance. Cortical learning simply reflects the average of past choices regardless of the reward. If the switch to reversal occurs much earlier, the model predicts that cortical learning is not sufficiently engaged, and reversal takes fewer trials. By contrast, the model predicts that after sufficiently large number of rewarded trials, cortical learning may lock the choice even if it becomes non-rewarded, or even punished. Such aversion-resistant behavior is shown in substance abuse disorders [103,104]. The ability of the system to stop an unwanted behavior depends on the strength of the cortical vs. BG inputs to the PMC. To avoid aversion-resistant behavior for non-addictive, natural reinforcers, it's necessary to assume that the cortico-cortical synaptic plasticity is further limited to the low values achieved in our simulations. Homeostatic mechanisms that counteract Hebbian potentiation are plentiful [105], and malfunctioning of these mechanisms may, therefore, lead to aversion-resistant behaviors.

On the other hand, our model will forget a choice that was rewarded once tens of trials in the past due to the decay in the cortico-striatal connections. While there may be situations in which rare decisions are kept in the memory, which are not reproduced by the model, these situations are probably kept by other memory systems (e.g. emotional memory). Other limitations of our mode are mostly related to the firing rate formalism: First, the formalism does not allow for direct modeling of the eligibility traces necessary for stimulus-reward association, and we account for that by PFC activity that persists for the duration of the trial. Second, the D1 and D2 MSNs in the model reproduce their balanced excitation by cortical inputs in vivo [21,53]. The balance is supported by a number of complex mechanisms, from differential effects of DA on excitability of the D1 and D2 MSNs [54] to their lateral inhibition and contribution of striatal fast spiking interneurons [32]. These mechanisms are very hard to implement in a rate model, and we calibrate the D1 and D2 MSN equations identically to reflect the balance. The balance is perturbed in the PD DA depleted state (see [Methods](#)). Third, reciprocal inhibition between actions is implemented only at the PMC level, and omitted at the GPe level as STN-GPe network dynamics was shown to be only weakly dependent on this inhibition [45,46]. Fourth, the model does not differentiate between premotor cortex and thalamus because their interaction is a very complex separate direction. Finally, the oscillations that the model shows in the PD state are highly regular in spite of the noise added to all model components. The firing rate model is defined in terms of averaged firing rates of neural populations, and therefore, does not aim to reproduce all the noise. By contrast, it reproduces dynamical mechanisms that underlie robust deterministic processes that determine signal processing in the networks.

BG is suggested to be one of the main brain structure that determines action selection in multiple tasks and contexts. Hence, BG dysfunctions are shown to be linked to a broad spectrum of diseases, from Parkinson, to drug abuse. Traditionally, research effort on these diseases are disconnected from one another, even though they concern the same circuitry. Combining these efforts, in particular by modeling, will give us a more comprehensive picture of mechanisms involved in action selection at different levels of the brain circuitry. Modelling such complex mechanisms require connecting multiple brain regions, including cortical and subcortical. As a future direction, this model will be used as a building block in simulations of this circuitry. In particular, separating the dorsomedial and dorsolateral

striatal circuits (and, correspondingly, the cortical regions that project to these circuits), will allow one to address the development of goal-directed and habitual behavior in simulations as the two parts of the striatal circuitry are associated with the two distinct types of behavior [106]. One can further separate thalamic and cortical circuits to take into account the contribution of their interaction to action selection [107]. Finally, the model may be generalized into more complex tasks with multiple stimulus-response mappings. Thus, the simplicity of our model allows for qualitative explanation of mechanisms and, simultaneously, building large scale models that involve multiple brain regions.

Altogether, we have modeled the function of the cortico-BG-thalamo-cortical loop in a 2 choice instrumental conditioning task and shown that disbalance of the direct and indirect pathways is the mechanism by which this function is disrupted in HD and PD conditions. The model predicts that, after long training, direct cortico-cortical connections are responsible for the choices, and the cortico-BG-thalamo-cortical loop conforms to previously learned choices. The model also predicts that reversal is easier to achieve after short training of the initial contingency, and may be greatly impeded after very large number of repetitions of the initially rewarded choice. We have predicted how in pathological states, when BG impedes these choices, GPi lesion or DBS restores them, but completely disrupts learning of new behavior. Along with other computational studies [36,98,99], our results further reconcile the apparent contradiction between the critical involvement of the BG in execution of previously learned actions and yet no impairment of these actions after BG output is ablated by lesions or reduced by DBS.

Acknowledgments

We are grateful to Profs. Molkov, Gurney, Czachowski, Graham, Worth, and Rubchinsky for useful discussions.

Author Contributions

Conceptualization: Brady Atwood, Alexey Kuznetsov.

Data curation: Brady Atwood.

Formal analysis: Garrett Mulcahy.

Investigation: Garrett Mulcahy, Alexey Kuznetsov.

Methodology: Alexey Kuznetsov.

Resources: Alexey Kuznetsov.

Software: Garrett Mulcahy.

Supervision: Alexey Kuznetsov.

Validation: Brady Atwood.

Visualization: Garrett Mulcahy.

Writing – original draft: Garrett Mulcahy.

Writing – review & editing: Brady Atwood, Alexey Kuznetsov.

References

1. Grillner S, Robertson B, Stephenson-Jones M. The evolutionary origin of the vertebrate basal ganglia and its role in action selection. *J Physiol.* 2013 Nov 15; 591(22):5425–31. <https://doi.org/10.1113/jphysiol.2012.246660> PMID: 23318875

2. Nelson AB, Kreitzer AC. Reassessing models of basal ganglia function and dysfunction. *Annu Rev Neurosci*. 2014; 37:117–35. <https://doi.org/10.1146/annurev-neuro-071013-013916> PMID: 25032493
3. Bhatia KP, Marsden CD. The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain*. 1994; 117(4):859–76.
4. DeLong MR. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci*. 1990 Jul; 13(7):281–5. [https://doi.org/10.1016/0166-2236\(90\)90110-v](https://doi.org/10.1016/0166-2236(90)90110-v) PMID: 1695404
5. Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, et al. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat Rev Neurosci*. 2010 Nov; 11(11):760–72. <https://doi.org/10.1038/nrn2915> PMID: 20944662
6. Packard MG, Knowlton BJ. Learning and Memory Functions of the Basal Ganglia. *Annu Rev Neurosci*. 2002 Mar; 25(1):563–93.
7. Foerde K, Shohamy D. The role of the basal ganglia in learning and memory: insight from Parkinson's disease. *Neurobiol Learn Mem*. 2011 Nov; 96(4):624–36. <https://doi.org/10.1016/j.nlm.2011.08.006> PMID: 21945835
8. Graybiel AM. Habits, Rituals, and the Evaluative Brain. *Annu Rev Neurosci*. 2008 Jul; 31(1):359–87.
9. Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci*. 2011 Feb; 14(2):154–62. <https://doi.org/10.1038/nn.2723> PMID: 21270784
10. Keiflin R, Janak PH. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*. 2015 Oct; 88(2):247–63. <https://doi.org/10.1016/j.neuron.2015.08.037> PMID: 26494275
11. Frank MJ. Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism. *J Cogn Neurosci*. 2005 Jan; 17(1):51–72. <https://doi.org/10.1162/0898929052880093> PMID: 15701239
12. Schultz W. Updating dopamine reward signals. *Curr Opin Neurobiol*. 2013 Apr; 23(2):229–38. <https://doi.org/10.1016/j.conb.2012.11.012> PMID: 23267662
13. Takahashi Y, Schoenbaum G, Niv Y. Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. *Front Neurosci*. 2008 Jul; 2(1):86–99. <https://doi.org/10.3389/neuro.01.014.2008> PMID: 18982111
14. Surmeier DJ, Plotkin J, Shen W. Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. *Curr Opin Neurobiol*. 2009 Dec; 19(6):621–8. <https://doi.org/10.1016/j.conb.2009.10.003> PMID: 19896832
15. Bamford NS, Wightman RM, Sulzer D. Dopamine's Effects on Corticostriatal Synapses during Reward-Based Behaviors. *Neuron*. 2018 Feb 7; 97(3):494–510. <https://doi.org/10.1016/j.neuron.2018.01.006> PMID: 29420932
16. Gerfen CR. The Neostriatal Mosaic: Multiple Levels of Compartmental Organization in the Basal Ganglia. *Annu Rev Neurosci*. 1992 Mar; 15(1):285–320.
17. Surmeier DJ, Song WJ, Yan Z. Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J Neurosci Off J Soc Neurosci*. 1996 Oct 15; 16(20):6579–91.
18. Kreitzer AC, Malenka RC. Striatal Plasticity and Basal Ganglia Circuit Function. *Neuron*. 2008 Nov; 60(4):543–54. <https://doi.org/10.1016/j.neuron.2008.11.005> PMID: 19038213
19. Alexander GE, Crutcher MD. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci*. 1990 Jul; 13(7):266–71. [https://doi.org/10.1016/0166-2236\(90\)90107-I](https://doi.org/10.1016/0166-2236(90)90107-I) PMID: 1695401
20. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol Cybern*. 2001 May 11; 84(6):401–10. <https://doi.org/10.1007/PL00007984> PMID: 11417052
21. Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM, et al. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature*. 2013 Feb 14; 494(7436):238–42. <https://doi.org/10.1038/nature11846> PMID: 23354054
22. Markowitz JE, Gillis WF, Beron CC, Neufeld SQ, Robertson K, Bhagat ND, et al. The Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection. *Cell*. 2018 28; 174(1):44–58.e17. <https://doi.org/10.1016/j.cell.2018.04.019> PMID: 29779950
23. London TD, Licholai JA, Szczot I, Ali MA, LeBlanc KH, Fobbs WC, et al. Coordinated Ramping of Dorsal Striatal Pathways preceding Food Approach and Consumption. *J Neurosci Off J Soc Neurosci*. 2018 Apr 4; 38(14):3547–58.
24. Lanciego JL, Luquin N, Obeso JA. Functional neuroanatomy of the basal ganglia. *Cold Spring Harb Perspect Med*. 2012 Dec 1; 2(12):a009621. <https://doi.org/10.1101/cshperspect.a009621> PMID: 23071379

25. Balleine BW, Liljeholm M, Ostlund SB. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res*. 2009 Apr; 199(1):43–52. <https://doi.org/10.1016/j.bbr.2008.10.034> PMID: 19027797
26. Munhoz RP, Cerasa A, Okun MS. Surgical Treatment of Dyskinesia in Parkinson's Disease. *Front Neurol* [Internet]. 2014 Apr 29 [cited 2019 Mar 8]; 5. <http://journal.frontiersin.org/article/10.3389/fneur.2014.00065/abstract>
27. Brown P, Eusebio A. Paradoxes of functional neurosurgery: clues from basal ganglia recordings. *Mov Disord Off J Mov Disord Soc*. 2008 Jan; 23(1):12–20; quiz 158.
28. Marsden CD, Obeso JA. The functions of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. *Brain J Neurol*. 1994 Aug; 117(Pt 4):877–97.
29. Piron C, Kase D, Topalidou M, Goillandeau M, Orignac H, N'Guyen T-H, et al. The globus pallidus pars interna in goal-oriented and routine behaviors: Resolving a long-standing paradox: GPI, Goal-directed behaviors and habits. *Mov Disord*. 2016 Aug; 31(8):1146–54. <https://doi.org/10.1002/mds.26542> PMID: 26900137
30. Topalidou M, Kase D, Boraud T, Rougier NP. A Computational Model of Dual Competition between the Basal Ganglia and the Cortex. *eneuro*. 2018; 5(6):ENEURO.0339-17.2018.
31. Humphries M. Capturing dopaminergic modulation and bimodal membrane behaviour of striatal medium spiny neurons in accurate, reduced models. *Front Comput Neurosci* [Internet]. 2009 [cited 2017 Dec 31]; 3. <http://journal.frontiersin.org/article/10.3389/fncom.2009/abstract>
32. Damodaran S, Evans RC, Blackwell KT. Synchronized firing of fast-spiking interneurons is critical to maintain balanced firing between direct and indirect pathway neurons of the striatum. *J Neurophysiol*. 2014 Feb 15; 111(4):836–48. <https://doi.org/10.1152/jn.00382.2013> PMID: 24304860
33. Reed MC, Best J, Nijhout HF. Passive and active stabilization of dopamine in the striatum. *Biosci Hypotheses*. 2009 Jan; 2(4):240–4.
34. Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD. Influence of Phasic and Tonic Dopamine Release on Receptor Activation. *J Neurosci*. 2010 Oct 20; 30(42):14273–83. <https://doi.org/10.1523/JNEUROSCI.1894-10.2010> PMID: 20962248
35. Humphries MD, Stewart RD, Gurney KN. A Physiologically Plausible Model of Action Selection and Oscillatory Activity in the Basal Ganglia. *J Neurosci*. 2006 Dec 13; 26(50):12921–42. <https://doi.org/10.1523/JNEUROSCI.3486-06.2006> PMID: 17167083
36. Schroll H, Vitay J, Hamker FH. Dysfunctional and compensatory synaptic plasticity in Parkinson's disease. *Eur J Neurosci*. 2014 Feb; 39(4):688–702. <https://doi.org/10.1111/ejn.12434> PMID: 24313650
37. Lindahl M, Hellgren Kotaleski J. Untangling Basal Ganglia Network Dynamics and Function: Role of Dopamine Depletion and Inhibition Investigated in a Spiking Network Model. *eneuro*. 2016 Nov; 3(6):ENEURO.0156-16.2016.
38. Leblois A. Competition between Feedback Loops Underlies Normal and Pathological Dynamics in the Basal Ganglia. *J Neurosci*. 2006 Mar 29; 26(13):3567–83. <https://doi.org/10.1523/JNEUROSCI.5050-05.2006> PMID: 16571765
39. Chakravarthy VS, Joseph D, Bapi RS. What do the basal ganglia do? A modeling perspective. *Biol Cybern*. 2010 Sep; 103(3):237–53. <https://doi.org/10.1007/s00422-010-0401-y> PMID: 20644953
40. Humphries MD, Obeso JA, Dreyer JK. Insights into Parkinson's disease from computational models of the basal ganglia. *J Neurol Neurosurg Psychiatry*. 2018 Nov; 89(11):1181–8. <https://doi.org/10.1136/jnnp-2017-315922> PMID: 29666208
41. Schroll H, Hamker FH. Basal Ganglia dysfunctions in movement disorders: What can be learned from computational simulations: Basal Ganglia Defects in Movement Disorders. *Mov Disord*. 2016 Nov; 31(11):1591–601. <https://doi.org/10.1002/mds.26719> PMID: 27393040
42. Kim T, Hamade KC, Todorov D, Barnett WH, Capps RA, Latash EM, et al. Reward Based Motor Adaptation Mediated by Basal Ganglia. *Front Comput Neurosci* [Internet]. 2017 Mar 31 [cited 2019 Mar 8]; 11. <http://journal.frontiersin.org/article/10.3389/fncom.2017.00019/full>
43. Reynolds JN, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. *Nature*. 2001 Sep 6; 413(6851):67–70. <https://doi.org/10.1038/35092560> PMID: 11544526
44. Buonomano DV, Merzenich MM. CORTICAL PLASTICITY: From Synapses to Maps. *Annu Rev Neurosci*. 1998 Mar; 21(1):149–86.
45. Wolters A, Sandbrink F, Schlottmann A, Kunesch E, Stefan K, Cohen LG, et al. A Temporally Asymmetric Hebbian Rule Governing Plasticity in the Human Motor Cortex. *J Neurophysiol*. 2003 May; 89(5):2339–45. <https://doi.org/10.1152/jn.00900.2002> PMID: 12612033
46. Koch C, Segev I, editors. *Methods in neuronal modeling: from ions to networks*. 2. ed., 3. printing. Cambridge, Mass.: MIT Press; 2001. 671 p. (Computational neuroscience).

47. Best J, Park C, Terman D, Wilson C. Transitions between irregular and rhythmic firing patterns in excitatory-inhibitory neuronal networks. *J Comput Neurosci*. 2007 Aug 22; 23(2):217–35. <https://doi.org/10.1007/s10827-007-0029-7> PMID: 17624604
48. Park C, Worth RM, Rubchinsky LL. Neural dynamics in Parkinsonian brain: The boundary between synchronized and nonsynchronized dynamics. *Phys Rev E [Internet]*. 2011 Apr 13 [cited 2019 Aug 24]; 83(4). <https://link.aps.org/doi/10.1103/PhysRevE.83.042901>
49. Pavlides A, Hogan SJ, Bogacz R. Computational Models Describing Possible Mechanisms for Generation of Excessive Beta Oscillations in Parkinson's Disease. Graham LJ, editor. *PLOS Comput Biol*. 2015 Dec 18; 11(12):e1004609. <https://doi.org/10.1371/journal.pcbi.1004609> PMID: 26683341
50. Schultz W. Behavioral dopamine signals. *Trends Neurosci*. 2007 May; 30(5):203–10. <https://doi.org/10.1016/j.tins.2007.03.007> PMID: 17400301
51. Gurney KN, Humphries MD, Redgrave P. A New Framework for Cortico-Striatal Plasticity: Behavioural Theory Meets In Vitro Data at the Reinforcement-Action Interface. Dayan P, editor. *PLoS Biol*. 2015 Jan 6; 13(1):e1002034. <https://doi.org/10.1371/journal.pbio.1002034> PMID: 25562526
52. Shindou T, Shindou M, Watanabe S, Wickens J. A silent eligibility trace enables dopamine-dependent synaptic plasticity for reinforcement learning in the mouse striatum. *Eur J Neurosci*. 2019 Mar; 49(5):726–36. <https://doi.org/10.1111/ejn.13921> PMID: 29603470
53. Fisher SD, Robertson PB, Black MJ, Redgrave P, Sagar MA, Abraham WC, et al. Reinforcement determines the timing dependence of corticostriatal synaptic plasticity in vivo. *Nat Commun [Internet]*. 2017 Dec [cited 2019 Aug 24]; 8(1). <http://www.nature.com/articles/s41467-017-00394-x>
54. Baladron J, Nambu A, Hamker FH. The subthalamic nucleus-external globus pallidus loop biases exploratory decisions towards known alternatives: a neuro-computational study. *Eur J Neurosci*. 2019 Mar; 49(6):754–67. <https://doi.org/10.1111/ejn.13666> PMID: 28833676
55. Mallet N. Cortical Inputs and GABA Interneurons Imbalance Projection Neurons in the Striatum of Parkinsonian Rats. *J Neurosci*. 2006 Apr 5; 26(14):3875–84. <https://doi.org/10.1523/JNEUROSCI.4439-05.2006> PMID: 16597742
56. Moyer JT, Wolf JA, Finkel LH. Effects of Dopaminergic Modulation on the Integrative Properties of the Ventral Striatal Medium Spiny Neuron. *J Neurophysiol*. 2007 Dec; 98(6):3731–48. <https://doi.org/10.1152/jn.00335.2007> PMID: 17913980
57. Kalia LV, Lang AE. Parkinson's disease. *The Lancet*. 2015 Aug; 386(9996):896–912.
58. Schapira AHV. Neurobiology and treatment of Parkinson's disease. *Trends Pharmacol Sci*. 2009 Jan; 30(1):41–7. <https://doi.org/10.1016/j.tips.2008.10.005> PMID: 19042040
59. Kita H, Kita T. Role of Striatum in the Pause and Burst Generation in the Globus Pallidus of 6-OHDA-Treated Rats. *Front Syst Neurosci [Internet]*. 2011 [cited 2019 Mar 11]; 5. <http://journal.frontiersin.org/article/10.3389/fnsys.2011.00042/abstract>
60. Miller WC, DeLong MR. Altered Tonic Activity of Neurons in the Globus Pallidus and Subthalamic Nucleus in the Primate MPTP Model of Parkinsonism. In: Carpenter MB, Jayaraman A, editors. *The Basal Ganglia II [Internet]*. Boston, MA: Springer US; 1987 [cited 2019 Mar 11]. p. 415–27. http://link.springer.com/10.1007/978-1-4684-5347-8_29
61. Pan HS, Walters JR. Unilateral lesion of the nigrostriatal pathway decreases the firing rate and alters the firing pattern of globus pallidus neurons in the rat. *Synapse*. 1988; 2(6):650–6. <https://doi.org/10.1002/syn.890020612> PMID: 3145582
62. Sterio D, Berić A, Dogali M, Fazzini E, Alfaro G, Devinsky O. Neurophysiological properties of pallidal neurons in Parkinson's disease. *Ann Neurol*. 1994 May; 35(5):586–91. <https://doi.org/10.1002/ana.410350512> PMID: 8179304
63. Doudet DJ, Gross C, Arluison M, Bioulac B. Modifications of precentral cortex discharge and EMG activity in monkeys with MPTP-induced lesions of DA nigral neurons. *Exp Brain Res*. 1990; 80(1):177–88. <https://doi.org/10.1007/bf00228859> PMID: 1972680
64. Bergman H, Wichmann T, Karmon B, DeLong MR. The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of parkinsonism. *J Neurophysiol*. 1994 Aug; 72(2):507–20. <https://doi.org/10.1152/jn.1994.72.2.507> PMID: 7983515
65. Hassani OK, Mouroux M, Féger J. Increased subthalamic neuronal activity after nigral dopaminergic lesion independent of disinhibition via the globus pallidus. *Neuroscience*. 1996 May; 72(1):105–15. [https://doi.org/10.1016/0306-4522\(95\)00535-8](https://doi.org/10.1016/0306-4522(95)00535-8) PMID: 8730710
66. Hutchison WD, Lozano AM, Davis KD, Saint-Cyr JA, Lang AE, Dostrovsky JO. Differential neuronal activity in segments of globus pallidus in Parkinson's disease patients. *Neuroreport*. 1994 Jul 21; 5(12):1533–7. <https://doi.org/10.1097/00001756-199407000-00031> PMID: 7948856
67. Wichmann T, Bergman H, Starr PA, Subramanian T, Watts RL, DeLong MR. Comparison of MPTP-induced changes in spontaneous neuronal discharge in the internal pallidal segment and in the

- substantia nigra pars reticulata in primates. *Exp Brain Res*. 1999 Apr; 125(4):397–409. <https://doi.org/10.1007/s002210050696> PMID: 10323285
68. Galvan A, Devergnas A, Wichmann T. Alterations in neuronal activity in basal ganglia-thalamocortical circuits in the parkinsonian state. *Front Neuroanat* [Internet]. 2015 Feb 5 [cited 2019 Mar 11]; 9. <http://journal.frontiersin.org/Article/10.3389/fnana.2015.00005/abstract>
 69. Cooper AJ, Stanford IM. Dopamine D2 receptor mediated presynaptic inhibition of striatopallidal GABAA IPSCs in vitro. *Neuropharmacology*. 2001 Jul; 41(1):62–71. [https://doi.org/10.1016/s0028-3908\(01\)00038-7](https://doi.org/10.1016/s0028-3908(01)00038-7) PMID: 11445186
 70. Stanford IM, Cooper AJ. Presynaptic mu and delta opioid receptor modulation of GABAA IPSCs in the rat globus pallidus in vitro. *J Neurosci Off J Soc Neurosci*. 1999 Jun 15; 19(12):4796–803.
 71. Ogura M, Kita H. Dynorphin Exerts Both Postsynaptic and Presynaptic Effects in the Globus Pallidus of the Rat. *J Neurophysiol*. 2000 Jun; 83(6):3366–76. <https://doi.org/10.1152/jn.2000.83.6.3366> PMID: 10848555
 72. Shen K-Z, Johnson SW. Presynaptic dopamine D₂ and muscarine M₃ receptors inhibit excitatory and inhibitory transmission to rat subthalamic neurones *in vitro*. *J Physiol*. 2000 Jun; 525(2):331–41.
 73. Shen K-Z, Zhu Z-T, Munhall A, Johnson SW. Dopamine receptor supersensitivity in rat subthalamus after 6-hydroxydopamine lesions. *Eur J Neurosci*. 2003 Dec; 18(11):2967–74. <https://doi.org/10.1046/j.1460-9568.2003.03058.x> PMID: 14656292
 74. Floran B, Floran L, Eriji D, Aceves J. Dopamine D4 receptors inhibit depolarization-induced [3H]GABA release in the rat subthalamic nucleus. *Eur J Pharmacol*. 2004 Sep; 498(1–3):97–102. <https://doi.org/10.1016/j.ejphar.2004.07.078> PMID: 15363981
 75. Rommelfanger KS. Extrastriatal dopaminergic circuits of the basal ganglia. *Front Neuroanat* [Internet]. 2010 [cited 2019 Aug 24]; 4. <http://journal.frontiersin.org/article/10.3389/fnana.2010.00139/abstract>
 76. Reiner A, Dragatsis I, Dietrich P. Genetics and Neuropathology of Huntington's Disease. In: *International Review of Neurobiology* [Internet]. Elsevier; 2011 [cited 2019 Mar 11]. p. 325–72. <https://linkinghub.elsevier.com/retrieve/pii/B9780123813282000146> <https://doi.org/10.1016/B978-0-12-381328-2.00014-6> PMID: 21907094
 77. Sharp AH, Ross CA. Neurobiology of Huntington's Disease. *Neurobiol Dis*. 1996 Feb; 3(1):3–15. <https://doi.org/10.1006/nbdi.1996.0002> PMID: 9173909
 78. Cohen JD, McClure SM, Yu AJ. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos Trans R Soc B Biol Sci*. 2007 May 29; 362(1481):933–42.
 79. Prescott TJ, Bryson JJ, Seth AK. Introduction. Modelling natural action selection. *Philos Trans R Soc B Biol Sci*. 2007 Sep 29; 362(1485):1521–9.
 80. Klaus A, Martins GJ, Paixao VB, Zhou P, Paninski L, Costa RM. The Spatiotemporal Organization of the Striatum Encodes Action Space. *Neuron*. 2017 Aug; 95(5):1171–1180.e7. <https://doi.org/10.1016/j.neuron.2017.08.015> PMID: 28858619
 81. Hulse SH, Egeth H, Deese J. *The psychology of learning*. 5th ed. New York: McGraw-Hill; 1980. 478 p. (McGraw-Hill series in psychology).
 82. Izquierdo A, Wiedholz LM, Millstein RA, Yang RJ, Bussey TJ, Saksida LM, et al. Genetic and dopaminergic modulation of reversal learning in a touchscreen-based operant procedure for mice. *Behav Brain Res*. 2006 Aug 10; 171(2):181–8. <https://doi.org/10.1016/j.bbr.2006.03.029> PMID: 16713639
 83. Linden J, James AS, McDaniel C, Jentsch JD. Dopamine D2 Receptors in Dopaminergic Neurons Modulate Performance in a Reversal Learning Task in Mice. *eNeuro*. 2018 Feb; 5(1).
 84. Cools R, Altamirano L, D'Esposito M. Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*. 2006 Jan; 44(10):1663–73. <https://doi.org/10.1016/j.neuropsychologia.2006.03.030> PMID: 16730032
 85. Peterson DA, Elliott C, Song DD, Makeig S, Sejnowski TJ, Poizner H. Probabilistic reversal learning is impaired in Parkinson's disease. *Neuroscience*. 2009 Nov; 163(4):1092–101. <https://doi.org/10.1016/j.neuroscience.2009.07.033> PMID: 19628022
 86. Grospe GM, Baker PM, Ragozzino ME. Cognitive Flexibility Deficits Following 6-OHDA Lesions of the Rat Dorsomedial Striatum. *Neuroscience*. 2018 Mar; 374:80–90. <https://doi.org/10.1016/j.neuroscience.2018.01.032> PMID: 29374536
 87. Soikkeli R, Partanen J, Soininen H, Pääkkönen A, Riekkinen P. Slowing of EEG in Parkinson's disease. *Electroencephalogr Clin Neurophysiol*. 1991 Sep; 79(3):159–65. [https://doi.org/10.1016/0013-4694\(91\)90134-p](https://doi.org/10.1016/0013-4694(91)90134-p) PMID: 1714807
 88. Cozac VV, Gschwandtner U, Hatz F, Hardmeier M, Rüegg S, Fuhr P. Quantitative EEG and Cognitive Decline in Parkinson's Disease. *Park Dis*. 2016; 2016:1–14.

89. van Albada SJ, Gray RT, Drysdale PM, Robinson PA. Mean-field modeling of the basal ganglia-thalamocortical system. II. *J Theor Biol.* 2009 Apr; 257(4):664–88. <https://doi.org/10.1016/j.jtbi.2008.12.013> PMID: 19154745
90. Wei W, Rubin JE, Wang X-J. Role of the Indirect Pathway of the Basal Ganglia in Perceptual Decision Making. *J Neurosci.* 2015 Mar 4; 35(9):4052–64. <https://doi.org/10.1523/JNEUROSCI.3611-14.2015> PMID: 25740532
91. Grahn JA, Parkinson JA, Owen AM. The role of the basal ganglia in learning and memory: Neuropsychological studies. *Behav Brain Res.* 2009 Apr; 199(1):53–60. <https://doi.org/10.1016/j.bbr.2008.11.020> PMID: 19059285
92. Lawrence AD, Hodges JR, Rosser AE, Kershaw A, French-Constant C, Rubinsztein DC, et al. Evidence for specific cognitive deficits in preclinical Huntington's disease. *Brain J Neurol.* 1998 Jul; 121 (Pt 7):1329–41.
93. Gutierrez-Garralda JM, Moreno-Briseño P, Boll M-C, Morgado-Valle C, Campos-Romo A, Diaz R, et al. The effect of Parkinson's disease and Huntington's disease on human visuomotor learning. *Eur J Neurosci.* 2013 Jun;n/a–n/a.
94. Heindel WC, Butters N, Salmon DP. Impaired learning of a motor skill in patients with Huntington's disease. *Behav Neurosci.* 1988 Feb; 102(1):141–7. <https://doi.org/10.1037//0735-7044.102.1.141> PMID: 2965592
95. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci.* 2004 Jan; 19(1):181–9. <https://doi.org/10.1111/j.1460-9568.2004.03095.x> PMID: 14750976
96. Wojtecki L, Groiss SJ, Hartmann CJ, Elben S, Omlor S, Schnitzler A, et al. Deep Brain Stimulation in Huntington's Disease—Preliminary Evidence on Pathophysiology, Efficacy and Safety. *Brain Sci.* 2016 Aug 30; 6(3).
97. Dostrovsky JO, Hutchison WD, Lozano AM. The Globus Pallidus, Deep Brain Stimulation, and Parkinson's Disease. *The Neuroscientist.* 2002 Jun; 8(3):284–90. <https://doi.org/10.1177/1073858402008003014> PMID: 12061508
98. Rubin JE. Computational models of basal ganglia dysfunction: the dynamics is in the details. *Curr Opin Neurobiol.* 2017 Oct; 46:127–35. <https://doi.org/10.1016/j.conb.2017.08.011> PMID: 28888856
99. Schroll H, Beste C, Hamker FH. 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.* 2015 May; 41(9):1227–44. <https://doi.org/10.1111/ejn.12868> PMID: 25778633
100. Marsh BT, Tarigoppula VSA, Chen C, Francis JT. Toward an Autonomous Brain Machine Interface: Integrating Sensorimotor Reward Modulation and Reinforcement Learning. *J Neurosci.* 2015 May 13; 35(19):7374–87. <https://doi.org/10.1523/JNEUROSCI.1802-14.2015> PMID: 25972167
101. An J, Yadav T, Hessburg JP, Francis JT. Reward Expectation Modulates Local Field Potentials, Spiking Activity and Spike-Field Coherence in the Primary Motor Cortex. *eneuro.* 2019 May; 6(3):ENEURO.0178-19.2019.
102. Neymotin SA, Chadderdon GL, Kerr CC, Francis JT, Lytton WW. Reinforcement Learning of Two-Joint Virtual Arm Reaching in a Computer Model of Sensorimotor Cortex. *Neural Comput.* 2013 Dec; 25(12):3263–93. https://doi.org/10.1162/NECO_a_00521 PMID: 24047323
103. Hopf FW, Lesscher HMB. Rodent models for compulsive alcohol intake. *Alcohol Fayettev N.* 2014 May; 48(3):253–64.
104. Hogarth L, Balleine BW, Corbit LH, Killcross S. Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Ann N Y Acad Sci.* 2013 Apr; 1282:12–24. <https://doi.org/10.1111/j.1749-6632.2012.06768.x> PMID: 23126270
105. Turrigiano GG, Nelson SB. Homeostatic plasticity in the developing nervous system. *Nat Rev Neurosci.* 2004 Feb; 5(2):97–107. <https://doi.org/10.1038/nrn1327> PMID: 14735113
106. Graybiel AM, Grafton ST. The Striatum: Where Skills and Habits Meet. *Cold Spring Harb Perspect Biol.* 2015 Aug; 7(8):a021691. <https://doi.org/10.1101/cshperspect.a021691> PMID: 26238359
107. Rikhye RV, Wimmer RD, Halassa MM. Toward an Integrative Theory of Thalamic Function. *Annu Rev Neurosci.* 2018 Jul 8; 41(1):163–83.