EJN European Journal of Neuroscience

FENS Federation of European Neuroscience Societies

European Journal of Neuroscience, Vol. 41, pp. 1227–1244, 2015 doi:10.1111/ejn.12868

CLINICAL AND TRANSLATIONAL NEUROSCIENCE

Combined lesions of direct and indirect basal ganglia pathways but not changes in dopamine levels explain learning deficits in patients with Huntington's disease

Henning Schroll,[1,2,3,4] Christian Beste[5] and Fred H. Hamker[2,4]
[1]Neurology, Charité – Universitätsmedizin Berlin, Berlin, Germany
[2]Bernstein Center for Computational Neuroscience, Charité – Universitätsmedizin Berlin, Berlin, Germany
[3]Psychology, Humboldt Universität zu Berlin, Berlin, Germany
[4]Computer Science, Chemnitz University of Technology, Straße der Nationen 62, 09111 Chemnitz, Germany
[5]Cognitive Neurophysiology, Department of Child and Adolescent Psychiatry, Faculty of Medicine of the TU Dresden, Dresden, Germany

Keywords: computational modeling, dopamine, neurodegeneration, stimulus–response learning, striatum

Abstract

Huntington's disease (HD) is a hereditary neurodegenerative disease of the basal ganglia that causes severe motor, cognitive and emotional dysfunctions. In the human basal ganglia, these dysfunctions are accompanied by a loss of striatal medium spiny neurons, dysfunctions of the subthalamic nucleus and globus pallidus, and changes in dopamine receptor binding. Here, we used a neuro-computational model to investigate which of these basal ganglia dysfunctions can explain patients' deficits in different stimulus–response learning paradigms. We show that these paradigms are particularly suitable for scrutinising the effects of potential changes in dopamine signaling and of potential basal ganglia lesions on overt behavior in HD. We find that combined lesions of direct and indirect basal ganglia pathways, but none of these lesions alone, reproduce patients' learning impairments. Degeneration of medium spiny neurons of the direct pathway accounts for patients' deficits in facilitating correct responses, whereas degeneration of indirect pathway medium spiny neurons explains their impairments in inhibiting dominant but incorrect responses. The empirical results cannot be explained by lesions of the subthalamic nucleus, which is part of the hyperdirect pathway, or by changes in dopamine levels. Overall, our simulations suggest combined lesions of direct and indirect pathways as a major source of HD patients' learning impairments and, tentatively, also their motor and cognitive deficits in general, whereas changes in dopamine levels are suggested to not be causally related to patients' impairments.

Introduction

Huntington's disease (HD) is an autosomal-dominant hereditary disease of the basal ganglia that causes progressive neuronal decline and leads to severe motor and cognitive impairments. The neuronal origins of these impairments have not yet been clearly identified, but a loss of striatal medium spiny neurons (MSNs) has been proposed to be critical (DeLong, 1990; Raymond *et al.*, 2011; Ehrlich, 2012). These striatal MSNs receive afferents from the cortex, and project to either the globus pallidus internus (GPi) and substantia nigra reticulata (SNr) or to the globus pallidus externus (GPe). Those MSNs that project to the GPi/SNr belong to the 'direct' basal ganglia pathway, which, via the thalamus, excites the cortex (Fig. 1A). Those MSNs that project to the GPe, in contrast, belong to the 'indirect' basal ganglia pathway, which inhibits cortical activity (Fig. 1A). In the presymptomatic and early symptomatic stages of HD, neuronal degeneration primarily affects indirect pathway MSNs (Albin *et al.*, 1992), but, later in the disease, direct pathway MSNs also deteriorate (Reiner *et al.*, 1988). As well as direct and indirect pathways, the basal ganglia have a third, 'hyperdirect', pathway (Fig. 1A) that proceeds to the GPi/SNr via the subthalamic nucleus (STN). Like the indirect pathway, this pathway inhibits cortical activity, but it simultaneously affects larger numbers of cortical neurons (Hazrati & Parent, 1992a,b; cf. Schroll & Hamker, 2013) and probably has different mechanisms of synaptic plasticity (Schroll *et al.*, 2014).

HD neurodegeneration affects not only the striatum, but also the STN and the globus pallidus (O'Kusky *et al.*, 1999; Tabrizi *et al.*, 2012, 2013; Delmaire *et al.*, 2013), suggesting degeneration of the hyperdirect pathway as another potential cause of patients' impairments. Moreover, changes in dopamine receptor binding have been reported (Antonini *et al.*, 1996; Weeks *et al.*, 1996; Glass *et al.*, 2000; Van Oostrom *et al.*, 2005).

Motor and cognitive deficits of HD patients have been described according to a variety of neuropsychological paradigms. They are related to learning (Lawrence *et al.*, 1999; Brown *et al.*, 2001;

Correspondence: Henning Schroll and Fred H. Hamker, [4]Computer Science, as above.
E-mails: henning.schroll@informatik.tu-chemnitz.de; fred.hamker@informatik.tu-chemnitz.de

Received 22 October 2014, revised 17 January 2015, accepted 6 February 2015

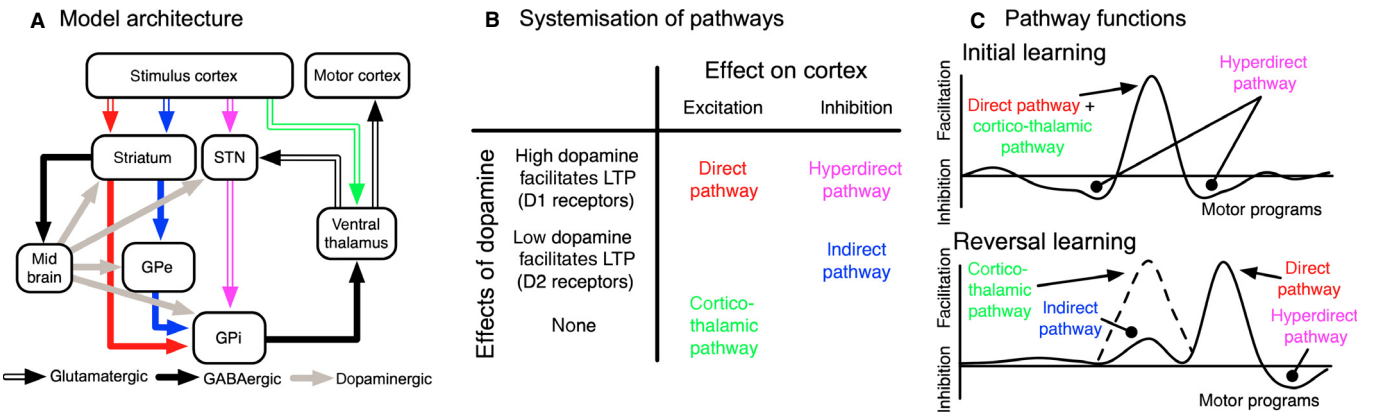


FIG. 1. Overview of the anatomy and functionality of our neuro-computational model (Schroll *et al.*, 2014). (A) Model architecture. The model contains several distinct layers corresponding to the cortex, basal ganglia, and thalamus. Each of these layers comprises a predefined number of artificial neurons (Table 1) that are interconnected in agreement with anatomical evidence (Braak & Del Tredici, 2008). The direct pathway is shown in red, the indirect pathway in blue, the hyperdirect pathway in magenta, and the cortico-thalamic pathway in green. (B) Scheme of relevant properties of basal ganglia pathways in the model. Each of these pathways either excites or inhibits the motor cortex. Moreover, dopamine either facilitates either LTP or LTD in each of these pathways. (C) Pathway functions in reward-based stimulus–response learning as they emerge from the scheme shown in B. The *y*-axis shows activation of cortical motor programs, where values above zero imply facilitation of motor programs and values below zero imply inhibition of motor programs. Headed arrows denote excitatory effects of basal ganglia pathways on motor program activity, and pointed arrows denote inhibitory effects.

Filoteo *et al.*, 2001), executive functions (Peinemann *et al.*, 2005; Beste *et al.*, 2007, 2008b, 2012, 2013), attention (Montoya *et al.*, 2006b), Go–NoGo performance (Sprengelmeyer *et al.*, 1995; Beste *et al.*, 2008a, 2010, 2011), verbal fluency (Bäckman *et al.*, 1997; Ho *et al.*, 2002), and episodic memory (Montoya *et al.*, 2006a), to name just a few. However, to what extent these deficits can be explained by specific basal ganglia dysfunctions has not yet been systematically explored.

Via simulations in a neuro-computational model, we here investigated whether lesions of direct pathway MSNs, indirect pathway MSNs or hyperdirect pathway STN neurons could reproduce patients' empirically observed deficits in stimulus–response learning paradigms. Moreover, we decreased and increased dopamine levels to simulate the effects of changes in dopamine receptor binding.

Materials and methods

Neuro-computational model

For our simulations, we used a neuro-computational model that has been described in detail by Schroll *et al.* (2014), who applied it to explain stimulus–response learning in healthy subjects and Parkinson's disease patients. Its architecture comprises a cortico-basal ganglia-thalamic loop that connects cortical areas related to stimulus processing to motor cortical areas (Fig. 1A). Although we did not specify the anatomical substrate of this stimulus cortex in Schroll *et al.* (2014), it most likely corresponds to prefrontal cortical areas that participate in the conscious processing of stimulus-related information. Anatomically, the prefrontal cortex has been shown to provide direct inputs to both the striatum (Alexander *et al.*, 1989) and the STN (Hartmann-von Monakow *et al.*, 1978; Canteras *et al.*, 1990). Within the basal ganglia, the model contains direct, indirect and hyperdirect pathways, as based on anatomical evidence (Braak & Del Tredici, 2008). These three pathways converge in the GPi, a basal ganglia output nucleus with tonically active neurons that inhibit the thalamus. Reductions in GPi activity increase thalamic firing and thereby result in increased motor cortical activity, whereas increases in GPi activity decrease thalamic and cortical firing. The direct pathway proceeds from the cortex via striatal MSNs to the GPi; activity within this pathway decreases GPi firing, resulting in increased cortical activity (Fig. 1B). The indirect pathway in our model proceeds from the cortex, via the striatum and the GPe, to the GPi; activity within this pathway excites GPi neurons. Note that our model contains the short route of the indirect pathway only – an additional longer route that traverses the STN (Smith *et al.*, 1998; cf. Schroll & Hamker, 2013) was omitted from the implementation as detailed in Schroll *et al.* (2014). Finally, the hyperdirect pathway proceeds from the cortex via the STN to the GPi; activity in this pathway also increases GPi firing (Fig. 1B).

Each modeled nucleus within the cortico-basal ganglia-thalamic loop contains a predefined number of artificial neurons (Table 1) whose membrane potentials and firing rates are determined via differential equations (Appendix A). In brief, each neuron's membrane potential is determined by the sum of its synaptic inputs (Eqn A1 in Appendix A). Firing rates are mostly equal to membrane potentials, but rectify negative values to zero (Eqn A2 in Appendix A). The strengths (i.e. weights) of synaptic connections between neurons are determined by an additional set of differential equations (Appendix A): according to Hebb's principle, synapses are strengthened whenever presynaptic and postsynaptic neurons are concurrently active. Synapses within the basal ganglia are additionally modulated by dopamine, as outlined below. To simulate stimulus–response learning performance in the model, stimulus representations are fed into the model's stimulus cortex, and its motor responses are read out of the motor cortex. Exploration of response alternatives is guaranteed by random noise terms that contribute to neurons' membrane potentials (Eqn A1 in Appendix A). This noise also causes variability in behavioral performance among otherwise equivalent networks.

Basal ganglia pathways learn to interconnect stimulus and motor cortices in a meaningful way on the basis of reward-related dopamine signals. Dopamine signals, provided by the substantia nigra compacta (SNc), encode reward prediction errors (i.e. difference signals between received and expected rewards): whenever the model experiences more rewards than predicted, dopamine activity increases above baseline by an amount proportional to the prediction error; whenever the model experiences fewer rewards than predicted, dopamine activity decreases below baseline. High levels of

TABLE 1. Numbers of neurons for each of the model's layers, and parameters used to compute the model's membrane potentials and firing rates

	No. of cells							
Cell type	Reversal learning	Sequence learning	Category learning	Model prediction	w^{ff}	w^{lat}	B	$\varepsilon_{i,t}$
Stimulus cortex	4	8	2	4	–	–	0.0	0.0
Motor cortex	2	3	2	5	$w_{i,j}^{Thal-Cx} = 1.0$	$w_{i,j}^{Cx-Cx} = -1.0$	0.0	[−1.0, 1.0]
Striatum (D1)	12	16	12	16		$w_{i,j}^{Str(D1)-Str(D1)} = -0.3$	0.4	[−0.1, 0.1]
Striatum (D2)	12	16	12	16		$w_{i,j}^{Str(D2)-Str(D2)} = -0.3$	0.4	[−0.1, 0.1]
Striatum (thalamus)	2	3	2	5	$w_{i,j}^{Thal-Str(Thal)} = 1.0$	$w_{i,j}^{Str(Thal)-Str(Thal)} = -0.3$	0.4	[−0.1, 0.1]
STN	12	16	12	16		$w_{i,j}^{STN-STN} = -0.3$	0.4	[−0.1, 0.1]
GPe	2	3	2	4	$w_{i,j}^{Str(Thal)-GPe} = -0.3$		1.0	[−1.0, 1.0]
GPi	2	3	2	5	$w_{i,j}^{GPe-GPi} = -1.5$ $w_{i,j}^{Str(Thal)-GPi} = -0.3$		2.4	[−1.0, 1.0]
Thalamus	2	3	2	5	$w_{i,j}^{GPi-Thal} = -1.5$	$w_{i,j}^{Thal-Thal} = -0.6$	1.0	[−0.1, 0.1]
SNc	1	1	1	1			0.1	0.0

For each modeled brain area and nucleus, numbers of simulated neurons for the four different neuro-psychological paradigms are shown. Also shown are hard-coded weights (both feedforward, w^{ff}, and lateral, w^{lat}), baseline membrane potentials (B), and the boundaries of uniform distributions from which error terms ($\varepsilon_{i,t}$) were randomly drawn. Striatum (thalamus): striatal neurons that receive thalamic feedback.

dopamine facilitate long-term potentiation (LTP) in synapses of the direct and hyperdirect pathways and long-term depression (LTD) in synapses of the indirect pathway (Fig. 1B); low levels of dopamine, in contrast, have exactly opposite effects on all pathways. These assumptions are based on findings by Shen *et al.* (2008), who reported that D1 dopamine agonists facilitated LTP in striatal D1 MSNs, whereas D1 antagonists facilitated LTD in these MSNs. Moreover, Shen *et al.* (2008) showed that D2 agonists favored LTD in striatal D2 MSNs, whereas D2 antagonists favored LTP in these MSNs. As a result of these different rules of synaptic plasticity and the pathways' excitatory vs. inhibitory effects on cortical activity, pathways self-organise in our model to fulfill the following distinct functions (Fig. 1C): the direct pathway learns to facilitate those stimulus–response associations that result in rewards, the hyperdirect pathway learns to inhibit competing stimulus–response associations whose non-execution results in rewards, and the indirect pathway learns to inhibit those stimulus–response associations that result in omissions of expected rewards (i.e. those that have been rewarded previously but that currently do not result in reward; Fig. 1C) (Schroll *et al.*, 2014). In addition, a cortico-thalamic pathway is capable of automatising previously learned stimulus–response associations, achieving faster response times as performance becomes habitual (Schroll *et al.*, 2014).

We adapted our neuro-computational model for application to the stimulus–response learning paradigms outlined below in a few minor respects. For each simulated paradigm, we set the number of neurons in the visual cortex to the number of possible stimuli and the number of neurons in the GPi, GPe, thalamus and motor cortex to the number of possible responses. Moreover, we adapted the post-synaptic activity threshold for synaptic plasticity within the thalamus to 0.3 for two thalamic neurons, to 0.47 for three thalamic neurons, to 0.55 for four thalamic neurons, and to 0.60 for five thalamic neurons. This became necessary because we adapted the number of thalamic neurons to the number of possible responses: synaptic plasticity in the thalamus is partly based on the average firing rate of this nucleus (Schroll *et al.*, 2014). This average firing rate, however, changes with the number of available but inactive neurons: for one active neuron with a firing rate of 1 and otherwise inactive neurons with firing rates of 0, the average firing rate becomes 0.5 for two thalamic neurons, 0.33 for three thalamic neurons, 0.25 for four thalamic neurons, and 0.20 for five thalamic neurons. We compensated for this effect by adapting the above-mentioned activity threshold accordingly. Overall, this activity threshold was chosen to be lower than in Schroll *et al.* (2014) (where we did not fit empirical data) to allow for faster cortico-thalamic automatisation of stimulus–response associations. For paradigms where only two response options were available, moreover, we fixed the lateral weights within the GPi to 1.25 rather than having them evaluated on the basis of synaptic plasticity, as the corresponding rule of synaptic plasticity requires at least three neurons to function adequately. Finally, we adapted the number of neurons in the two segments of the striatum (belonging to the direct and indirect pathways) and in the STN to the complexities of the simulated behavioral paradigms. For adequate learning performance, each stimulus–response association of a paradigm has to be represented by a separate cluster of neurons in each of these nuclei in our model. Neuronal numbers were set to 12 for paradigms that involved relatively few stimulus–response associations (i.e. the reversal learning paradigm and the category learning paradigm as outlined below) and to 16 for more complex paradigms (i.e. the sequence learning paradigm and the paradigm used to investigate model predictions, also as outlined below). Table 1 gives an overview of implemented numbers of neurons for each of the four different learning paradigms.

Simulating basal ganglia dysfunctions

In our neuro-computational model, we separately introduced four different lesions of the basal ganglia, and investigated whether any of these reproduced HD patients' behavioral deficits on the neuro-psychological paradigms outlined below. Following empirical

evidence that neuronal degeneration primarily affects MSNs of the striatum, we lesioned MSNs of the direct pathway, MSNs of the indirect pathway, and both of these MSNs simultaneously. To explore the effects of potential neurodegeneration within the hyperdirect pathway, we lesioned the STN. Lesions were implemented by setting the outputs of 50% of relevant neurons to zero. This percentage was chosen because piloting runs had shown it to result in obvious behavioral deficits without completely disrupting performance. Of course, effect sizes for a given relative amount of neuronal loss depend on the number of initially available neurons. This number is much reduced in our model as compared with the human brain (in line with the fact that our model has to cope with only a very limited range of behavioral tasks). Therefore, we would not claim that, with 50% neuronal loss in target structures, human patients would show deficits that are quantitatively equivalent to those revealed by our networks. We do claim, however, that qualitatively all effects should be the same.

To investigate the consequences of changes in dopamine levels, resulting in increased or reduced dopamine receptor binding, we either increased or decreased dopamine levels (both tonic and phasic) by 10%. Again, we wanted to investigate whether these changes reproduced patients' behavioral deficits qualitatively, not quantitatively.

To simulate which of these basal ganglia dysfunctions best reproduced patients' stimulus–response learning deficits, we employed an abductive or 'reverse-engineering' approach. As defined by Maia & Frank (2011) in their seminal article on computational neurology, such an approach is based on reasoning from a computational reproduction of behavioral effects to the neuronal origins of these effects.

Basal ganglia pathway effectiveness

To provide insights into the model's functioning, we recorded the outputs of basal ganglia and cortico-thalamic pathways while networks learned our model prediction task outlined below. For each trial, we computed a measure of functional pathway effectiveness as follows. (i) We first computed each pathway's outputs on the different response neurons of our model. This involved three steps. In the first step, we recorded all synaptic weights and presynaptic activities of striato-GPi synapses (direct pathway), striato-GPe synapses (indirect pathway), subthalamo-GPi synapses (hyperdirect pathway) and cortico-thalamic synapses (cortico-thalamic pathway) precisely at the time points where responses were chosen. In the second step, for each synaptic contact, we then multiplied its synaptic weight by the activity of the corresponding presynaptic neuron. In the third step, we summed up these products across all synapses that targeted a particular postsynaptic response neuron, separately for each pathway and postsynaptic neuron. (ii) For each pathway, we then determined task-relevant and task-irrelevant responses as explained below. (iii) Finally, we computed the relative number of networks whose outputs on neurons that corresponded to relevant responses superseded their average outputs on all other neurons by at least one order of magnitude. This percentage of networks in which pathway functions had developed appropriately gave us an estimation of the pathways' functional effectiveness across networks.

For the direct pathway, relevant neurons were those that encoded currently correct responses. This is because we had shown previously that this pathway's function is to specifically facilitate correct responses in our model (Fig. 1C) (Schroll *et al.*, 2014). For the indirect pathway, relevant neurons were those that targeted previously correct responses, based on previous evidence that this pathway specifically inhibits such responses in our model (Fig. 1C) (Schroll *et al.*, 2014). For the hyperdirect pathway, those postsynaptic neurons were relevant that encoded currently incorrect responses, as we had shown that the hyperdirect pathway's function is to suppress incorrect responses in a surround-inhibition manner in our model (Fig. 1C) (Schroll *et al.*, 2014). Finally, for the cortico-thalamic pathway, neurons that encoded currently correct responses and those that encoded previously correct responses were relevant – in line with previous findings that the cortico-thalamic pathway facilitates correct responses in our model and maintains this facilitation even if these responses become incorrect later.

Stimulus–response learning paradigms

For both practical and theoretical reasons, we focused on stimulus–response learning paradigms for our comparisons of empirical data with model performance. From a practical point of view, our computational model was specifically designed to be capable of reward-based stimulus–response learning (Schroll *et al.*, 2014). It therefore appeared most natural to use it in this respect also for our simulations of potential neuronal dysfunctions in HD, rather than adapting it for coping with different behavioral paradigms. From a theoretical point of view, stimulus–response learning paradigms are well suited for simulating both the effects of lesions of basal ganglia pathways and the effects of changes in dopamine levels. Whereas the behavioral consequences of basal ganglia lesions could have been simulated on any kind of neuropsychological paradigm, changes in dopamine levels are more challenging: as the only function that dopamine is generally acknowledged to contribute to in the brain is reward-based learning (e.g. Waelti *et al.*, 2001; Frank *et al.*, 2007; Glimcher, 2011), we required such a paradigm for our simulations.

We found three empirical reports on HD patients' reinforcement-based stimulus–response learning performance, namely a reinforcement-based reversal learning task (Lawrence *et al.*, 1999), a reinforcement-based sequence learning task (Brown *et al.*, 2001), and a reinforcement-based category learning task (Filoteo *et al.*, 2001).

In all of our implementations, irrespective of timings in the original paradigms, stimuli were presented for 1000 ms before responses were recorded (single time point) and feedback was presented for 500 ms, followed by inter-trial intervals of 100 ms. For each modeled basal ganglia dysfunction, we simulated 100 networks that passed our learning criteria, irrespective of the original number of subjects, thereby achieving more reliable performance estimates.

Reinforcement-based reversal learning

Lawrence *et al.* (1999) devised a reinforcement-based stimulus–response learning task that involved phases of initial learning and reversal learning (Fig. 2A). In this task, two stimuli were simultaneously presented on a computer screen. In each trial, subjects had to choose one of these stimuli for which the 'correct' stimulus resulted in positive feedback in 80% of trials and in negative feedback in 20% of trials, and the incorrect stimulus resulted in negative feedback throughout. Stimuli were presented in two of four possible locations on the screen (left, right, top, and bottom). When subjects had completed 40 trials of this task, reinforcement contingencies were reversed, and the previously unrewarded stimulus now resulted in probabilistic rewards for 40 trials. For both initial and reversal learning phases, Lawrence *et al.* (1999) analysed the number of trials required by subjects to reach a criterion of eight correct responses in a row.

Lawrence *et al.* (1999) tested 21 symptomatic patients (mean age, 48.2 years; mean disease duration, 4.8 years) and 21 age-matched and

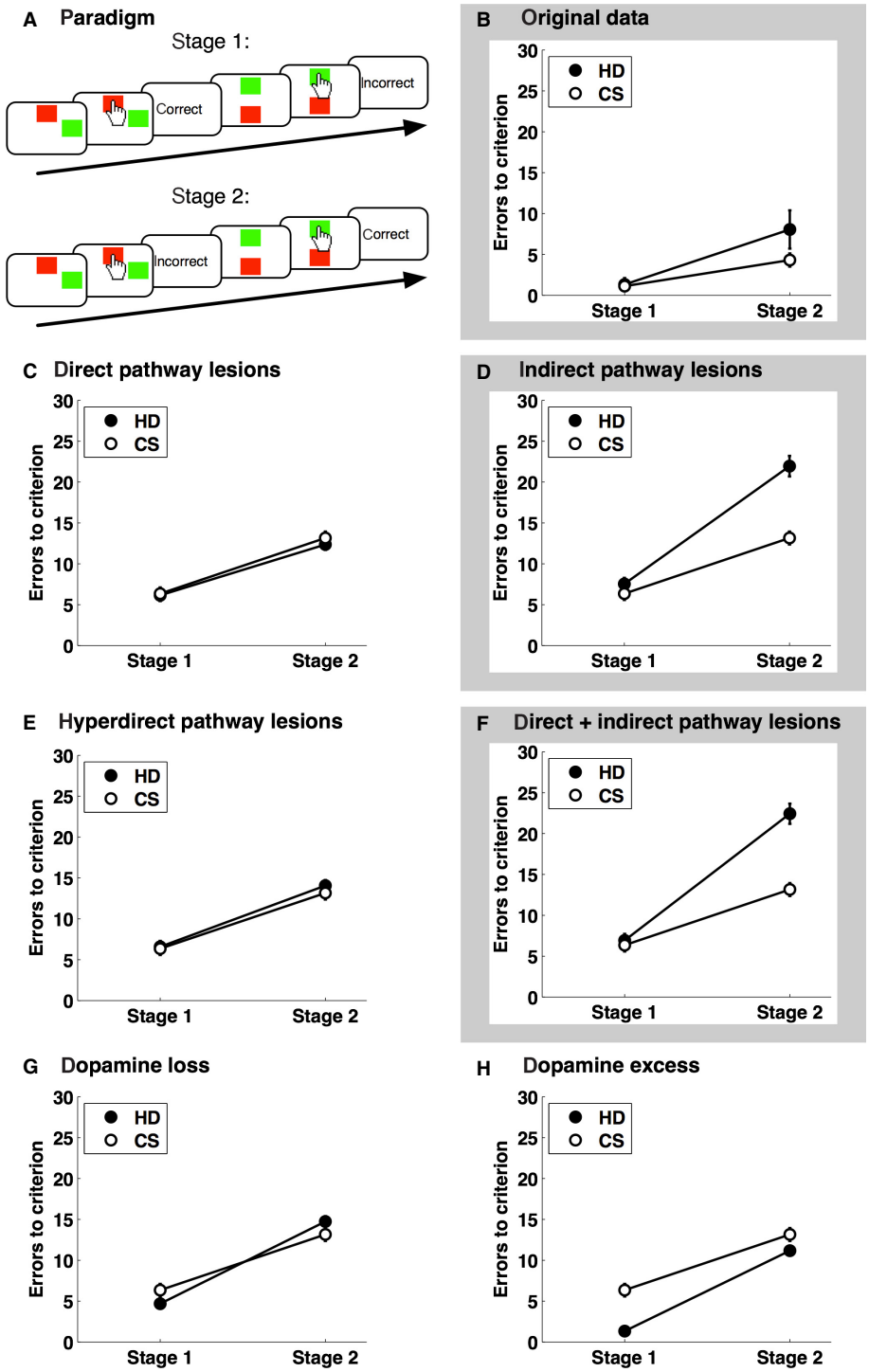


FIG. 2. Stimulus–response learning paradigm and behavioral results as reported by Lawrence *et al.* (1999). (A) Stimulus–response learning paradigm. Subjects had to learn touching the red stimulus in stage 1 of the experiment, but to switch to touching the green stimulus in stage 2, based on probabilistic feedback. Details are given in the main text. (B) Original empirical results as reported by Lawrence *et al.* (1999). The number of incorrectly responded trials before learning criterion was reached are shown for both stage 1 (initial learning) and stage 2 (reversal learning) of this paradigm, separately for HD patients (filled circles) and matched controls (open circles). The criterion was defined as eight correct responses in a row. (C–H) Simulated results for lesions of the direct, indirect, hyperdirect and direct plus indirect pathways, and for reduced and increased dopamine levels. Error bars represent standard errors of the mean. Gray background shading indicates panels with results that are qualitatively equivalent to those shown in B. CS, control subjects.

IQ-matched controls on this paradigm. Patients' mean score on the Unified Huntington's Disease Rating Scale was 21.0; no patient fulfilled the criteria for dementia on the Mini Mental State Examination.

For our implementation, we simplified this paradigm, in that we did not model the presentation of stimuli in different spatial locations, thereby holding stimulus positions constant across trials. We

opted for this approach because there is no way of presenting red and green stimuli in different spatial locations to our model such that it would recognise the equivalence of red and green across these locations. In the human brain, this ability most likely does not rely on the basal ganglia, so we think that our simplification is justified for our model.

Reinforcement-based sequence learning

In a paradigm used by Brown *et al.* (2001), subjects had to learn a simple sequence of eight letters (URLDUDRL). In the first trial of each run, the first letter of this sequence (i.e. U) appeared highlighted on a computer screen (Fig. 3A). Subjects could respond by pressing any of the response buttons corresponding to the three other letters. If they chose the correct button, this new stimulus became highlighted on the screen. If not, the previous stimulus remained highlighted, and subjects had to choose again. At the end of each eight-letter sequence, the whole sequence started again. Subjects continued with this sequence for 10 repetitions in each of 10 blocks. Brown *et al.* (2001) measured the proportion of trials in which subjects responded correctly upon first choice, separately for each block. Moreover, they determined the length of the longest sequence of first-choice correct responses in each block.

Brown *et al.* (2001) tested 16 symptomatic patients with HD on this paradigm (mean age, 42.8 years, mean disease duration, 6.4 years) who scored 41.8 points on average on the Unified Huntington's Disease Rating Scale motor section. A control group was matched with regard to sex, age, education, class of occupation, and handedness.

For our computational implementation, we adapted the paradigm in the following ways. In the original paradigm, the sequence consisted of four distinct stimuli (i.e. U, D, L, and R), each of which was presented twice in each eight-character sequence run (i.e. URLDUDRL). In our implementation, the sequence's eight characters were handled as eight distinct stimuli (that is, identical letters within the first half and the second half of the sequence were handled as different inputs). Thereby, the model could differentiate between identical stimuli at different positions of the sequence. We think that this simplification is justified for a model that focuses on the functions and dysfunctions of basal ganglia pathways, and that therefore does not contain any memory system: differentiation between identical stimuli at different sequence positions would have required a working memory system that could store stimuli presented immediately before. The number of response options, however, was kept at three (corresponding to the three stimuli not currently highlighted). Networks that did not finish the task after 5.000 trials were abandoned, but fully scored.

Reinforcement-based category learning

Filoteo *et al.* (2001) asked their subjects to categorise simple line stimuli on the basis of reinforcement signals. Stimuli consisted of a horizontal line and a vertical line that were interconnected at their upper and left ends, respectively, and that were independently varied in length. In two separate experiments, category boundaries were determined on the basis of either linear or non-linear rules. Our neuro-computational model can solve only the linear condition. In each trial of this condition, subjects saw one of the above-described stimuli and had to indicate whether it belonged to category 1 or to category 2 (Fig. 5A). The paradigm comprised six blocks of 100 trials each. One hundred different stimuli were generated and shown in six different orderings across blocks.

Filoteo *et al.* (2001) recruited seven HD patients for this paradigm (mean age, 49.9 years) and six healthy controls who did not differ significantly from patients with regard to age or years of education. Subjects' correct-response rates were analysed blockwise.

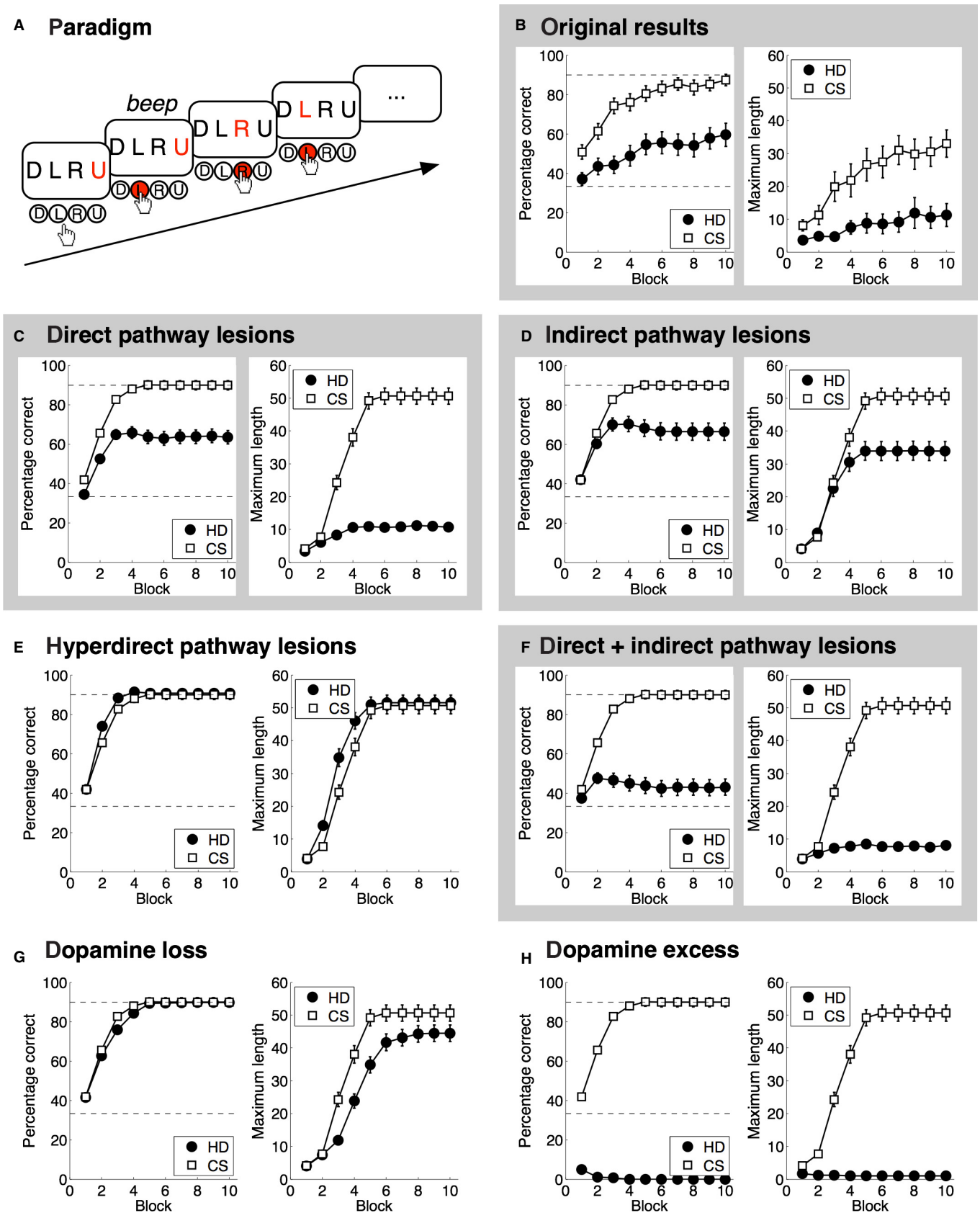
When implementing this paradigm, we assumed that stimuli were represented by two neurons in the stimulus cortex in a binary fashion: the first neuron became active whenever the length of the vertical line of the stimulus exceeded the length of the horizontal line, and the second neuron became active whenever the horizontal line was longer. In line with neurophysiological evidence (e.g. Freedman *et al.*, 2001), the definition of category boundaries was implemented to reside within the prefrontal stimulus cortex. This allowed basal ganglia pathways to learn associations between given categories and response options in a stimulus–response-like manner; it does not preclude the possibility that different basal ganglia loops (not modeled here) are involved in category learning as well (Antzoulatos & Miller, 2011). The lengths of the two lines were drawn from bivariate normal distributions as specified by Filoteo *et al.* (2001) (category 1, mean of dimension 1 = 0.38, standard deviation = 0.1, mean of dimension 2 = 0.62, standard deviation = 0.1; category 2, dimensions 1 and 2 exchanged), but separately for each simulated network, such that our simulations did not depend on specific values of these randomly drawn variables.

Paradigm to investigate model predictions

To highlight empirically testable predictions of our computational model, we simulated the model's performance on an additional stimulus–response learning paradigm that, to our knowledge, has not yet been used to empirically investigate learning deficits of HD patients. The use of this paradigm in an empirical study on HD patients and healthy controls in the future will allow disentanglement of the effects of the dysfunctions of direct and indirect pathways in HD that are described in this work, and may provide further support for our interpretations.

In a first phase of this paradigm, associations between three abstract arbitrary stimuli and four response buttons have to be learned via rewarding feedback (Fig. 6A). Whenever stimulus 1 is presented, button *A* should be pressed; whenever stimulus 2 is presented, button *B* should be pressed; and whenever stimulus 3 is

FIG. 3. Sequence learning paradigm and behavioral results as reported by Brown *et al.* (2001). (A) Sequence learning paradigm as reported by Brown *et al.* (2001). Subjects had to learn an eight-letter sequence (URLDUDRL) across 10 blocks of trials, where each block consisted of 10 completed sequences. In each trial, a letter was presented on a computer screen, and subjects had to indicate the subsequent letter by pressing the corresponding button. If they decided correctly, this letter replaced the previous one on the screen; if not, the previous letter remained, and subjects were required to choose again. Thus, subjects moved through the sequence, starting again from the beginning whenever the last letter was reached. Details are given in the main text. Dashed lines within the left subplot show the chance rate of 33.33% correct responses and the criterion rate of 90% correct responses, respectively. (B) Original results as reported by Brown *et al.* (2001). The left subplot shows the development of average correct response rates across blocks (where each block consists of 10 repetitions of the eight-letter sequence), separately for HD patients (filled circles) and healthy controls (open circles). Only trials in which subjects responded correctly upon first choice were counted as correct. The right subplot shows, for each block, the length of subjects' longest sequence of first-choice correct responses, again separately for HD patients (circles) and healthy controls (squares). (C–H) Simulation results for partial lesions of the direct, indirect, hyperdirect and direct plus indirect pathways, and for reduced and increased dopamine levels. Subplots are organised as in B. Error bars represent standard errors of the mean. Gray background shading indicates panels with results that are qualitatively equivalent to those shown in B. CS, control subjects.



presented, button *C* should be pressed. In a second step of the paradigm, starting when phase 1 has been reliably learned and continuing for 500 trials, networks are presented with these same three stimuli and one additional stimulus. The previously known stimulus 1 now becomes associated with response *D* (reversal), stimulus 2 remains associated with response *B* (maintenance), and stimulus 3 no longer results in reward, no matter which response is selected (extinction). The novel stimulus 4 becomes associated with response *A*. For this second phase of the experiment, our model makes a couple of specific predictions, as outlined in Results.

We simulated our model's performance for 'healthy', i.e. unlesioned, networks, for networks with partial lesions of the indirect pathway (corresponding to early stages of HD in humans), and for networks with partial lesions of both the direct pathway and the indirect pathway (corresponding to later stages of HD in humans).

Results

Reinforcement-based reversal learning

In the study of Lawrence *et al.* (1999), 14 of 21 patients (66.6%) and 20 of 21 controls (95.2%) passed the initial learning stage (stage 1) of the stimulus–response learning paradigm. For just these patients, a two-factor repeated-measures ANOVA on numbers of trials required to reach the learning criterion revealed a significant main effect of stage (i.e. initial vs. reversal learning), no significant effect of group (i.e. HD patients vs. healthy controls), and a marginally significant stage × group interaction, which could be resolved such that HD patients required significantly more trials on the reversal stage than controls, but not on the initial learning stage (Fig. 2B).

In our simulations, 72.5% of the unlesioned control networks reached the learning criterion. Lesions of the direct pathway reproduced the finding of Lawrence *et al.* (1999) of a reduced number of patients reaching the learning criterion (51.8% successful networks), as did combined lesions of the direct and indirect pathways (57.5%) and increased levels of dopamine (56.5%), but not lesions of the hyperdirect (71.9%) or indirect pathway (76.3%) or decreased levels of dopamine (79.4%).

For networks passing stage 1, data were analysed further. Overall, networks required more trials for learning criteria to be reached than the subjects of Lawrence *et al.* (1999) s (Fig. 2C–H). Their conclusions, however, were based on differences between conditions, independently of overall performance levels. We applied the same statistical tests as used by Lawrence *et al.* (1999) to our simulation data. Lesions of the indirect pathway produced a significant main effect of stage ($F_{1,198} = 152.89$, $P < 0.0001$), a significant main effect of group ($F_{1,198} = 19.85$, $P < 0.0001$), and a significant stage × group interaction ($F_{1,198} = 8.86$, $P = 0.003$ (Fig. 2D), as did combined lesions of the direct and indirect pathways (stage, $F_{1,198} = 164.12$, $P < 0.0001$; group, $F_{1,198} = 17.49$, $P < 0.0001$; stage × group interaction, $F_{1,198} = 15.07$, $P = 0.0001$). Reduced and increased levels of dopamine produced a significant main effect of stage ($F_{1,198} > 187$, $P < 0.0001$) and a significant interaction ($F_{1,198} > 6.8$, $P < 0.01$). Lesions of the direct and hyperdirect pathways, in contrast, did not produce a significant interaction (Fig. 2C and E). Additionally taking into account visual comparisons between the original data and our simulated dysfunctions (Fig. 2), we conclude that lesions of the indirect pathway and combined lesions of the direct and indirect pathways best reproduced the results of Lawrence *et al.* (1999).

Reinforcement-based sequence learning

Brown *et al.* (2001) found that HD patients were significantly impaired in sequence learning as compared with healthy controls (Fig. 3B). In a repeated-measures ANOVA on correct-response rates, they found significant main effects of group (i.e. HD patients vs. controls) and block (i.e. learning progress), and a marginally significant group × block interaction. In an equivalent ANOVA on subjects' longest sequences of first-choice correct responses, they found significant main effects of group and block, and a significant group × block interaction (Fig. 3B). Moreover, they reported that only two of 16 patients (12.5%), but 13 of 16 controls (81.3%), reached the learning criterion of 90% correct responses in any of the 10 blocks (Fig. 4A).

We applied the same statistical tests as used by Brown *et al.* (2001) to analyse our simulation data. We found that lesions of direct-pathway MSNs reproduced patients' deficits well (Fig. 3C). In a repeated-measures ANOVA on correct-response rates, there were significant main effects of group ($F_{1,1782} = 45.30$, $P < 0.0001$) and block ($F_{9,1782} = 212.33$, $P < 0.0001$), and a significant group × block interaction ($F_{9,1782} = 15.39$, $P < 0.0001$). Similarly, combined lesions of the direct and indirect pathways (Fig. 3F) reproduced the main effects of group ($F_{1,1782} = 100.23$, $P < 0.0001$) and block ($F_{9,1782} = 83.42$, $P < 0.0001$), and the interaction ($F_{9,1782} = 63.05$, $P < 0.0001$). Lesions of the indirect pathway (Fig. 3D) also reproduced these effects (group, $F_{1,1782} = 20.96$, $P < 0.0001$; block, $F_{9,1782} = 124.13$, $P < 0.0001$; group × block, $F_{9,1782} = 16.5$, $P < 0.0001$). Lesions of the hyperdirect pathway (Fig. 3E), in contrast, reproduced the significant main effect of block ($F_{9,1782} = 447.52$, $P < 0.0001$) and the significant group × block interaction ($F_{9,1782} = 3.43$, $P = 0.0003$), but not the significant main effect of group ($F_{1,1782} = 0.93$, $P = 0.34$). Decreased dopamine levels (Fig. 4G) reproduced the significant main effect of block ($F_{9,1782} = 395.62$, $P < 0.0001$), but neither the main effect of group ($F_{1,1782} = 0.54$, $P < 0.46$) nor the interaction ($F_{9,1782} = 1.88$, $P = 0.05$). Finally, with increased levels of dopamine (Fig. 3H), networks were unable to complete the task, sticking to incorrect sequence elements infinitely. We therefore refrained from performing statistical tests for these networks.

Regarding the networks' longest sequences of first-choice correct responses, again, lesions of the direct pathway reproduced the significant main effects of group ($F_{1,1782} = 240.24$, $P < 0.0001$) and block ($F_{9,1782} = 200.93$, $P < 0.0001$), and the group × block interaction ($F_{9,1782} = 116.59$, $P < 0.0001$), as did combined lesions of the direct and indirect pathways (group, $F_{1,1782} = 264.37$, $P < 0.0001$; block, $F_{9,1782} = 177.99$, $P < 0.0001$; group × block, $F_{9,1782} = 137.09$, $P < 0.0001$), lesions of the indirect pathway (group, $F_{1,1782} = 14.85$, $P = 0.0002$; block, $F_{9,1782} = 225.95$, $P < 0.0001$; group × block, $F_{9,1782} = 15.21$, $P < 0.0001$), and decreased dopamine levels (group, $F_{1,1782} = 10.74$, $P = 0.001$; block, $F_{9,1782} = 286.95$, $P < 0.0001$; group × block, $F_{9,1782} = 5.96$, $P < 0.0001$). Lesions of the hyperdirect pathway, in contrast, reproduced the main effect of block ($F_{9,1782} = 338.61$, $P < 0.0001$) and the group × block interaction ($F_{9,1782} = 3.67$, $P = 0.0001$), but not the significant main effect of group ($F_{1,1782} = 1.52$, $P < 0.22$). Increased levels of dopamine were not analysed further.

For the number of networks reaching the 90% performance criterion (Fig. 4), we found that 94.0% of control networks were successful. HD patients' impairments in reaching this criterion were reproduced by lesions of the direct pathway (15.0% of networks successful), by lesions of the indirect pathway (69.0% successful), and by combined lesions of the direct and indirect pathways (12.0%

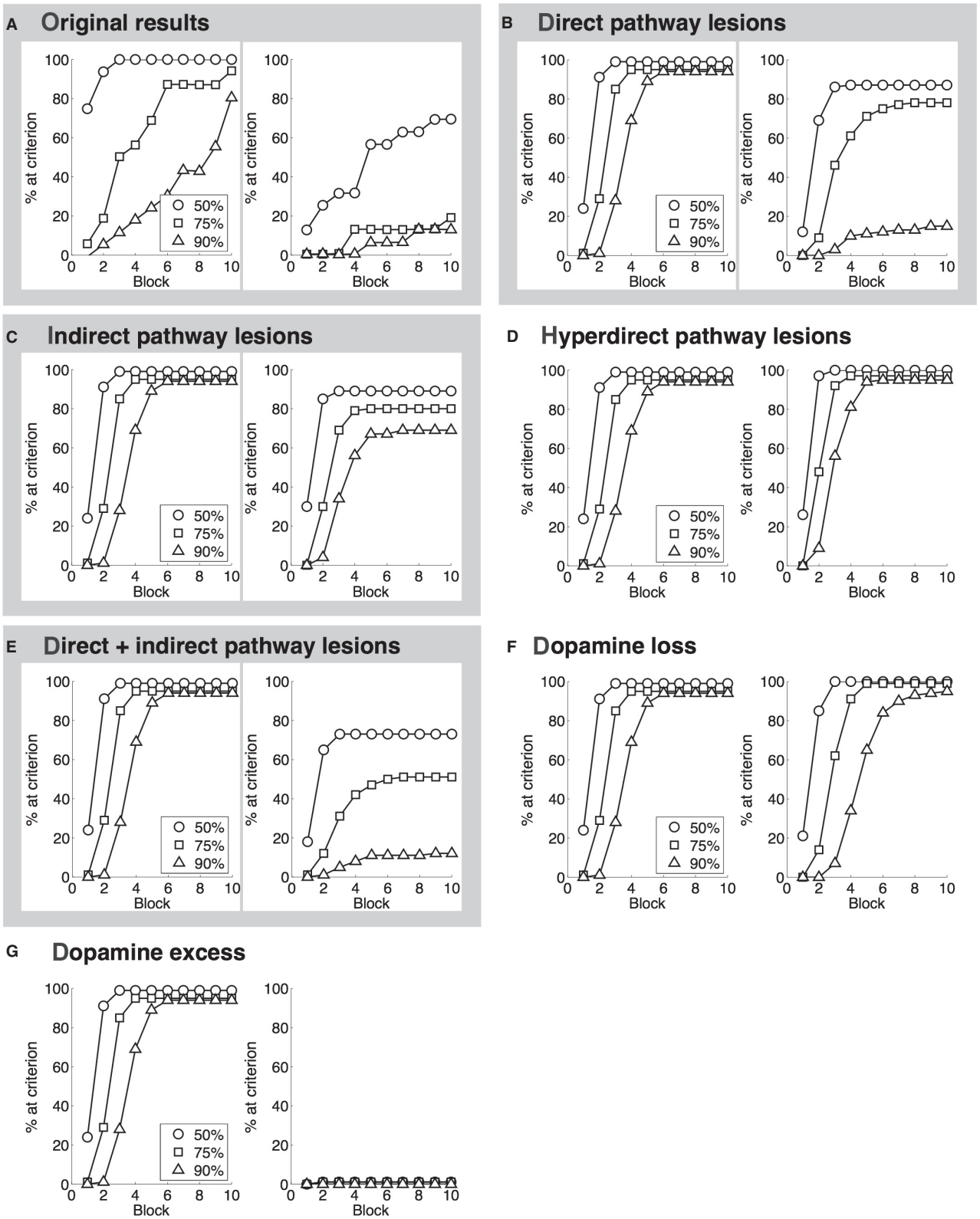


FIG. 4. Additional results of the sequence learning paradigm presented by Brown *et al.* (2001). (A) Original results as reported by Brown *et al.* (2001). For both healthy controls (left) and HD patients (right), subplots show the percentage of subjects across blocks who reached the learning criteria of 50% (circles), 75% (squares) and 90% (triangles) correct responses. (B–G) Simulation results for partial lesions of the direct, indirect, hyperdirect and direct plus indirect pathways, and for reduced and increased dopamine levels. Subplots are organised as in A. Gray background shading indicates panels with results that are qualitatively equivalent to those shown in A.

successful), but not by lesions of the hyperdirect pathway (100.0% successful) or by decreased dopamine levels (95.0% successful). With increased dopamine levels, as outlined above, no network finished the task, excluding this condition from all statistical analyses. Overall, therefore, lesions of the direct pathway, lesions of the indirect pathway and combined lesions of the direct and indirect pathways reproduced the results best. Visual comparisons between empirical and simulated results particularly highlighted combined lesions of the direct and indirect pathways as reproducing results well.

Reinforcement-based category learning

For their category learning paradigm, Filoteo *et al.* (2001) performed repeated-measures ANOVAS on accuracy rates, with the between-subjects factor group (i.e. HD patients vs. controls) and the within-subjects factor block (i.e. learning progress), for both their full six blocks of 100 trials each and for the first block only, with subdivision into 10 blocks of 10 trials each (Fig. 5B). For these first 100 trials, they reported a significant main effect of group, reflecting HD patients' overall impaired performance as compared with healthy controls, and a significant main effect of block, reflecting subjects' overall improvements across blocks, but no significant group × block interaction. For the full 600 trials, they again reported significant main effects of group and block, and also a significant interaction, reflecting stronger improvement of HD patients than of controls across blocks.

When we performed repeated-measures ANOVAS on our simulated data as used by Filoteo *et al.* (2001), we found that lesions of direct-pathway MSNs (Fig. 5C) reproduced the significant main effect of group for both the first 100 trials ($F_{1,1782} = 11.42$, $P = 0.0009$) and the full 600 trials ($F_{1,990} = 32.34$, $P < 0.0001$), as did combined lesions of the direct and indirect pathways (Fig. 5F) ($F_{1,1782} = 5.71$, $P = 0.018$ and $F_{1,990} = 35.32$, $P < 0.0001$, respectively). Decreased levels of dopamine reproduced a significant main effect of group for the full 600 trials (Fig. 5G) ($F_{1,990} = 13.52$, $P = 0.0003$), but not for the first 100 trials ($F_{1,1782} = 0.21$, $P = 0.65$), whereas increased levels of dopamine reproduced the main effect for the first 100 trials (Fig. 5H) ($F_{1,1782} = 9.19$, $P = 0.003$), but not for the full 600 trials ($F_{1,990} = 0.73$, $P = 0.40$). Lesions of the indirect pathway alone (Fig. 5D) and lesions of the hyperdirect pathway (Fig. 5E) did not reproduce any significant main effect of group (all $F < 2.96$, $P > 0.08$). All simulated dysfunctions reproduced the significant main effects of block for both the first 100 trials and the full 600 trials ($F > 66.26$, $P < 0.0001$). A significant interaction for the full 600 trials was reproduced by lesions of the indirect pathway ($F_{1,990} = 2.82$, $P = 0.02$) and increased levels of dopamine ($F_{1,990} = 9.16$, $P < 0.0001$). These simulated dysfunctions, however, did not reproduce the more important main effect of group as outlined above. Overall, we therefore conclude that lesions of the direct pathway and combined lesions of direct and indirect pathways reproduced results best.

Model predictions

In the preceding subsections, we have shown that only combined lesions of the direct and indirect pathways can reproduce HD patients' deficits on all simulated stimulus–response learning paradigms. In the following, we report the model's performance on a novel stimulus–response learning paradigm for these types of lesion to generate empirically testable predictions (Fig. 6A). We simulated three model conditions: healthy (unlesioned networks), early HD (networks with 50% lesions of the indirect pathway), and advanced HD (networks with 50% lesions of both the direct pathway and the indirect pathway). Our predictions relate to learning impairments of HD patients during the reversal stage of the paradigm (i.e. stage 2; Fig. 6B) and to alterations in the functional effectiveness of basal ganglia and cortico-thalamic pathways during the reversal stage (Fig. 6C).

For a better understanding of HD patients' predicted deficits in the effectiveness of basal ganglia pathways, we will first repeat these pathways' functions in 'healthy', i.e. unlesioned, networks. As outlined in Materials and methods and shown in Fig. 1C, the direct pathway's function is to learn facilitation of currently correct responses. Therefore, it increases its effectiveness when novel stimulus–response associations are learned in unlesioned networks, i.e. during reversal learning and learning of novel stimulus–response associations (circles in Fig. 6C, first line of subplots, stimuli 1 and 4). As outlined in Materials and Methods, the indirect pathway's function is to suppress previously correct responses (Fig. 1C). Therefore, its effectiveness increases in our paradigm whenever previously correct stimulus–response associations have to be suppressed, as is the case in reversal learning and extinction trials, in particular during early blocks (circles in Fig. 6C, second line of subplots, stimuli 1 and 3). Importantly, suppression of previously correct responses is a prerequisite for learning novel correct responses in reversal learning trials (stimulus 1). The hyperdirect pathway's function is to suppress currently incorrect responses, thereby supporting the selection of correct responses via the direct pathway (see Materials and methods and Fig. 1C). Therefore, it becomes effective in our paradigm whenever novel stimulus–response associations are learned (circles in Fig. 6C, third line of subplots, stimuli 1 and 3). This suppression becomes unnecessary once the direct pathway has learned to reliably facilitate correct responses, accounting for the reduction in effectiveness of the hyperdirect pathway as performance improves. Finally, the cortico-thalamic pathway's function is to 'automatize' currently correct stimulus–response associations and to implement a long-term memory system of previously correct responses (see Materials and methods) (Schroll *et al.*, 2014). In line with this function, it increases its effectiveness in facilitating currently correct responses in our paradigm whenever novel stimulus–response associations are learned (Fig. 6C, last line of subplots, stimuli 1 and 3, lower subpanels). Moreover, it remains effective in facilitating previously correct (but now incorrect) responses, thereby implementing long-term memory (circles in Fig. 6C, last line of subplots, stimulus 1, upper subpanel).

HD patients in relatively early stages of the disease, where lesions of indirect-pathway striatal MSNs are predominant (Reiner *et al.*, 1988; Albin *et al.*, 1992), are predicted to show deficits in reversal learning, i.e. in associating previously known stimuli with novel responses (triangles in Fig. 6B, stimulus 1) and in extinction learning, i.e. in suppressing previously established stimulus–response associations altogether (triangles in Fig. 6B, stimulus 3). As shown in Fig. 6C, both of these deficits are predicted to result from reduced functional effectiveness of the indirect pathway (triangles in Fig. 6C, second line of subplots, stimuli 1 and 3), i.e. from this pathway's failure to inhibit previously correct responses. Because this inhibition is a prerequisite for learning novel stimulus–response associations, the direct pathway is predicted to show a deficit in facilitating novel correct responses, i.e. to be less functionally effective as well (triangles in Fig. 6C, first line of subplots, stimulus 1), although this pathway is not lesioned itself. Patients in early stages of the disease are predicted to not be impaired in learning correct responses to novel stimuli (triangles in Fig. 6B, stimulus 4) or in

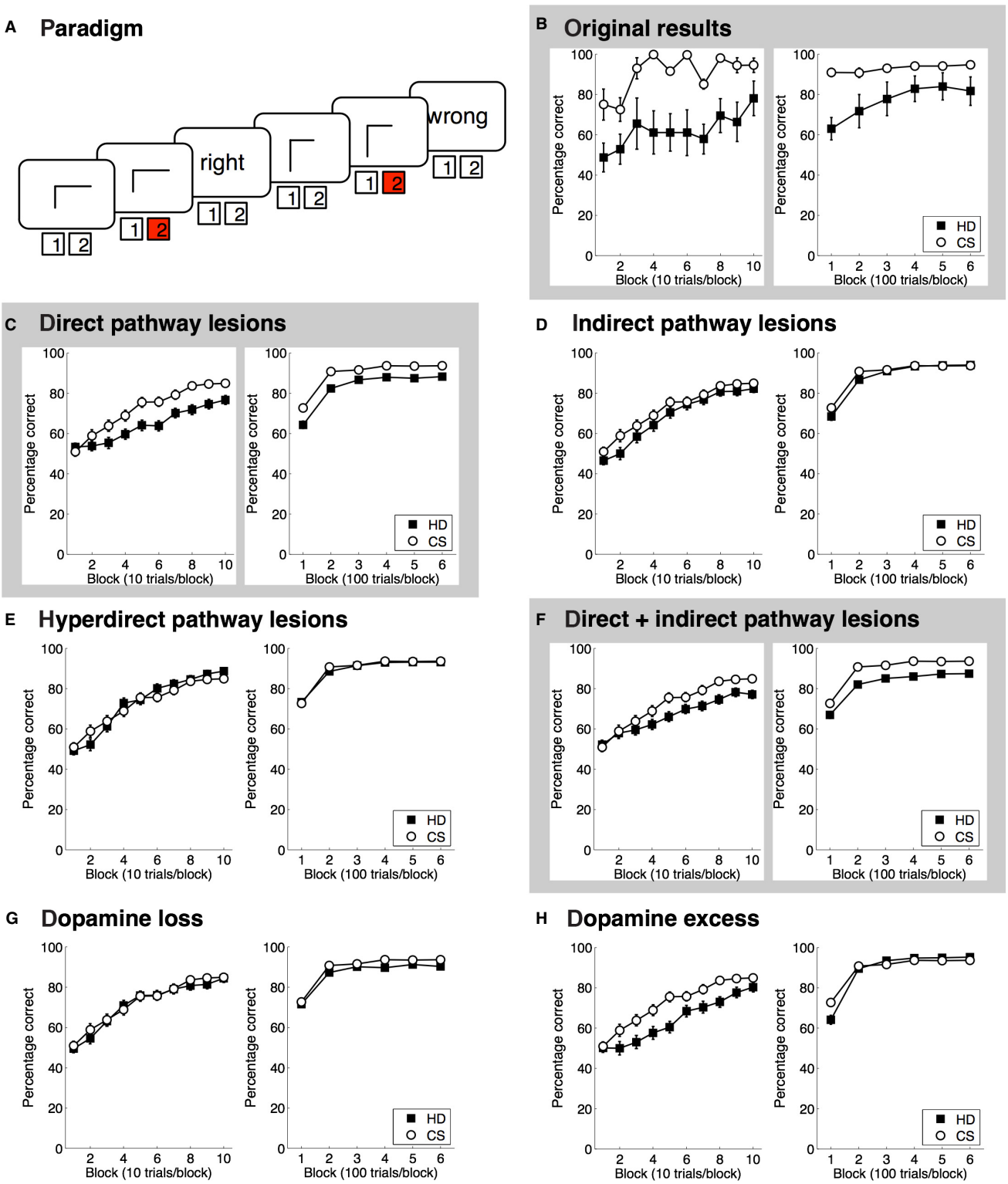


FIG. 5. Category learning paradigm and behavioral results as reported by Filoteo *et al.* (2001). (A) Category learning paradigm. Subjects were presented with six blocks of 100 trials each. In each trial, a line stimulus had to be categorised as belonging to either category 1 or category 2, where the relative lengths of a horizontal line and a vertical line constituted the decisive factor. Details are given in the main text. (B) Original empirical results as reported by Filoteo *et al.* (2001). Subplots show HD patients' (squares) and healthy controls' (circles) percentages of correct responses across six blocks of 100 trials each (right) and across the first 10 subblocks of 10 trials each (left). (C–H) Simulated results for lesions of the direct, indirect, hyperdirect and direct plus indirect pathways, and for reduced and increased dopamine levels. Error bars represent standard errors of the mean. Subplots are organised as in B. Gray background shading indicates panels with results that are qualitatively equivalent to those shown in B. CS, control subjects.

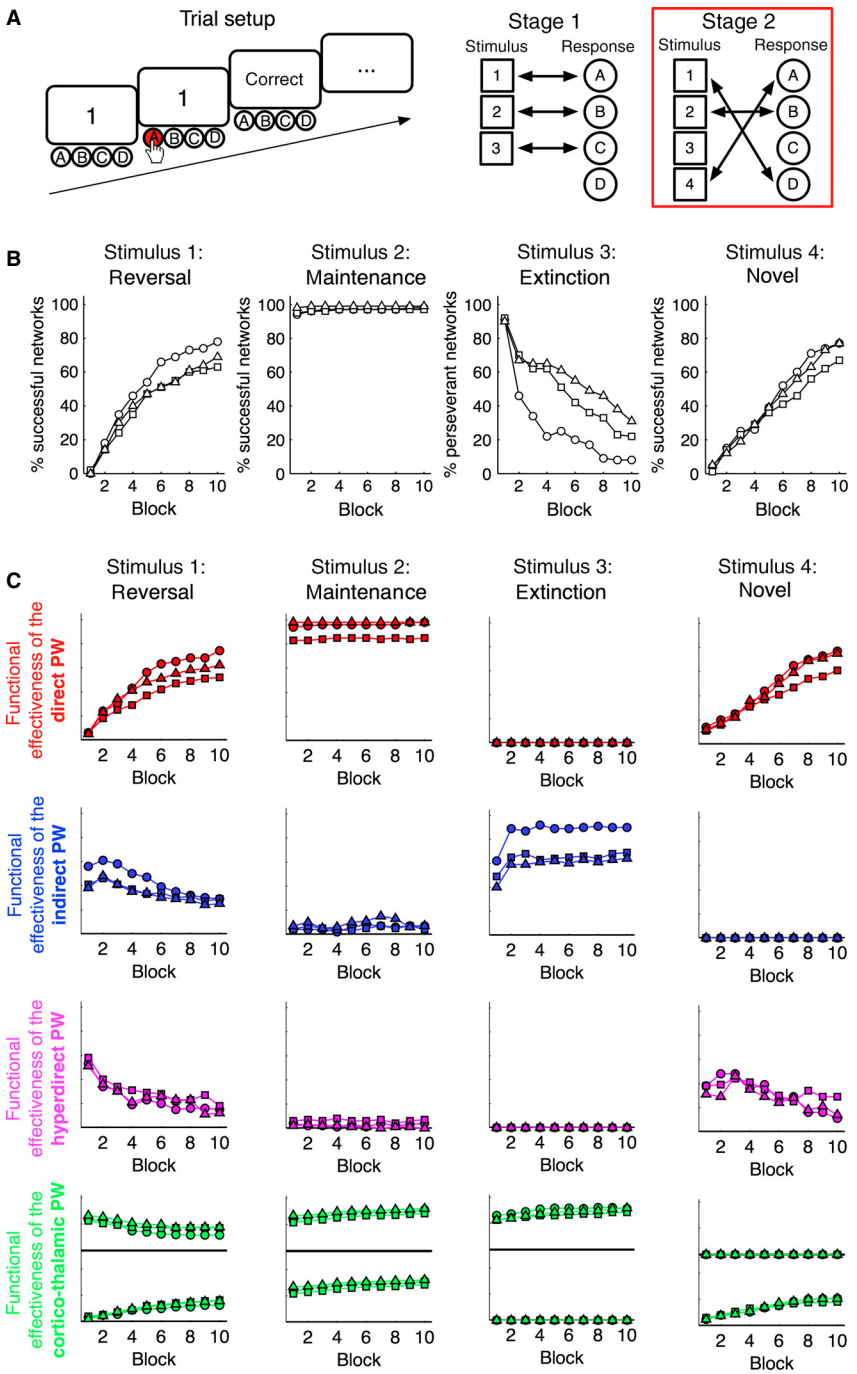


FIG. 6. Stimulus–response learning paradigm used to test our model predictions. (A) Trial setup and stimulus–response associations. In each trial, networks were presented with one of four stimuli (1–4), the networks' response (*A–D*) was recorded, and they received feedback on whether their response was correct. During stage 1 of the paradigm, the correct response for stimulus 1 was *A*, that for stimulus 2 was *B*, and that for stimulus 3 was *C*. In the second stage of the paradigm, stimulus 1 became associated with response *D* (reversal), whereas stimulus 2 remained associated with response *B* (maintenance), and stimulus 3 no longer resulted in reward, independently of response choice (extinction). A novel stimulus 4 became associated with response *A*. (B) Simulated model performance on stage 2 of the paradigm shown in A for healthy control networks (circles), networks with lesions of the indirect pathway (triangles), and networks with combined lesions of the direct and indirect pathways. For stimuli 1, 2, and 4, subplots show the percentage of successful networks, i.e. the percentage of networks satisfying a learning criterion of at least 90% correct responses in a given block of learning. For stimulus 3, the subplot shows the percentage of perseverant networks, i.e. the percentage of networks satisfying a criterion of at least 50% perseverant responses in a given block of learning. Each block of learning consisted of 50 trials. (C) Pathway dysfunctions explain the learning deficits of networks with HD-like lesions. Subplots show developments in the functional effectiveness of basal ganglia and cortico-thalamic pathways across 10 blocks of learning – separately for unlesioned networks (circles), networks with lesions of the indirect pathway (triangles), and networks with combined lesions of the direct and indirect pathways (squares). We determined the pathways' effectiveness by measuring their outputs on neurons encoding relevant responses, and comparing them with their average outputs on all other neurons, as detailed in Materials and methods. For the direct pathway, correct responses were relevant; for the indirect pathway, previously correct (but now incorrect) responses were relevant; for the hyperdirect pathway, currently incorrect responses were relevant; and for the cortico-thalamic pathway, both currently correct responses (lower subpanel) and previously correct responses (upper subpanel) were relevant. Details on computation of the pathways' functional effectiveness are given in Materials and methods. PW, pathway.

maintaining previously learned stimulus–response associations (triangles in Fig. 6B, stimulus 2). In line with these behavioral observations, the pathways' effectiveness does not differ from that of unlesioned networks for the corresponding types of trial (triangles in Fig. 6C, stimuli 2 and 4).

HD patients in relatively advanced stages of the disease who suffer not only from lesions of the indirect pathway but also from additional lesions of the direct pathway (Reiner *et al.*, 1988) are predicted to show impairments not only in reversal learning (squares in Fig. 6B, stimulus 1) and extinction learning (squares in Fig. 6B, stimulus 3) as compared with healthy controls, but also in learning novel stimulus–response associations (squares in Fig. 6B, stimulus 4). That is, these patients' additional lesions of the direct pathway are predicted to cause deficits in learning novel stimulus–response associations. These additional deficits are predicted to result from reduced functional effectiveness of the direct pathway (squares in Fig. 6C, first line of subplots, stimulus 4), i.e. from a deficit of this pathway in facilitating currently correct responses that results in reduced execution of these correct responses.

Discussion

We showed that combined lesions of direct pathway and indirect pathway striatal MSNs, but none of these lesions alone, reproduced HD patients' deficits in stimulus–response learning on three distinct neuro-psychological paradigms. Lesions of the hyperdirect pathway and changes in dopamine levels could not account for patients' impairments. Our results therefore suggest that combined lesions of the direct and indirect pathways, rather than lesions of the hyperdirect pathway or changes in dopamine, are responsible for patients' learning deficits, and tentatively also for their motor and cognitive deficits in general, as will be discussed below.

Loss of striatal MSNs explains HD patients' stimulus–response learning deficits

Via immunohistochemical and autoradiographic techniques, striatal MSNs have been repeatedly confirmed to degenerate in human HD brains (Reiner *et al.*, 1988; Albin *et al.*, 1992; Glass *et al.*, 2000). Striato-GPe MSNs of the indirect pathway die in the early and presymptomatic stages of the disease, whereas striato-GPi/striato-SNr MSNs of the direct pathway die only as the disease advances (Reiner *et al.*, 1988; Albin *et al.*, 1992; Glass *et al.*, 2000). The early loss of striato-GPe MSNs of the (inhibitory) indirect pathway has been suggested to explain patients' hyperkinetic choreic movements in the early stages of the disease, whereas the later loss of (excitatory) direct pathway MSNs might explain their hypokinetic symptoms as the disease advances (DeLong, 1990; Storey & Beal, 1993).

In our simulations, combined lesions of the direct and indirect pathways reproduced HD patients' stimulus–response learning deficits well. Lesions of the direct and indirect pathways, however, contributed to these deficits in different ways. Lesions of the indirect pathway reproduced patients' reversal learning deficits because of less reliable inhibition of previously correct responses. These deficits in inhibiting unwanted but still dominant response tendencies are conceptually similar to patients' choreatic movements in the early stages of manifest HD. As outlined above, choreatic movements have indeed been argued to result from a loss of indirect pathway MSNs (DeLong, 1990). Lesions of the direct pathway, in contrast, reproduced patients' deficits in facilitating correct responses well, specifically their impairments in initiating correct sequence elements in the sequence learning paradigm and in choosing correct categories in the category learning paradigm. Such deficits in facilitating correct responses are conceptually similar to the hypokinetic symptoms that occur in some patients with advanced HD (Denny-Brown, 1960; Thompson *et al.*, 1988). As outlined above, hypokinetic symptoms have indeed been proposed to result from lesions of the direct basal ganglia pathway (e.g. André *et al.*, 2010, 2011). To the extent that stimulus–response learning paradigms are well suited for investigating HD dysfunctions (see Materials and methods), our results corroborate combined lesions of the direct and indirect pathways as the neuronal origin of HD motor and cognitive impairments in general. Changes in dopamine levels, in contrast, do not account for patients' deficits in our simulations. The empirically reported changes in dopamine receptor binding may therefore be corrolaries of MSN degeneration, according to our results.

Comparisons with previous computational models of HD

Beste *et al.* (2014) presented a computational model that explains the neuronal mechanisms behind improved response selection in HD, as empirically reflected in both reduced error rates and reduced reaction times (Beste *et al.*, 2008c, 2014; Tomkins *et al.*, 2013). These authors suggest that a loss of connectivity among striatal neurons impairs response selection in HD to a minor extent, while, at the same time, increased sensitivity of *N*-methyl-D-aspartate (NMDA) receptors improves response selection. In HD, these improvements are proposed to outweigh the above-mentioned impairments, resulting in overall improved response selection. In contrast, in benign hereditary chorea, which is a distinct disorder, the loss of connectivity among striatal MSNs is proposed to prevail over increased NMDA receptor sensitivity, resulting in overall impaired response selection. Thus, in contrast to our approach, Beste *et al.* (2014) did not investigate the neuronal origins of behavioral impairments in HD, but focused on a specific behavioral improvement relative to healthy controls. Therefore, their conclusions do not contrast with our results: whereas increased NMDA receptor sensitivity might explain improved response selection in HD (Beste *et al.*, 2014), combined lesions of the direct and indirect pathways might account for patients' deficits in stimulus–response learning. Speculatively, increased NMDA receptor sensitivity might be a compensatory mechanism to alleviate the effects of MSN lesions.

The model of Beste *et al.* (2014) does not include synaptic plasticity, and is therefore incapable of reward-based stimulus–response learning. Moreover, it does not include hyperdirect and cortico-thalamic pathways. Our model, in contrast, allows us to focus on learning, synaptic plasticity, the effects of dopamine, and lesions of all three basal ganglia pathways, thereby surmounting the capabilities of their model in these respects.

Model predictions

A specific model prediction is detailed in Results. In general, our model predicts that lesions of direct pathway striatal MSNs result in reduced facilitation of intended responses, whereas lesions of indirect pathway MSNs cause deficits in inhibiting dominant response tendencies (i.e. response tendencies that were rewarded in the past, but that no longer result in rewards). The effects of lesions of the indirect pathway may be evident not only from deficits in extinction learning, but also from impairments in reversing stimulus–response associations, i.e. in learning new response associations for known stimuli (Fig. 6B). These predictions are more specific than those put forward in previous models that had suggested overshooting motor activity as such to be a consequence of lesions of indirect pathway

MSNs (e.g. Albin *et al.*, 1989; DeLong, 1990). According to our predictions, HD patients should show overshooting motor activity (i.e. loss of inhibition) specifically when a dominant, overlearned response option exists. Reduced motor activity, in contrast, should be prominent primarily in novel or unusual contexts.

As basal ganglia pathways are known to control not only motor but also cognitive and motivational functions in separate cortico-basal ganglia-thalamic loops (Haber, 2003), our model predictions can be generalised to cognitive and motivational functions. The difficulties in activating cognitive and motivational states that are experienced by HD patients in addition to their motor symptoms should result from lesions of direct pathway MSNs in cognitive and motivational loops, respectively, whereas deficits in inhibiting dominant cognitive or motivational states should be caused by lesions of indirect pathway MSNs.

Limitations and advantages of our approach

Computational modeling offers the opportunity to study the effects of various basal ganglia dysfunctions on behavior. In the brain, such investigations would involve major technical challenges, would be ethically suitable only in animal models of HD, and would therefore be restricted to simple behavioral paradigms. As a downside, however, neuro-computational simulations can only be as good as the assumptions used in the underlying computational model. Although, of course, any of our assumptions might have to be revised in the future on the basis of novel empirical insights, we carefully designed each assumption according to a wealth of existing empirical data (Schroll *et al.*, 2014). Moreover, we developed and published our neuro-computational model well before we simulated the effects of HD-like lesions. Therefore, our model assumptions were not influenced by considerations on how to reproduce the stimulus–response learning deficits of HD patients. Further underlining the validity of our assumptions, moreover, our model has already performed well in predicting the effects of dopamine loss in Parkinson's disease and in reproducing findings on how Parkinsonian symptoms can be alleviated by specific basal ganglia lesions (Verleger *et al.*, 2013; Schroll *et al.*, 2014).

Our model is capable of coping with only a specific subset of neuro-psychological paradigms, namely those that involve stimulus–response learning. Our simulations show that combined lesions of striatal MSNs of the direct and indirect basal ganglia pathways can reproduce HD patients' deficits on these paradigms, as outlined above. However, we could not investigate whether combined lesions of the direct and indirect pathways might also explain patients' deficits on different types of paradigm. For instance, HD patients also show deficits in attention, executive function, Go–NoGo performance, and episodic memory (see Introduction). As outlined in Materials and methods, however, we consider stimulus–response learning paradigms to be well suited for investigations of the potential neuronal origins of HD patient's cognitive and motor dysfunctions in general.

Finally, our finding that only combined lesions of the direct and indirect pathways account for patients' learning impairments might exclusively relate to relatively advanced stages of the disease, as the reproduced empirical results of Lawrence *et al.* (1999), Brown *et al.* (2001) and Filoteo *et al.* (2001) were based on data from patients with clearly manifest HD. It is known that relatively advanced stages of the disease are associated with a loss of MSNs in both the direct and indirect basal ganglia pathways (Reiner *et al.*, 1988). Premanifest patients or patients in very early symptomatic stages of the disease, in contrast, might have shown different learning deficits that might have been better explained by lesions of the indirect pathway alone (Reiner *et al.*, 1988). Changes in dopamine levels or dysfunctions of the hyperdirect pathway, however, most probably would not have explained these patients' deficits either.

Conclusions

By performing neuro-computational simulations, we found that combined lesions of direct pathway and indirect pathway striatal MSNs can explain HD patients' stimulus–response learning deficits. Whereas lesions of the direct pathway explain their impairments in facilitating novel desired responses, lesions of the indirect pathway explain their deficits in inhibiting undesired but previously relevant (and therefore still dominant) response tendencies. Lesions of the hyperdirect pathway and changes in dopamine levels did not explain patients' impairments. Tentatively, these results suggest that combined lesions of the direct and indirect pathways constitute the most prominent cause of HD patients' cognitive and motor impairments in general.

Acknowledgements

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft), grant 'Deep Brain Stimulation. Mechanisms of Action, Cortex-Basal Ganglia-Physiology and Therapy Optimization' (KFO 247: DFG HA2630/7-2), and also in part by grant 'German-Japanese Collaboration Computational Neuroscience: The Function and Role of Basal Ganglia Pathways: From Single to Multiple Loops' (DFG HA2630/8-1) and by grants BE4045/10-1 and 10-2 (Deutsche Forschungsgemeinschaft).

Abbreviations

GPe, globus pallidus externus; GPi, globus pallidus internus; HD, Huntington's disease; LTD, long-term depression; LTP, long-term potentiation; MSN, medium spiny neuron; NMDA, *N*-methyl-D-aspartate; SNc, substantia nigra compacta; SNr, substantia nigra reticulata; STN, subthalamic nucleus.

References

Albin, R.L., Young, A.B. & Penney, J.B. (1989) The functional anatomy of basal ganglia disorders. *Trends Neurosci.*, **12**, 166–175.

Albin, R.L., Reiner, A., Anderson, K.D., Dure, L.S., Handelin, B., Balfour, R., Whetsell, B.O., Penney, J.B. & Young, A.B. (1992) Preferential loss of striato-external pallidal projection neurons in presymptomatic Huntington's disease. *Ann. Neurol.*, **31**, 425–430.

Alexander, G.E., Crutcher, M.D. & DeLong, M.R. (1989) Basal ganglia–thalamocortical circuits: parallel substrates for motor, oculomotor, 'prefrontal' and 'limbic' functions. *Prog. Brain Res.*, **85**, 119–146.

André, V.M., Cepeda, C. & Levine, M.S. (2010) Dopamine and glutamate in Huntington's disease: a balancing act. *CNS Neurosci. Ther.*, **16**, 163–178.

André, V.M., Fisher, Y.E. & Levine, M.S. (2011) Altered balance of activity in the striatal direct and indirect pathways in mouse models of Huntington's disease. *Front. Syst. Neurosci.*, **5**, 46.

Antonini, A., Leenders, K.L., Spiegel, R., Meier, D., Vontobel, P., Weigell-Weber, M., Snachez-Pernaute, R., de Yébenez, J.G., Boesiger, P., Weindl, A. & Maguire, R.P. (1996) Striatal glucose metabolism and dopamine D2 receptor binding in asymptomatic gene carriers and patients with Huntington's disease. *Brain*, **119**, 2085–2095.

Antzoulatos, E.G. & Miller, E.K. (2011) Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron*, **71**, 243–249.

Bäckman, L., Robins-Wahlin, T.B., Lundin, A., Ginovart, N. & Farde, L. (1997) Cognitive deficits in Huntington's disease are predicted by dopaminergic PET markers and brain volumes. *Brain*, **120**, 2207–2217.

Beste, C., Saft, C., Andrich, J., Müller, T., Gold, R. & Falkenstein, M. (2007) Time processing in Huntington's disease: a group-control study. *PLoS One*, **2**, e1263.

Beste, C., Saft, C., Andrich, J., Gold, R. & Falkenstein, M. (2008a) Response inhibition in Huntington's disease – a study using ERPs and sLORETA. *Neuropsychologia*, **46**, 1290–1297.

Beste, C., Saft, C., Andrich, J., Gold, R. & Falkenstein, M. (2008b) Stimulus–response compatibility in Huntington's disease: a cognitive–neurophysiological analysis. *J. Neurophysiol.*, **99**, 1213–1223.
Beste, C., Saft, C., Güntürkün, O. & Falkenstein, M. (2008c) Increased cognitive functioning in symptomatic Huntington's disease as revealed by behavioral and event-related potential indices of auditory sensory memory and attention. *J. Neurosci.*, **28**, 11695–11702.
Beste, C., Willemssen, R., Saft, C. & Falkenstein, M. (2010) Response inhibition subprocesses and dopaminergic pathways: basal ganglia disease effects. *Neuropsychologia*, **48**, 366–373.
Beste, C., Ness, V., Falkenstein, M. & Saft, C. (2011) On the role of fronto-striatal neural synchronization processes for response inhibition – evidence from ERP phase-synchronization analyses in pre-manifest Huntington's disease gene mutation carriers. *Neuropsychologia*, **49**, 3484–3493.
Beste, C., Ness, V., Lukas, C., Hoffmann, R., Stüwe, S., Falkenstein, M. & Saft, C. (2012) Mechanisms mediating parallel action monitoring in fronto-striatal circuits. *NeuroImage*, **62**, 137–146.
Beste, C., Stock, A.K., Ness, V., Hoffmann, R., Lukas, C. & Saft, C. (2013) A novel cognitive–neurophysiological state biomarker in premanifest Huntington's disease validated on longitudinal data. *Sci. Rep.*, **3**, 1797.
Beste, C., Humphries, M. & Saft, C. (2014) Striatal disorders dissociate mechanisms of enhanced and impaired response selection – evidence from cognitive neurophysiology and computational modelling. *Neuroimage Clin.*, **4**, 623–634.
Braak, H. & Del Tredici, K. (2008) Cortico-basal ganglia–cortical circuitry in Parkinson's disease reconsidered. *Exp. Neurol.*, **212**, 226–229.
Brown, R.G., Redondo-Verge, L., Chacon, J.R., Lucas, M.L. & Channon, S. (2001) Dissociation between intentional and incidental sequence learning in Huntington's disease. *Brain*, **124**, 2188–2202.
Canteras, N.S., Shammah-Lagnado, S.J., Silva, B.A. & Ricardo, J.A. (1990) Afferent connections of the subthalamic nucleus: a combined retrograde and anterograde horseradish peroxidase study in the rat. *Brain Res.*, **513**, 43–59.
Delmaire, C., Dumas, E.M., Sharman, M.A., den Bogaard, S.J., Valabregue, R., Jauffret, C., Justo, D., Reilmann, R., Stout, J.C., Craufurd, D., Trabrizi, S.J., Roos, R.A.C., Durr, A. & Lehéricy, S. (2013) The structural correlates of functional deficits in early Huntington's disease. *Hum. Brain Mapp.*, **34**, 2141–2153.
DeLong, M.R. (1990) Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.*, **13**, 281–285.
Denny-Brown, D. (1960) Diseases of the basal ganglia: their relation to disorders of movement. *Lancet*, **276**, 1099–1105.
Ehrlich, M.E. (2012) Huntington's disease and the striatal medium spiny neuron: cell-autonomous and non-cell-autonomous mechanisms of disease. *Neurotherapeutics*, **9**, 270–284.
Filoteo, J.V., Maddox, W.T. & Davis, J.D. (2001) A possible role of the striatum in linear and nonlinear category learning: evidence from patients with Hungtington's disease. *Behav. Neurosci.*, **115**, 786–798.
Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T. & Hutchison, K.E. (2007) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. USA*, **104**, 16311–16316.
Freedman, D.J., Riesenhuber, M., Poggio, T. & Miller, E.K. (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, **291**, 312–316.
Glass, M., Dragunow, M. & Faull, R.L.M. (2000) The pattern of neurodegeneration in Huntington's disease: a comparative study of cannabinoid, dopamine, adenosine and $GABA_A$ receptor alterations in the human basal ganglia in Huntington's disease. *Neuroscience*, **97**, 505–519.
Glimcher, P.W. (2011) Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. USA*, **108**, 15647–15654.
Haber, S.N. (2003) The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.*, **26**, 317–330.
Hartmann-von Monakow, K., Akert, K. & Künzle, H. (1978) Projections of the precentral motor cortex and other cortical areas of the frontal lobe to the subthalamic nucleus in the monkey. *Exp. Brain Res.*, **33**, 395–403.
Hazrati, L.N. & Parent, A. (1992a) Convergence of subthalamic and striatal efferents at pallidal level in primates: an anterograde double-labeling study with biocytin and PHA-L. *Brain Res.*, **569**, 336–340.
Hazrati, L.N. & Parent, A. (1992b) The striatopallidal projection displays a high degree of anatomical specificity in the primate. *Brain Res.*, **592**, 213–227.
Ho, A.K., Sahakian, B.J., Robbins, T.W., Barker, R.A., Rosser, A.E. & Hodges, J.R. (2002) Verbal fluency in Huntington's disease: a longitudinal analysis of phonemic and semantic clustering and switching. *Neuropsychologia*, **40**, 1277–1284.
Lawrence, A.D., Sahakian, B.J., Rogers, R.D., Hodges, J.R. & Robbins, T.W. (1999) Discrimination, reversal, and shift learning in Huntington's disease: mechanisms of impaired response selection. *Neuropsychologia*, **37**, 1359–1374.
Maia, T.V. & Frank, M.J. (2011) From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.*, **14**, 154–162.
Montoya, A., Pelletier, M., Menear, M., Duplessis, E., Richer, F. & Lepage, M. (2006a) Episodic memory impairment in Huntington's disease: a meta-analysis. *Neuropsychologia*, **44**, 1984–1994.
Montoya, A., Price, B.H., Menear, M. & Lepage, M. (2006b) Brain imaging and cognitive dysfunctions in Huntington's disease. *J. Psychiatr. Neurosci.*, **31**, 21–29.
O'Kusky, J.R., Nasir, J., Cicchetti, F., Parent, A. & Hayden, M.R. (1999) Neuronal degeneration in the basal ganglia and loss of pallido-subthalamic synapses in mice with targeted disruption of the Huntington's disease gene. *Brain Res.*, **818**, 468–479.
Peinemann, A., Schuller, S., Pohl, C., Jahn, T., Weindl, A. & Kassubek, J. (2005) Executive dysfunction in early stages of Huntington's disease is associated with striatal and insular atrophy: a neuropsychological and voxel-based morphometric study. *J. Neurol. Sci.*, **239**, 11–19.
Raymond, L.A., André, V.M., Cepeda, C., Gladding, C.M., Milnerwood, A.J. & Levine, M.S. (2011) Pathophysiology of Huntington's disease: time-dependent alterations in synaptic and receptor function. *Neuroscience*, **198**, 252–273.
Reiner, A., Albin, R.L., Anderson, K.D., D'Amato, C.J., Penney, J.B. & Young, A.B. (1988) Differential loss of striatal projection neurons in Huntington disease. *Proc. Natl. Acad. Sci. USA*, **85**, 5733–5737.
Schroll, H. & Hamker, F.H. (2013) Computational models of basal-ganglia pathway functions: focus on functional neuroanatomy. *Front. Syst. Neurosci.*, **7**, 122.
Schroll, H., Vitay, J. & Hamker, F.H. (2014) Dysfunctional and compensatory synaptic plasticity in Parkinson's disease. *Eur. J. Neurosci.*, **39**, 688–702.
Shen, W., Flajolet, M., Greengard, P. & Surmeier, D.J. (2008) Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*, **321**, 848–851.
Smith, Y., Bevan, M.D., Shink, E. & Bolam, J.P. (1998) Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*, **86**, 353.
Sprengelmeyer, R., Lange, H. & Hömberg, V. (1995) The pattern of attentional deficits in Huntington's disease. *Brain*, **118**, 145–152.
Storey, E. & Beal, M.F. (1993) Neurochemical substrates of rigidity and chorea in Huntington's disease. *Brain*, **116**, 1201–1222.
Tabrizi, S.J., Reilmann, R., Roos, R.A., Durr, A., Leavitt, B., Owen, G., Jones, R., Johnson, H., Craufurd, D., Hicks, S.L., Kennard, C., Landwehrmeyer, B., Stout, J.C., Borowsky, B., Scahill, R.I., Frost, C. & Langbehn, D.R. (2012) Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. *Lancet Neurol.*, **11**, 42–53.
Tabrizi, S.J., Scahill, R.I., Owen, G., Durr, A., Leavitt, B.R., Roos, R.A., Borowsky, B., Landwehrmeyer, B., Frost, C., Johnson, H., Craufurd, D., Reilmann, R., Stout, J.C. & Langbehn, D.R. (2013) Predictors of phenotypic progression and disease onset in premanifest and early-stage Huntington's disease in the TRACK-HD study: analysis of 36-month observational data. *Lancet Neurol.*, **12**, 637–649.
Thompson, P.D., Berardelli, A., Rothwell, J.C., Day, B.L., Dick, J.P.R., Benecke, R. & Marsden, C.D. (1988) The coexistence of bradykinesia and chorea in Huntington's disease and its implications for theories of basal ganglia control of movement. *Brain*, **111**, 223–244.
Tomkins, A., Vasilaki, E., Beste, C., Gurney, K. & Humphries, M.D. (2013) Transient and steady-state selection in the striatal microcircuit. *Front. Comput. Neurosci.*, **7**, 192.
Van Oostrom, J.C.H., Maguire, R.P., Verschuuren-Bemelmans, C.C., Veenma-Van Der Duin, L., Pruim, J., Roos, R.A.C. & Leenders, K.L. (2005) Striatal dopamine D2 receptors, metabolism, and volume in preclinical Huntington disease. *Neurology*, **65**, 941–943.
Verleger, R., Schroll, H. & Hamker, F.H. (2013) The unstable bridge from stimulus processing to correct responding in Parkinson's disease. *Neuropsychologia*, **51**, 2512–2525.
Waelti, P., Dickinson, A. & Schultz, W. (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, **412**, 43–48.
Weeks, R.A., Piccini, P., Harding, A.E. & Brooks, D.J. (1996) Striatal D1 and D2 dopamine receptor loss in asymptomatic mutation carriers of Huntington's disease. *Ann. Neurol.*, **40**, 49–54.

Appendix A

We here present the full mathematical setup of our neuro-computational model. All equations are equivalent to those described in Schroll *et al.* (2014), except for the changes described in Materials and methods.

Membrane potentials and firing rates

Membrane potentials of postsynaptic neurons, $m_{i,t}^{\text{post}}$, are determined via

$$\tau \cdot \frac{\mathrm{d}m_{i,t}^{\text{post}}}{\mathrm{d}t} + m_{i,t}^{\text{post}} = \sum_{j \in \text{bre}} (w_{i,j,t}^{\text{pre-post}} \cdot r_{j,t}^{\text{pre}}) + B + \varepsilon_{i,t} \qquad (A1)$$

where τ = 10 ms is a time constant, $w_{i,j,t}^{\text{pre-post}}$ is the weight from the presynaptic neuron j to the postsynaptic neuron i at time t, $r_{i,t}^{\text{pre}}$ is the firing rate of the presynaptic neuron j, $w_{i,j,t}^{\text{pre-post}}$ is the weight from the presynaptic to the postsynaptic neuron, B is the postsynaptic neuron's baseline membrane potential, and $\varepsilon_{i,t}$ is a random noise term, drawn from a uniform distribution. This noise term ensures the exploration of network states, and thereby guarantees the exploration of different response alternatives. It is the origin of variability in behavioral performance among networks. Specific values for the parameters of Eqn A1 are shown in Table 1.

As an exception, membrane potentials of SNc neurons, $m_{i,t}^{\text{SNc}}$, are determined via

$$\tau \cdot \frac{\mathrm{d}m_{i,t}^{\text{SNc}}}{\mathrm{d}t} + m_{i,t}^{\text{SNc}} = P_t \cdot \left(R_t + Q_t \cdot \sum_{j \in \text{Str}(D1)} (w_{i,j,t}^{\text{Str}(D1)-\text{SNc}} \cdot r_{j,t}^{\text{Str}(D1)}) \right) + B \qquad (A2)$$

where τ = 10 ms is a time constant, and P_t is set to 1 whenever reward can occur and to 0 otherwise, R_t is set to $(1 - B)$ whenever reward can occur and to 0 otherwise, and Q_t is set to 1 whenever reward can occur to 10 otherwise, as described in Schroll *et al.* (2014). B = 0.1 is the baseline membrane potential of SNc neurons.

Firing rates of postsynaptic neurons, $r_{i,t}^{\text{post}}$, are computed from membrane potentials via

$$r_{i,t}^{\text{post}} = (m_{i,t}^{\text{post}})^{+} \qquad (A3)$$

where $()^{+}$ determines that all negative values are set to 0. As an exception, firing rates of thalamic neurons, $r_{i,t}^{\text{Thal}}$, are computed from thalamic membrane potentials via

$$r_{i,t}^{\text{Thal}} = \begin{cases} (m_{i,t}^{\text{Thal}})^{+} & \text{if } m_{i,t}^{\text{Thal}} \leq 1 \\ \left(0.5 + \frac{1}{1 + e^{\frac{1 - m_{i,t}^{\text{Thal}}}{2}}}\right)^{+} & \text{else} \end{cases} \qquad (A4)$$

as described in Schroll *et al.* (2014).

Strengths of synaptic contacts

The strengths of synaptic contacts between presynaptic and postsynaptic neurons are determined via a different set of differential equations. Below, we show these equations separately for each type of connection in the model. This allows for easier reading than the table-based format used in Schroll *et al.* (2014).

Cortico-thalamic synapses

The strengths of synaptic contacts between stimulus cortical and thalamic neurons, $w_{i,j,t}^{\text{Cx-Thal}}$, are computed via

$$\eta \cdot \frac{\mathrm{d}w_{i,j,t}^{\text{Cx-Thal}}}{\mathrm{d}t} = \text{Ca}_{i,j,t}^{\text{Cx-Thal}} - \alpha_{i,t}^{\text{Cx-Thal}} \cdot \text{Ca}_{i,j,t}^{\text{Cx-Thal}} \qquad (A5)$$

With

$$\alpha_{i,t}^{\text{Cx-Thal}} = (m_{i,t}^{\text{Thal}} - m^{\text{MAX}})^{+} \qquad (A6)$$

where η = 2000 is a time constant, $\text{Ca}_{i,j,t}^{\text{Cx-Thal}}$ denotes a calcium trace as defined in Eqn A7, and $\alpha_{i,t}^{\text{Cx-Thal}}$ ensures that synaptic contacts cannot increase in strength infinitely, but may reach a dynamic threshold as the postsynaptic membrane potential crosses a threshold defined by m^{MAX} = 0.9. Cortico-thalamic weights are not allowed to become negative.

Calcium traces are determined via

$$\text{Ca}_{i,j,t}^{\text{Cx-Thal}} = (r_{j,t}^{\text{Cx}} - \overline{\text{Cx}_{\text{t}}}) \cdot (r_{i,t}^{\text{Thal}} - \overline{\text{Thal}_{\text{t}}} - \gamma^{\text{Thal}})^{+} \qquad (A7)$$

where $r_{j,t}^{\text{Cx}}$ is the firing rate of the presynaptic cortical neuron j at time t, $\overline{\text{Cx}_{\text{t}}}$ is the average firing rate of the stimulus cortex at time t, $r_{i,t}^{\text{Thal}}$ is the firing rate of the thalamic neuron i at time t, $\overline{\text{Thal}_{\text{t}}}$ is the average firing rate of the thalamus at time t, γ^{Thal} is the postsynaptic activity threshold (which was set to 0.3 for two thalamic neurons, to 0.47 for three thalamic neurons, to 0.55 for four thalamic neurons, and to 0.60 for five thalamic neurons, as described in Materials and methods), and $()^{+}$ determines that negative values are set to zero.

Cortico-striatal synapses

The strengths of connections between the stimulus cortex and striatal MSNs, $w_{i,j,t}^{\text{Cx-Str}}$, are determined via

$$\eta \cdot \frac{\mathrm{d}w_{i,j,t}^{\text{Cx-Str}}}{\mathrm{d}t} = f_{\text{DA}}(\text{DA}_t - 0.1) \cdot \text{Ca}_{i,j,t}^{\text{Cx-Str}} - \alpha_{i,t}^{\text{Cx-Str}} \cdot \text{Ca}_{i,j,t}^{\text{Cx-Str}} \qquad (A8)$$

with

$$\alpha_{i,t}^{\text{Cx-Str}} = (m_{i,t}^{\text{Str}} - m^{\text{MAX}})^{+} \qquad (A9)$$

where Str can refer either to D1 MSNs or to D2 MSNs. DA_t is the dopamine level at time t, $\text{Ca}_{i,j,t}^{\text{Cx-Str}}$ is the calcium trace for the synapse of interest, and $\alpha_{i,t}^{\text{Cx-Str}}$ determines that weights decrease as the postsynaptic membrane potential crosses a threshold defined by m^{MAX} = 1.0. Both formulae are equivalent for cortical synapses to D1 and D2 MSNs. Calcium traces are computed via

$$\eta^{\text{Ca}} \cdot \frac{\mathrm{d}\text{Ca}_{i,j,t}^{\text{Cx-Str}}}{\mathrm{d}t} + \text{Ca}_{i,j,t}^{\text{Cx-Str}} = (r_{j,t}^{\text{Cx}} - \overline{\text{Cx}_t} - \gamma^{\text{Cx}}) \cdot (r_{i,t}^{\text{Str}} + \overline{\text{Str}_t})^{+} \qquad (A10)$$

with

$$\eta^{\mathrm{Ca}} = \begin{cases} 250 & \text{if } \eta^{\mathrm{Ca}} \cdot \frac{\mathrm{dCa}_{i,j,t}^{\mathrm{Cx-Str}}}{\mathrm{d}t} + \mathrm{Ca}_{i,j,t}^{\mathrm{Cx-Str}} = 0 \\ 1 & \text{else} \end{cases} \quad (A11)$$

again equivalently for D1 and D2 MSNs, where $r_{j,t}^{\mathrm{Cx}}$ is the firing rate of the presynaptic stimulus cortical neuron j at time t, $\overline{\mathrm{Cx}_t}$ is the average firing rate of the stimulus cortex at time t, $\gamma^{\mathrm{Cx}} = 0.15$ is the presynaptic activity threshold, $r_{i,t}^{\mathrm{Str}}$ is the firing rate of the postsynaptic striatal neuron i at time t, and $\overline{\mathrm{Str}_t}$ is the average firing rate of the striatal neurons in question (i.e. the average firing rate of all D1 MSNs if the evaluated synapse belongs to a D1 MSN, and the average firing rate of all D2 MSNs if the synapse belongs to a D2 MSN).

The effects of dopamine on synaptic plasticity differ between D1 and D2 MSNs. For D1 MSNs, the following function determines the effects of dopamine:

$$f_{\mathrm{DA}}(x) = \begin{cases} 2x & \text{if } x > 0 \\ 0.8x & \text{if } x < 0 \cap \\ & ((\mathrm{Ca}_{i,j,t}^{\mathrm{Cx-Str}(D1)} > 0 \cap w_{i,j,t}^{\mathrm{Cx-Str}(D1)} > 0) \\ & \cup (\mathrm{Ca}_{i,j,t}^{\mathrm{Cx-Str}(D1)} < 0 \cap w_{i,j,t}^{\mathrm{Cx-Str}(D1)} < 0)) \\ 0 & \text{else} \end{cases} \quad (A12)$$

whereas for D2 MSNs, the following function holds true, as described in Schroll *et al.* (2014):

$$f_{\mathrm{DA}}(x) = \begin{cases} -2x & \text{if } x < 0 \\ -0.8x & \text{if } x > 0 \cap \\ & ((\mathrm{Ca}_{i,j,t}^{\mathrm{Cx-Str}(D2)} > 0 \cap w_{i,j,t}^{\mathrm{Cx-Str}(D2)}) \\ & \cup (\mathrm{Ca}_{i,j,t}^{\mathrm{Cx-Str}(D2)} < 0 \cap w_{i,j,t}^{\mathrm{Cx-Str}(D2)} < 0)) \\ 0 & \text{else} \end{cases} \quad (A13)$$

Cortico-subthalamic synapses

The strengths of cortico-subthalamic synapses are determined in exactly the same way as the strengths of synapses between the cortex and D1 MSNs.

Striato-nigral synapses

The strengths of synapses between striatal D1 MSNs and the SNc, $w_{i,j,t}^{\mathrm{Str}(D1)-\mathrm{SNc}}$, are determined via

$$\eta \cdot \frac{\mathrm{d}w_{i,j,t}^{\mathrm{Str}(D1)-\mathrm{SNc}}}{\mathrm{d}t} = -f_{\mathrm{DA}}(\mathrm{DA}_t - 0.1) \cdot (r_{j,t}^{\mathrm{Str}(D1)} - \overline{\mathrm{Str}(D1)_t})^{+} \quad (A14)$$

with

$$f_{\mathrm{DA}}(x) = \begin{cases} x & \text{if } x > 0 \\ 3x & \text{else} \end{cases} \quad (A15)$$

where η = 100 000 is a time constant, DA_t is the dopamine level at time t, $r_{j,t}^{\mathrm{Str}(D1)}$ is the firing rate of the presynaptic striatal D1 MSN j at time t, $\overline{\mathrm{Str}(D1)_t}$ is the average firing rate of all striatal D1 MSNs at time t, and $()^{+}$ determines that all negative values are set to 0. As striato-nigral synapses are GABAergic, their weights are prevented from increasing above 0. Functionally, Eqns A14 and A15 determine that striato-nigral synapses increase in strength after phasic increases in dopamine (that is, they become more negative after unexpected rewards). As a consequence, the magnitudes of future phasic increases in SNc firing are reduced as rewards become predictable.

Striato-pallidal synapses

The strengths of synaptic contacts between striatal D1 MSNs and the GPi, and between striatal D2 MSNs and the GPe, $w_{i,j,t}^{\mathrm{Str-GP}}$, are determined via

$$\eta \cdot \frac{\mathrm{d}w_{i,j,t}^{\mathrm{Str-GP}}}{\mathrm{d}t} = -f_{\mathrm{DA}}(\mathrm{DA}_t - 0.1) \cdot \mathrm{Ca}_{i,j,t}^{\mathrm{Str-GP}} - \alpha_{i,t}^{\mathrm{Str-GP}} \cdot \mathrm{Ca}_{i,j,t}^{\mathrm{Str-GP}} \quad (A16)$$

with

$$\alpha_{i,t}^{\mathrm{Str-GP}} = -(-m_{i,t}^{\mathrm{GP}} - m^{\mathrm{MAX}})^{+} \quad (A17)$$

where Str can refer to either D1 or D2 MSNs, and GP can either refer to the GPi or the GPe. η = 50 is a learning rate, DA_t is the dopamine level at time t, $\mathrm{Ca}_{i,t}^{\mathrm{Str-GP}}$ is the calcium trace between the presynaptic neuron j and the postsynaptic neuron i at time t as determined from Eqn A18, and $\alpha_{i,t}^{\mathrm{Str-GP}}$ reduces synaptic weights as postsynaptic firing crosses a threshold defined by m^{MAX} (with m^{MAX} = 1.0 for synapses between D1 MSNs and GPi neurons, and m^{MAX} = 2.0 for synapses between D2 MSNs and GPe neurons). Weights are not allowed to increase above 0.

Calcium traces are computed via

$$\eta^{\mathrm{Ca}} \cdot \frac{\mathrm{dCa}_{i,j,t}^{\mathrm{Str-GP}}}{\mathrm{d}t} + \mathrm{Ca}_{i,j,t}^{\mathrm{Str-GP}} = (r_{j,t}^{\mathrm{Str}} - \overline{\mathrm{Str}_t})^{+} \cdot (-r_{i,t}^{\mathrm{GP}} + \overline{\mathrm{GP}_t} - \gamma^{\mathrm{GP}}) \quad (A18)$$

with

$$\eta^{\mathrm{Ca}} = \begin{cases} 250 & \text{if } \eta^{\mathrm{Ca}} \cdot \frac{\mathrm{dCa}_{i,j,t}^{\mathrm{Str-GP}}}{\mathrm{d}t} + \mathrm{Ca}_{i,j,t}^{\mathrm{Str-GP}} = 0 \\ 1 & \text{else} \end{cases} \quad (A19)$$

where $r_{j,t}^{\mathrm{Str}}$ is the firing rate of the striatal MSN j at time t, $\overline{\mathrm{Str}_t}$ is the average firing rate of all striatal MSNs of the respective type (i.e. either D1 or D2 MSNs), $r_{i,t}^{\mathrm{GP}}$ is the firing rate of the GPi or GPe neuron i at time t, $\overline{\mathrm{GP}_t}$ is the average firing rate of all GPi or GPe neurons, as appropriate, and γ^{GP} = 0.15 is the postsynaptic activity threshold.

The dopamine function $f_{\mathrm{DA}}(x)$ differs between synapses that connect D1 MSNs to the GPi and those that connect D2 MSNs to the GPe. For connections between D1 MSNs and GPi neurons, the following function is valid:

$$f_{\mathrm{DA}}(x) = \begin{cases} 2x & \text{if } x > 0 \\ 0.8x & \text{if } x < 0 \cap \mathrm{Ca}_{i,j,t}^{\mathrm{Str}(D1)-\mathrm{GPi}} > 0 \\ 0 & \text{else.} \end{cases} \quad (A20)$$

whereas for connections between D2 MSNs and GPe neurons, the following function holds true:

$$f_{\mathrm{DA}}(x) = \begin{cases} -2x & \text{if } x<0 \\ -0.8x & \text{if } x>0 \cap \mathrm{Ca}_{i,j,t}^{\mathrm{Str}(D2)-\mathrm{GPe}} > 0 \\ 0 & \text{else.} \end{cases} \quad (A21)$$

Subthalamo-pallidal synapses

Subthalamo-pallidal weights, $w_{i,j,t}^{\mathrm{STN-GPi}}$, are determined in strength via

$$\eta \cdot \frac{\mathrm{d}w_{i,j,t}^{\mathrm{STN-GPi}}}{\mathrm{d}t} = f_{\mathrm{DA}}(\mathrm{DA}_t - 0.1) \cdot \mathrm{Ca}_{i,j,t}^{\mathrm{STN-GPi}} - \alpha_{i,t}^{\mathrm{STN-GPi}} \cdot \mathrm{Ca}_{i,j,t}^{\mathrm{STN-GPi}} \quad (A22)$$

with

$$\alpha_{i,t}^{\mathrm{STN-GPi}} = (m_{i,t}^{\mathrm{GPi}} - m^{\mathrm{MAX}})^{+} \quad (A23)$$

and

$$f_{\mathrm{DA}}(x) = \begin{cases} 2x & \text{if } x>0 \\ 0.8x & \text{if } x<0 \cap \mathrm{Ca}_{i,j,t}^{\mathrm{STN-GPi}} > 0 \\ 0 & \text{else.} \end{cases} \quad (A24)$$

where $\eta = 50$ is a learning rate, DA_t is the dopamine level at time t, $\mathrm{Ca}_{i,j,t}^{\mathrm{STN-GPi}}$ is the calcium trace between the STN neuron j and the GPi neuron i at time t, and $\alpha_{i,t}^{\mathrm{STN-GPi}}$ prevents further increases in synaptic weights as the postsynaptic membrane potential crosses a threshold defined by $m^{\mathrm{MAX}} = 1.5$. Weights are not allowed to decrease below 0.

Calcium traces are computed via

$$\eta^{\mathrm{Ca}} \cdot \frac{\mathrm{dCa}_{i,j,t}^{\mathrm{STN-GPi}}}{\mathrm{d}t} + \mathrm{Ca}_{i,j,t}^{\mathrm{STN-GPi}} = (r_{j,t}^{\mathrm{STN}} - \overline{\mathrm{STN}_t})^{+} \cdot (r_{i,t}^{\mathrm{GPi}} - \overline{\mathrm{GPi}_t} + \gamma^{\mathrm{GPi}}) \quad (A25)$$

with

$$\eta^{\mathrm{Ca}} = \begin{cases} 250 & \text{if } \eta^{\mathrm{Ca}} \cdot \frac{\mathrm{dCa}_{i,j,t}^{\mathrm{STN-GPi}}}{\mathrm{d}t} + \mathrm{Ca}_{i,j,t}^{\mathrm{STN-GPi}} = 0 \\ 1 & \text{else} \end{cases} \quad (A26)$$

where η^{Ca} is a learning rate, $r_{j,t}^{\mathrm{STN}}$ is the firing rate of STN neuron j at time t, $\overline{\mathrm{STN}_t}$ is the average firing rate of the STN at time t, $r_{i,t}^{\mathrm{GPi}}$ is the firing rate of the GPi neuron i at time t, $\overline{\mathrm{GPi}_t}$ is the average firing rate of the GPi at time t, $\gamma^{\mathrm{GPi}} = 0.15$ is the postsynaptic activity threshold, and $()^{+}$ determines that negative values are set to zero.

Pallido-pallidal synapses

Weights that control lateral competition in the GPi, $w_{i,j,t}^{\mathrm{GPi-GPi}}$, are determined via

$$\eta \cdot \frac{\mathrm{d}w_{i,j,t}^{\mathrm{GPi-GPi}}}{\mathrm{d}t} = (-r_{j,t}^{\mathrm{GPi}} + \overline{\mathrm{GPi}_t})^{+} \cdot (-r_{i,t}^{\mathrm{GPi}} + \overline{\mathrm{GPi}_t})^{+} - \alpha_{i,t}^{\mathrm{GPi-GPi}} \cdot (-r_{j,t}^{\mathrm{GPi}} + \overline{\mathrm{GPi}_t})^{+} \cdot (r_{i,t}^{\mathrm{GPi}} - \overline{\mathrm{GPi}_t})^{+} \quad (A27)$$

with

$$\alpha_{i,t}^{\mathrm{GPi-GPi}} = (m_{i,t}^{\mathrm{GPi}})^{+} \quad (A28)$$

where $r_{j,t}^{\mathrm{GPi}}$ is the firing rate of the presynaptic GPi neuron j at time t, $\overline{\mathrm{GPi}_t}$ is the average firing rate of the GPi at time t, $r_{i,t}^{\mathrm{GPi}}$ is the activity of the postsynaptic GPi neuron i at time t, $\alpha_{i,t}^{\mathrm{GPi-GPi}}$ determines that weights do not increase infinitely, and $()^{+}$ determines that negative values are set to 0. Note that, for the reversal learning and the category learning paradigms, where only two GPi neurons are present (Table 1), lateral competition in the GPi was set to 1.25 (that is, synaptic plasticity was not evaluated), as described in Materials and methods. Pallido-pallidal weights are prevented from decreasing below zero.

Overt responses

Overt responses are determined from motor cortical firing rates via a softmax rule, whereby the probability of the correct response i at time t, $P_{i,t}$, is given by

$$P_{i,t} = \frac{r_{i,t}^{\mathrm{motor\,Cx}} + \theta}{\sum_{j \in \mathrm{motor\,Cx}} (r_{j,t}^{\mathrm{motor\,Cx}} + \theta)} \quad (A29)$$

where $r_{i,t}^{\mathrm{motor\,Cx}}$ is the firing rate of motor cortical neuron i at time t, and $\theta = 10^{-10}$ prevents the denominator from becoming 0.