ARTICLE **Communicated by Samuel McDougle**

Differential Dopamine Receptor-Dependent Sensitivity Improves the Switch Between Hard and Soft Selection in a Model of the Basal Ganglia

Olivier Codol
codol.olivier@gmail.com
Department of Psychology and Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON N6A 3K7, Canada

Paul L. Gribble
pgribble@uwo.ca
Department of Psychology and Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, University of Western Ontario, London, ON N6A 3K7, Canada, and Haskins Laboratories, New Haven, CT 06511, U.S.A.

Kevin N. Gurney
k.gurney@sheffield.ac.uk
Department of Psychology, University of Sheffield, Sheffield S10 2TN, U.K.

The problem of selecting one action from a set of different possible actions, simply referred to as the problem of action selection, is a ubiquitous challenge in the animal world. For vertebrates, the basal ganglia (BG) are widely thought to implement the core computation to solve this problem, as its anatomy and physiology are well suited to this end. However, the BG still display physiological features whose role in achieving efficient action selection remains unclear. In particular, it is known that the two types of dopaminergic receptors (D1 and D2) present in the BG give rise to mechanistically different responses. The overall effect will be a difference in sensitivity to dopamine, which may have ramifications for action selection. However, which receptor type leads to a stronger response is unclear due to the complexity of the intracellular mechanisms involved. In this study, we use an existing, high-level computational model of the BG, which assumes that dopamine contributes to action selection by enabling a switch between different selection regimes, to predict which of D1 or D2 has the greater sensitivity. Thus, we ask, Assuming dopamine enables a switch between action selection

Oliver Codol is the corresponding author.

Neural Computation **34, 1588–1615** (2022)
https://doi.org/10.1162/neco_a_01517

regimes in the BG, what functional sensitivity values would result in improved action selection computation? To do this, we quantitatively assessed the model's capacity to perform action selection as we parametrically manipulated the sensitivity weights of D1 and D2. We show that differential (rather than equal) D1 and D2 sensitivity to dopaminergic input improves the switch between selection regimes during the action selection computation in our model. Specifically, greater D2 sensitivity compared to D1 led to these improvements.

1 Introduction

The need for selection usually arises when a system needs to use a quantitatively limited resource. In the case of action selection, several potential actions may require the same body effectors, making those actions mutually exclusive and requiring the brain to implement a solution to the ensuing action selection problem. The basal ganglia (BG) are now widely accepted as the brain structure that implements the computation pertaining to action selection by acting as the nervous system's central switch (Redgrave, Prescott, & Gurney, 1999; Ringo, 1991; Gurney, Prescott, & Redgrave, 2001a; Humphries, Stewart, & Gurney, 2006; Friend & Kravitz, 2014; Mink, 1996).

1.1 Overview of the Basal Ganglia Architecture. The primate BG are composed of several distinct nuclei with a complex but reasonably well-understood connectivity (Mink, 1996). The principal nuclei are the striatum, the subthalamic nucleus (STN), the globus pallidus (GP), and the substantia nigra (SN; Figure 1a). The GP is also subdivided into the internal (GPi) and external (GPe) segments and the SN into the pars reticulata (SNr) and pars compacta (SNc; see Figure 1b). The largest neural population in the SNc are dopaminergic neurons sending diffuse projections to the striatum (Smith, Bennett, Bolam, Parent, & Sadikot, 1994) and connecting to two different striatal neuron populations with a different subtype of dopamine receptors, called D1 and D2 receptors. Interestingly, D1-expressing neurons in the striatum project predominantly to the GPi/SNr while D2-expressing neurons project mainly to the GPe (see Figure 1a), with all these striatal projections being inhibitory (Smith, Bevan, Shink, & Bolam, 1998; Ericsson et al., 2013). The intrinsic circuitry of the GPe is rather complex. Several subpopulations of neurons are contained within it (Mallet et al., 2012; Sadek, Magill, & Bolam, 2007; Suryanarayana, Hellgren-Kotaleski, Grillner, & Gurney, 2019), with each subpopulation sending self-inhibitory projections and local inhibitory projections to other subpopulations within the GPe. They also send inhibitory projections outside the GPe to all other BG nuclei (Sadek et al., 2007; Suryanarayana et al., 2019), except for the arkypallidal neurons that send (diffuse) inhibitory projections only to D1- and D2-expressing striatal neurons without contacting the STN or the GPi/SNr

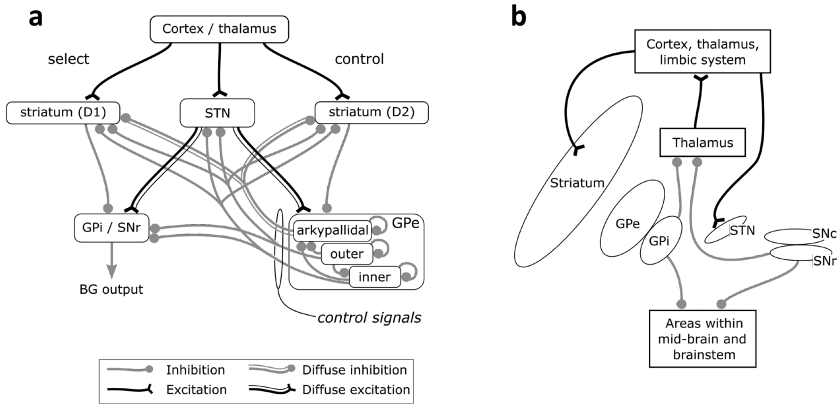

Figure 1: An overview of the anatomical architecture of the BG. (a) Intrinsic connectivity of the BG. (b) Extrinsic connectivity of the BG, showing the baso-thalamo-cortical loops. Figure modified from Gurney et al. (2001a).

(Mallet et al., 2012). Finally, the STN sends diffuse excitatory projections to all GPe subpopulations, as well as to the GPi/SNr (Shink, Bevan, Bolam, & Smith, 1996; Parent & Hazrati, 1993).

The BG receive projections from nearly the whole cortex, as well as the thalamus and limbic system (Nakano, Kayahara, Tsutsumi, & Ushiro, 2000; Coizet et al., 2009; Feger, Bevan, & Crossman, 1994; Lanciego et al., 2004; Sharpe et al., 2019; Monakow, Akert, & Kunzle, 1978), providing the required contextual information to properly define the relative value of each action (e.g., prominence of the stimulus or immediate utility) to perform the action selection computation. These various inputs enter the BG through the striatum and the STN. The main output nuclei of the BG are the GPi and SNr. The BG's default output is a tonic inhibition from the GPi/SNr to the thalamus (see Figures 1a and 1b). Actions are permitted in the BG by selectively releasing this inhibition, allowing their expression (Chevalier & Deniau, 1990).

1.2 The Basal Ganglia Architecture Supports the Action Selection Hypothesis. This anatomy of the BG is well suited for action selection (Bariselli, Fobbs, Creed, & Kravitz, 2019; Kwak & Jung, 2019). This was illustrated by Gurney, Prescott, and Redgrave in a computational model of the BG (Gurney et al., 2001a, 2001b; Suryanarayana et al., 2019), henceforth referred to as the GPR model, that reproduces the functional architecture in Figure 1a. A main assumption of the GPR model is that each action is represented in the BG by a specific stream of sensory, cognitive, and emotional information called a "channel," which represents one possible action to perform. Channels are organized in parallel, in line with physiological

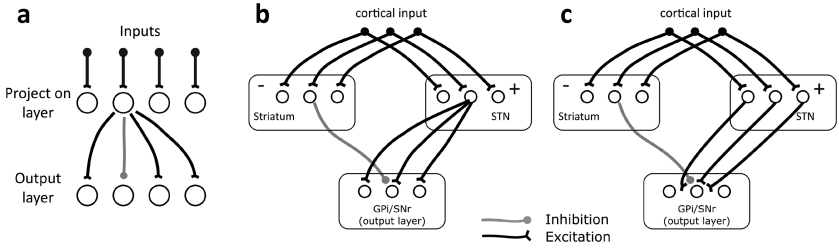

Figure 2: An illustration of a generic feedforward network that produces on-center, off-surround competition. (a) Functional architecture of a four-channel feedforward net implementing an off-center, on-surround computation. Only the output from the second neuron of the projection layer is explicitly shown. (b) Implementation of a three-channel feedforward net in the BG. Note that the focused inhibitory and diffuse excitatory component of the projection layer are split in two distinct neural populations—respectively, the striatum and the STN. (c) At the same time, each neuron in the output layer also receives excitatory input from all the neurons in the STN. Figure modified from (Gurney et al., 2001a).

and histological evidence (Graybiel, Ragsdale, Yoneoka, & Elde, 1981; Alexander, DeLong, & Strick, 1986). Those channels then compete with each other through off-center/on-surround computation in the BG circuit, with the "winning" channel eventually suppressing the other channels' activity to promote its own selection (Girard, Lienard, Gutierrez, Delord, & Doya, 2020).

A functional center/surround organization has been proposed in the BG (Mink & Thach, 1993) as a result of diffuse excitatory drive from the STN and focused inhibitory drive from the striatum to the GPi/SNr (see Figure 2a and 2b for a simplified illustration; Gurney et al., 2001a). In the GPR model, if a channel receives a stronger input from the cortex, all channels in the output layer will be more activated through that channel's diffuse STN→GPi/SNr projection (see Figure 2b), with only the more activated channel receiving equally strong inhibition from the striatum. Therefore, that channel's output will be less than the other channels' output, leading to disinhibition of the action represented by that channel. Of note, the striatum's local circuitry includes lateral inhibition, which may also contribute to the center/surround organization, although this is not included in the GPR model.

1.3 Dopamine and Action Selection. Dopaminergic input in the BG originates from the SNc, and one of its proposed functions is to modulate the action selection computation (Groves, Linder, & Young, 1994). In the GPR model, this translates into allowing the BG to switch between two selection regimes (Gurney, Humphries, Wood, Prescott, & Redgrave, 2004): a

so-called hard selection regime, which promotes selection of single actions, and a soft selection regime, which allows multiple actions to be selected. The latter may correspond in real life as a situation where competing actions are not mutually exclusive, such as picking up a fruit in each hand.

As mentioned, striatal neurons are segregated into two distinct populations expressing either D1 or D2 receptors. D1 receptors increase corticostriatal transmission, whereas D2 receptors attenuate it (Surmeier et al., 2007). They also display quantitatively different molecular affinities to dopamine, with D2 receptors binding more easily to dopamine (90% of D2 receptors in high-affinity state against a selective agonist) than D1 receptor (20% in a high-affinity state; Arbuthnott & Wickens, 2007; Richfield, Penney, & Young, 1989). However, these numbers have been challenged more recently, and what mechanistically underlies binding affinities of D1 and D2 is still unclear (Cumming, 2011; Skinbjerg, Sibley, Javitch, & Abi-Dargham, 2012). Additionally, the intracellular activation pathways differ between D1 and D2. Therefore, the overall response to dopamine at either receptor type will be the result of a complex cascade of influences, which may even work in opposing directions to the difference in proportion of high-affinity states expressed (Kenakin, 2013). It is not surprising, therefore, that attempts to measure or simulate intracellular pathway amplification have given inconclusive and sometimes opposite results (Yapo et al., 2017; Marcott, Mamaligas, & Ford, 2014; Watts et al., 1998; Watts & Neve, 1996; Kim et al., 2004; Dumartin et al., 2000). Therefore, it is currently unknown if the imbalance in the proportion of high-affinity states really translates into an imbalance in response to dopamine in D1 versus D2 receptors—and if it does, which receptor actually shows a stronger response.

In this study, we take this problem from a top-down, functional approach instead and ask, Assuming the function of the BG is to perform action selection and that dopamine modulation allows for a switch between soft and hard selection regime, what is the optimal relationship between D1 and D2 sensitivity? By "sensitivity," we encompass the amount of receptors expressed, binding affinity, and intracellular pathway amplification together and refer to it as "functional weight." To address this question, we added the different D1 and D2 functional weight to the GPR model, resulting in an augmented differential-weight model. We then varied the value of both D1 and D2 weights independently and quantified the augmented model's performance for soft and hard selection given each combination of weights. Our results supply a prediction for the overall functional sensitivity of each of the D1 and D2 responses.

2 Model Construction

While the GPR model was first proposed in Gurney et al. (2001a, 2001b), here we build up on a version of the same model described in Suryanarayana et al. (2019), which incorporates a substantially more detailed GPe

local circuitry and projections, based on more recently available anatomical data. Therefore, we refer to the model in Suryanarayana et al. (2019) as the "original model" here and use it as our control condition.

2.1 Model Code and Analysis Code Availability. The code for implementing the model, running the simulations for this study, analyzing the resulting data, and plotting the figures is freely available online at https://github.com/OlivierCodol/GPR-augmented-model.

2.2 Input to the Model. The GPR model focuses on the computations intrinsic to the BG to implement action selection. Consequently, cortical and thalamic structures are not represented in this model. Rather, it is assumed that the information entering the BG has already been preprocessed into a scalar termed "salience," as it is expected to scale with how urgent or salient a potential action may be given cognitive, sensory, and emotional evidence processed at the cortical level (Gurney et al., 2001b; Humphries et al., 2006). Here, each channel receives a specific, user-defined salience input that can vary over time. Additionally, since no noise is introduced in the model, the model is fully deterministic.

2.3 Implementation of the Differential Affinity of D1 and D2 Receptors. In the original model, the *i*th channel receives the same salience drive, c_i at both D1- and D2-striatal modules, yielding module inputs u_i^{D1} and u_i^{D2}, respectively. This input is also modulated by the shared dopaminergic input λ and multiplied by a shared corticostriatal weight w_s^{cortex} according to the following equations:

$$u_i^{D1} = w_s^{cortex}(1 + \lambda)\, c_i + \mathrm{GPe}, \tag{2.1}$$

$$u_i^{D2} = w_s^{cortex}(1 - \lambda)\, c_i + \mathrm{GPe}, \tag{2.2}$$

where u_i^{D1} and u_i^{D2} are the net, postsynaptic input to the D1- and D2-expressing neural population in the striatum. The term *GPe* here refers to the contribution from GPe projections that are not detailed here (see appendix A and see Figure 1a). Note that the dopaminergic modulation given by λ facilitates input for the striatum D1 and attenuates it for the striatum D2, in line with biological evidence (Surmeier, Ding, Day, Wang, & Shen, 2007). Also note that the formulation of dopaminergic modulation allows for the model to keep processing cortical inputs even in the absence of any dopaminergic drive, since for $\lambda = 0$, we have $(1 \pm \lambda) = 1$.

Since the model alteration in this study is at the level of the dopaminergic input λ, we focus in this section on equations 2.1 and 2.2. The full model's architecture is illustrated in Figure 1a, and the detailed mathematical formulation of the model is available in appendix A. Briefly, each nucleus in

the model contains a set of neural units, with each neural unit representing one channel. These channels receive a sum of weighted inputs, which is then passed through a saturating nonlinear output function (see equation A.2 in appendix A) to bound the output range.

To introduce the differential sensitivity to dopamine at D1 and D2 synapses, we added connectivity weights w_{D1}, w_{D2}, which combine with λ to produce receptor-type dependent modulation. Thus, in the augmented model, equations 2.1 and 2.2 become

$$u_i^{D1} = w_s^{cortex}(1 + w_{D1}\lambda)\, c_i \; + \text{GPe}, \tag{2.3}$$

$$u_i^{D2} = w_s^{cortex}(1 - w_{D2}\lambda)\, c_i \; + \text{GPe}. \tag{2.4}$$

Our experiments seek to assess the effects of these modulating factors on model selection. However, it is first necessary to define more precisely what we mean by "selection" so that we have a quantitative measure of this function that can be evaluated for different values of w_{D1}, w_{D2}.

3 Quantifying Selection Performance

3.1 Defining Selection. This section follows the method described in Gurney et al. (2004). Since the BG select via disinhibition, selection occurs on channel i if there is sufficient reduction in that channel's GPi/SNr output Y_i, compared to its tonic value Y^0. The tonic value Y^0 is the same for all channels (channel-independent) and is the output of the GPi/SNr in the absence of any input (see Figure 3). To quantify this reduction from the tonic value, we define a selection threshold θ, where $Y^0 > \theta \geq 0$, and consider that selection has occurred if $Y_i \leq \theta$. Specifically, in this instantiation of the model, we observe $Y^0 = 0.1032$ for the parameter values specified in Table 2, and we used $\theta = 0$. Finally, we consider a "distortion" threshold θ_d, with $Y^0 > \theta_d > \theta$. If we have $\theta_d > Y_i > \theta$, the outcome is not a clear selection of channel i but a distorted selection, which may lead to intereference with other selected channels, as detailed below. In this study, we used $\theta_d = 0.1 \times Y^0$.

If two channels have nonzero salience input, they are effectively competing. To evaluate the outcome of this scenario, we run the following experimental protocol (see Figure 3). At $t = 0$, no input is presented to the model for a time sufficient to let tonic outputs to be expressed and reach equilibrium. Then, at $t = 1$, one channel receives an input of salience c_1. Because of interchannel interactions, we can observe that the input to the first channel alters the activity of the second channel despite c_2 remaining constant. At $t = 2$, the output of both channels $Y_1^{t=2}$ and $Y_2^{t=2}$ has already reached equilibrium, and the second channel receives an input c_2. The simulation finishes at $t = 3$, when both the first and second channel outputs $Y_1^{t=3}$ and $Y_2^{t=3}$ have reached equilibrium again. At this point there are six possible outcomes.

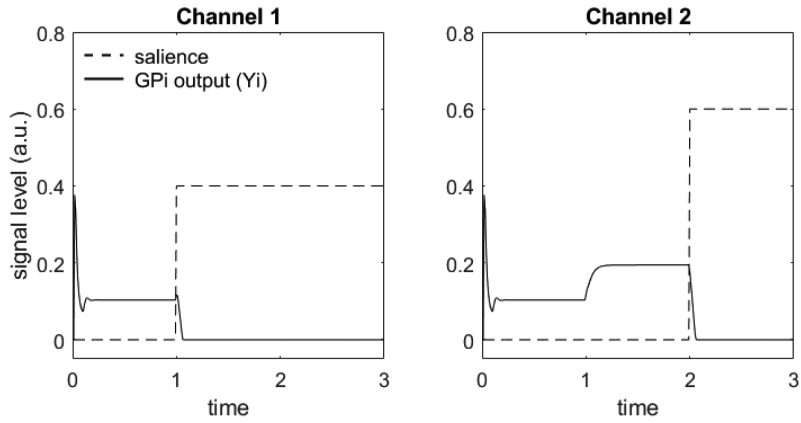

Figure 3: Salience input from the cortex to the model and dynamics of the model's GPi/SNr layer output. Note that the change in salience input in channel 1 at $t = 1$ leads to a change in output value for channel 2, and vice versa at $t = 2$. The tonic output (i.e., for no input at all in any channel) can be seen from $t = 0$ to $t = 1$.

1. *No selection.* Neither channel output has reached the selection threshold so no action is selected:

$$Y_1^{t=2} > \theta \quad \text{and} \quad Y_1^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} > \theta. \tag{3.1}$$

2. *Single-channel selection.* Only one channel is selected:

$$Y_1^{t=2} \leq \theta \quad \text{and} \quad Y_1^{t=3} \leq \theta \quad \text{and} \quad Y_2^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} > \theta_d \tag{3.2}$$

$$\text{or} \quad Y_1^{t=2} > \theta \quad \text{and} \quad Y_1^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} \leq \theta \quad \text{and} \quad Y_1^{t=3} > \theta_d. \tag{3.3}$$

3. *Dual-channel selection.* Both channels are selected:

$$Y_1^{t=2} \leq \theta \quad \text{and} \quad Y_1^{t=3} \leq \theta \quad \text{and} \quad Y_2^{t=3} \leq \theta. \tag{3.4}$$

4. *Interference.* Channel 1 is selected at first, but channel 2 forces deselection of channel 1 and is not selected in return:

$$Y_1^{t=2} \leq \theta \quad \text{and} \quad Y_1^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} > \theta. \tag{3.5}$$

5. *Channel switching.* Channel 1 is selected at first, but channel 2 forces deselection of channel 1 and is selected in return without interference:

$$Y_1^{t=2} \leq \theta \quad \text{and} \quad Y_1^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} \leq \theta \quad \text{and} \quad Y_1^{t=3} > \theta_d. \tag{3.6}$$

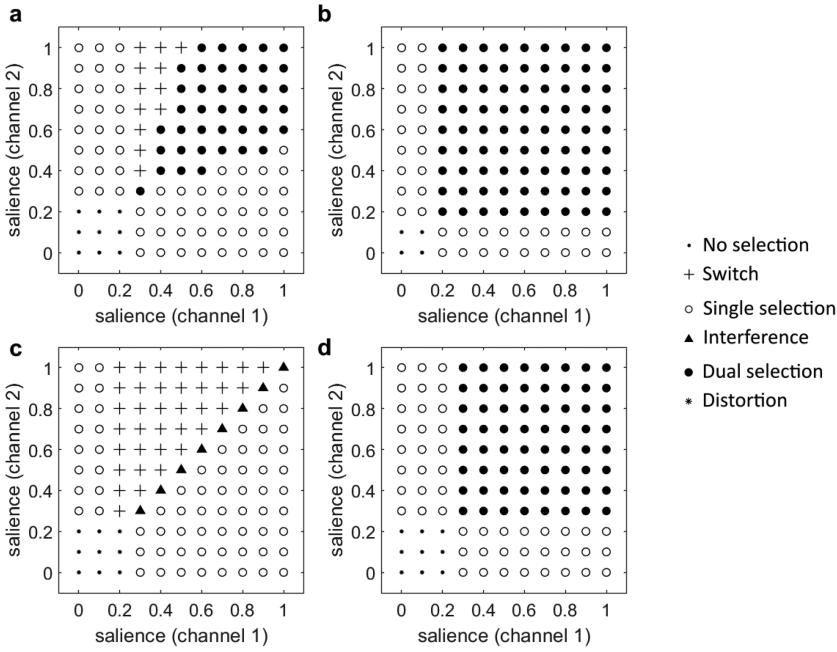

Figure 4: Templates for hard and soft selection regimes. (a, b) Simulation outcomes with two competing channels and the original model, with dopamine level $\lambda = 0.294$ (a) and $\lambda = 0.818$ (b). (c, d) Optimal outcomes for hard (c) and soft (d) action selection regimes as defined in the text.

6. *Distortion.* A single channel may be selected or switching may occur, but the losing channel is not clearly deselected:

$$Y_1^{t=2} \leq \theta \quad \text{and} \quad Y_1^{t=3} \leq \theta \quad \text{and} \quad Y_2^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} \leq \theta_d \tag{3.7}$$

$$\text{or} \quad Y_1^{t=2} > \theta \quad \text{and} \quad Y_1^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} \leq \theta \quad \text{and} \quad Y_1^{t=3} \leq \theta_d \tag{3.8}$$

$$\text{or} \quad Y_1^{t=2} \leq \theta \quad \text{and} \quad Y_1^{t=3} > \theta \quad \text{and} \quad Y_2^{t=3} \leq \theta \quad \text{and} \quad Y_1^{t=3} \leq \theta_d. \tag{3.9}$$

To discover the properties of selection more generally for a given set of model parameters, we repeat this experiment for all combinations of c_1 and c_2 between 0 and 1 with a step of 0.1. The outcome of this set of 121 simulations is reported on a 2D grid, with axes for c_1 and c_2 and a symbol at each grid point denoting which of the six possible outcomes occurred (as defined above; see Figure 4). Panels a and b show such results for typical low and high dopamine levels in the original model.

3.2 Dopamine-Driven Switch between Hard and Soft Selection. To quantify selection performance, it is necessary to reduce the grid of 121 experimental outcomes obtained in the previous section to a simpler set of numeric quantities and in turn characterize how these quantities change with dopamine modulation. To that end, we follow here the method given in Gurney et al. (2004) and Suryanarayana et al. (2019), which depends on the notions of hard and soft selection introduced in section 1.3. In the original model with low dopamine drive (see Figure 4a), most of the experiments produced either no selection, single-channel selection, or a channel selection switch; only a handful produced dual-channel selection when competition is strong at very high salience. This general pattern would therefore match hard selection more than soft selection due to the lower occurrence of dual selection. Thus, to characterize the archetypal variant of hard selection, we take this outcome to a limiting case shown in Figure 4c. Here there are no instances of dual selection, and the number of no-selection case is minimized (there have to be some, as disinhibition beyond threshold cannot occur for very small inputs). We refer to the outcome in Figure 4c as the hard selection template. Conversely, with high dopamine (see Figure 4b), most of experiments produced dual-channel selection, which is typical of a soft-selection regime. This is idealized in the soft selection template shown in Figure 4d. Note that these templates are optimized by hand rather than algorithmically and are identical to those used in previous iterations of the model to enable meaningful comparisons (Gurney et al., 2004; Suryanarayana et al., 2019).

We can now use these templates to characterize the outcome of a complete set of 121 experiments by its degree of similarity with each of the templates. Thus, for each experimentally obtained grid, let N_h be the number of grid points with the same outcome as the hard selection template and put $P_h = 100 \times N_h/121$. In a similar way, we define P_s as the degree of similarity with the soft selection template.

The values P_h, P_s will depend on the level of dopamine (as exemplified in Figure 4). The GPR model assumes this transition from hard to soft selection to be a feature of dopamine control of basal ganglia function (Blanco & Sloutsky, 2021; Costa, 2007; Costa et al., 2006; Wickens, Horvitz, Costa, & Killcross, 2007; Gizer, Ficks, & Waldman, 2009; Bogacz, 2020). If the basal ganglia supports such a mechanism, then values of the parameters w_{D1}, w_{D2}, which will optimize this transitional control, will reflect the relative biological dopamine sensitivity these parameters aim to capture. To proceed, we now require a quantitative measure of the transitional control from hard to soft selection.

3.3 Quantifying Dopamine-Dependent Selection Control. We first obtain the similarity measures P_h, P_s for each level of dopamine. Dopamine is represented as a scalar λ in the model, but a quantity which is perhaps easier

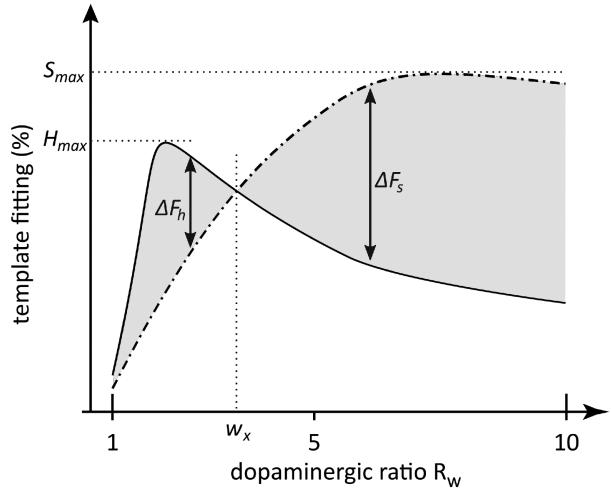

Figure 5: Cartoon representing the variables used to define the template fitting functions' main features. Hard and soft selection templates P_h and P_s are represented in solid and dashed lines, respectively. Figure modified from (Gurney et al., 2004).

to relate to cortico-striatal modulation is the ratio, R_w, of signal amplification/attenuation in the original model

$$R_w = \frac{1+\lambda}{1-\lambda}. \tag{3.10}$$

Note that R_w is bounded below by 1 since λ is bounded below by 0, but it has no theoretical upper bound as λ approaches 1 (see section A.4 in appendix A for more details). Then, for a given dopaminergic ratio, R_w, we run a grid set of experiments and obtain $P_h(R_w)$ and $P_s(R_w)$. We can then plot these functions as shown in Figure 5.

In an ideal system that allows control of the hard/soft selection via the parameter R_w, we might expect the following characteristics. First, there will be well-defined regimes for hard and soft selection, separated by some crossover point at $R_w = w_x$. For $1 \leq R_w < w_x$, hard selection will dominate, and so we have $P_h(R_w) > P_s(R_w)$; for $w_x < R_w$, soft selection prevails, and so we have $P_h(R_w) < P_s(R_w)$ (see Figure 5). Notice that the observed behavior of the original model (see Figure 6a) complies with this description. The degree to which the two regimes are well defined and contrasted may be captured in any number of ways, but we adopt a feature-based approach as shown in the figure. The features are the maximum value of P_h in the hard selection regime, $H_{\max}$; the corresponding maximum for soft selection, $S_{\max}$; the mean value ΔF_h of the difference $P_h - P_s$ for hard selection; and the mean

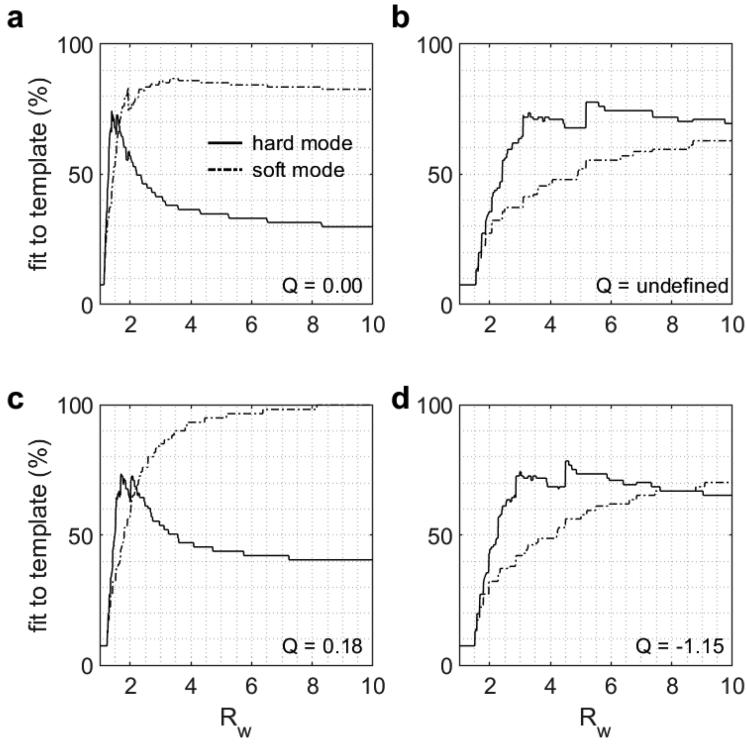

Figure 6: Sample of representative template fitting functions for different versions of the model. The weights of each version are indicated in Table 1. (a) The original model ($Q = 0$). (b) A failed model (Q undefined). (c) The best model obtained across all experiments ($Q = 0.18$). (d) A model that evaluated worse than the original model ($Q = -1.15$). The solid and dashed lines represent P_h and P_s, respectively.

value ΔF_s of $P_s - P_h$ up to a practical maximum value of $R_w = 10$. While this maximum value is somewhat arbitrary, we expect that physiological constraints will also limit any biological correlate of this ratio.

Ideally, we would expect all these features to have as large a value as possible. In addition, we might expect the transition point w_x to be significantly larger than 1; otherwise, there is little room for adopting the hard selection regime under noisy control of the parameter R_w. We therefore consider w_x as a member of the feature set describing selection.

To facilitate comparison with the original model, we used the ratio of the feature values for any new model with respect to those for the original model. Thus, if f is one of the five features $\{w_x, \Delta F_h, \Delta F_s, H_{\max}, S_{\max}\}$ we define $r = f/f^G$ where f^G is the original model value. We also log-transform,

defining $\tilde{f} = \log(r)$—for example, $\Delta\tilde{F}_h = \log(\Delta F_h / \Delta F_h^G)$, where ΔF_h^G is the value for the GPR model. Then $\tilde{f} > 0$ implies a feature value larger than that of the GPR model, which, by dint of the way the features were defined, means a better switch between hard and soft selection performance.

The features can then be used to define a metric, Q, which measures the quality of the transitional control between the two regimes. Thus, let $\tilde{f}_i$ be the (log-transformed) ith feature ($1 \leq i \leq 5$ since there are five features). Then assuming all features are well defined and that no $r_i = 0$, we can define $Q = \sum_i \tilde{f}_i$. From all the definitions above, the GPR model therefore yields $Q = 0$, and positive/negative values of Q imply a better/worse performance than the baseline of GPR. If, however, any model feature is not defined or zero, then Q is also not defined. This indicates a failure of the model to display sensible hard or soft selection behavior. More details about the rationale behind these features and their definition are in Gurney et al. (2004).

4 Results

A series of experiments was done by sampling w_{D1}, w_{D2} on a grid over their respective ranges of values. Note that we must have $w_{D2} \cdot \lambda \leq 1$ since values larger than this can allow for and overall change the sign of the corticostriatal input in equation 2.4, which is biologically implausible. Since we ran our simulations with $R_w \leq 10$, rearranging equation 3.10 gives us $\lambda \leq 9/11$, and so we must have $w_{D2} \leq 11/9$. Conversely, w_{D1} is not bounded as such, but performance peaked at values well under $w_{D1} = 10$, which therefore was a practical upper bound for our simulations. Thus, we varied w_{D1}, w_{D2} so that $0 \leq w_{D1} \leq 10$ and $0 \leq w_{D2} \leq 11/9$. We ran the experiments with 401 and 50 sampling points for parameters w_{D1} and w_{D2}, respectively, to ensure at least a sampling point every 0.0250 step. This resulted in 20,050 complete evaluations of the model, each yielding a set of features and Q value indicating its performance relative to the original model.

4.1 The Sensitivity Weightings Significantly Alter Switching Performance. Some representative template fitting functions, obtained from our grid sampling of w_{D1}, w_{D2}, are given in Figure 6 and Table 1, with results for the original model shown in panel a. Panel b shows an instance where there is no hard-soft transition within the domain explored and so the Q metric is undefined. In contrast, panels c and d show two cases of successful models with $Q > 0$ and $Q < 0$, indicating better and worse performance than the original model, respectively (in fact, the result in panel c was for the best model obtained).

The better performance of the model in panel c is due to greater fit to the optimal soft template displayed in Figure 4d ($S_{\max}$ in Table 1), while the fit to hard selection $H_{\max}$ is similar to the original model (although marginally

Table 1: Feature Set, Merit Q, and Parameter Values of Each Model from Figure 6.

Model Fig. 6 Panel	Original Panel a	Failed Panel b	Best Panel c	Worst Panel d
w_{D1}	1.00	0.57	0.28	0.55
w_{D2}	1.00	0.00	1.07	0.07
$H_{\max}$	74.38	77.69	73.55	78.51
$S_{\max}$	86.78	62.81	100.00	70.25
ΔF_h	14.29	16.36	13.75	15.02
ΔF_s	47.86	undefined	49.83	4.18
w_x	1.66	undefined	2.22	1.51
Q	0.00	undefined	0.18	−1.15

lower). The difference between hard and soft selection (ΔF_h and ΔF_s) is similar to that of the original model. Finally, w_x is greater than the original model (from 1.66 to 2.22), illustrating that the hard selection regime is more expressed along the dopaminergic axis.

4.2 Model Switching Performance Improves for Different Rather Than Equal w_{D1}, w_{D2} Weights. Next, we assessed the dependence of each feature on w_{D1}, w_{D2} (see Figure 7). Note that in this section, we discuss the features as a function of the original model ($\tilde{w_x}$ instead of w_x) to allow for better comparison, as mentioned in section 3.3. Thus, a feature value above/below 0 indicates better/worse performance than the original model for the corresponding w_{D1}, w_{D2} pair.

A first observation is that $w_{D1} < 0.2$ and $w_{D2} < 0.2$ lead to a failed model (signaled by white in Figure 7), and that this is not due to a specific feature failing but rather to all features failing similarly (see Figures 7a, 7b, 7e, and 7f, bottom left corners), with the exception of $\tilde{H}_{\max}$ and $\tilde{S}_{\max}$ (see Figures 7c and 7d). Some features even fail for $w_{D2} < 0.5$ values, such as $\Delta\tilde{F}_s$ and $\tilde{w}_x$ (see Figures 7b and 7e). Overall, small w_{D1} and w_{D2} values effectively shrink the effect of dopamine on the cortico-striatal input, eventually forcing the model to behave as it does when λ is always close to 0 (see equations 2.3 and 2.4). This hinders the model's ability to switch to soft selection at high λ values, leading to models behaving similarly to what we observe in Figure 6b, although to a more exteme case.

The opposite occurs as w_{D1} and w_{D2} take very large values. The λ modulation on cortico-striatal input is magnified, forcing the model to quickly transition to a soft selection mode. This can be seen by looking at $\Delta\tilde{F}_h$ in Figure 7a, which quickly reaches a (positive) peak value at low w_{D1} values and then drops and stabilizes around 0 as w_{D1} keeps increasing. Similarly, we see that $\tilde{w_x}$ initially takes high, positive values and then drops to

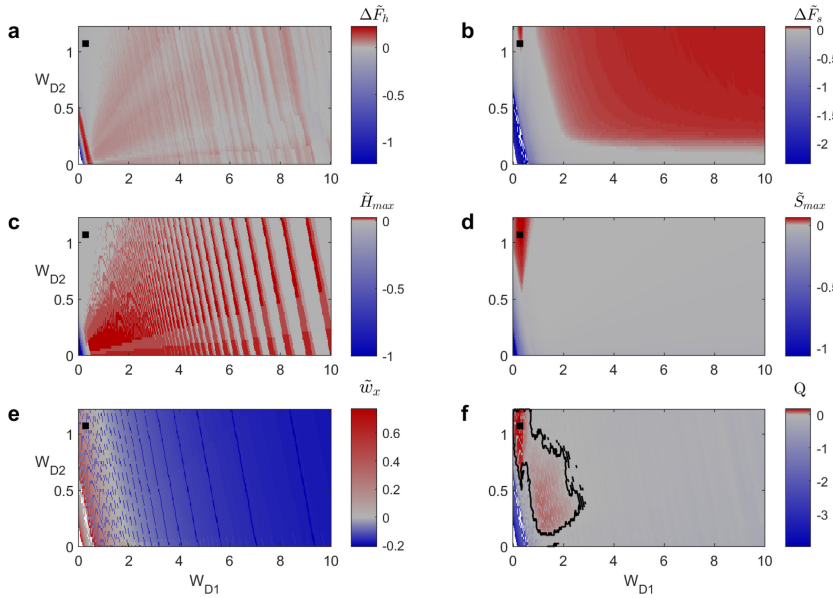

Figure 7: Variation of (a) $\Delta\tilde{F}_h$, (b) $\Delta\tilde{F}_s$, (c) $\tilde{H}_{\max}$, (d) $\tilde{S}_{\max}$, (e) $\tilde{w}_x$, and (f) Q as a function of w_{D1} and w_{D2} for the different versions of the augmented model. Blue and red represent low and high values, respectively. Empty (white) values mean that the variable is undefined. The black square mark on each panel indicates the location of the best Q value. The solid black line in panel f surrounds the region for which the augmented model is better than the control model ($Q > 0$). Note that the parameter value to color space mapping is nonlinear and centered on 0.

negative values as w_{D1} increases as well, indicating that the transition occurs for lower R_w values than in the original model (see Figure 7e).

$\Delta\tilde{F}_s$ shows the opposite pattern: the larger w_{D1} and w_{D2}, the better this feature performs. However, the improvement from the original model is very marginal, meaning that this feature did not drive up the Q value in that region (see Figure 7f) and was therefore not particularly relevant for improving the model's switching capabilities. Finally, the remaining two features $\tilde{H}_{\max}$ and $\tilde{S}_{\max}$ both displayed a sweet spot: for w_{D1} around 1.5 and w_{D2} around 0.2 ($\tilde{H}_{\max}$) and for w_{D1} around 0.3 and w_{D2} around 1.0 ($\tilde{S}_{\max}$).

Thus, the overall Q metric has two reasonably distinct regions where $Q > 0$, corresponding to the improvement in $\tilde{H}_{\max}$ and $\tilde{S}_{\max}$. However, the improvement in $\tilde{S}_{\max}$ is quantitatively greater than that of $\tilde{H}_{\max}$, which is why we clearly observe our best model in the region matching better $\tilde{S}_{\max}$ performance, with $Q = 0.18$ (see Table 1 and Figure 6c).

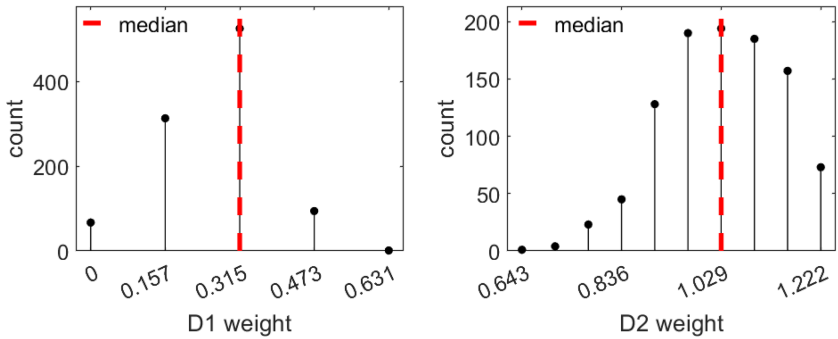

Figure 8: Distribution of the best model's w_{D1} (left) and w_{D2} (right) values for each of a thousand altered-connectivity models. The red vertical dashed lines indicate the median of the distribution.

4.3 The Optimal w_{D1}, w_{D2} Weight Values Are Robust to Random Changes in the Model Connectivity. The presence of two regions with improved model performance suggests that slightly different models may lean toward one region or another. In other words, is the GPR model's observed preference for low w_{D1} and high w_{D2} values robust? To answer this question, we repeated the experiment, giving rise to Figure 7a thousand times while randomly varying the value of each connectivity weight in the model within ±10% of its original value. We excluded from these variations w_{D1}, w_{D2}, and the cortico-striatal and cortico-STN input weight w_s^{cortex} and w_{stn}^{cortex}, which were fixed at their default value to ensure the input to the model was identical across simulations (see appendix A and Table 2 for detail). All remaining 24 weights were redrawn independently from each other for each simulation run using a uniform distribution. From this, we can see the distributions of the best w_{D1}, w_{D2} weight values are centered on low w_{D1} and high w_{D2} similar to the unaltered model (Figure 8), indicating that the result we observe is robust to marginal random changes in the model's connectivity.

4.4 The D1/D2 Weight Ratio Does Not Define Model Performance. We next assessed how our models performed as a function of the weight ratio w_{D1}/w_{D2} rather than the absolute value of each weight independently. We do observe that the best merit values occur with a ratio lower than 0.6 for the augmented model compared to the original model ($Q > 0$; see Figure 9a). This was also true in the set of altered-connectivity models, where the best model for each run always displayed a ratio lower than 0.6 (see Figure 9b). However, this ratio alone was not sufficient for a model to perform better, as many models with a ratio under 0.6 still performed poorly, and even performed worse than many models with a larger ratio (see Figure 9a).

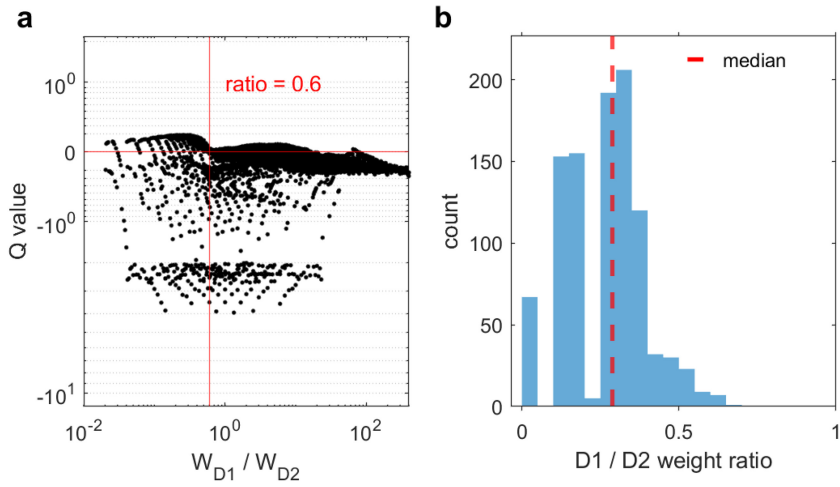

Figure 9: (a) Model merit Q as a function of w_{D1}/w_{D2} ratio. Note that the axes are in a logarithm base 10 scale. The red horizontal line indicates $Q = 0$. The red vertical line indicates a ratio of 0.6 (see text). (b) Distribution of the w_{D1}/w_{D2} ratio for each of the best altered-connectivity models. The red dashed line indicates the median of the distribution.

Therefore, it appears that a ratio lower than 0.6 is required for better model performance, but that one such ratio does not guarantee a well-performing model in itself. Another striking result is that while the most performant models all have a ratio lower than 0.6, models with different ratios within that region can perform similarly well compared to each other (e.g., a ratio of 0.4 versus a ratio of 0.5), indicating that there is no optimum ratio value.

5 Discussion

From the behavior of the model as we varied w_{D1}, w_{D2}, we observe that both $w_{D1} \approx 0.2$ with $w_{D2} \approx 1.0$ and $w_{D1} \approx 1.5$ with $w_{D2} \approx 0.2$ lead to an improved switch between action selection regimes compared to the original model. However, the former combination resulted in greater improvement compared to the latter, yielding the largest merit value in this study. To assess the robustness of this result, we performed a thousand random alterations of the model connectivity and observed that the best w_{D1}, w_{D2} weight value pairs for each run remained in a similar range to that of the unaltered model. This indicates that differential, rather than equal, sensitivity to dopamine is a benefit to this model of the BG to perform a switch between hard and soft selection regimes. Finally, we observe that this improved performance is only weakly related to the relative w_{D1}/w_{D2} ratio, as it also depended on the absolute value these weights take. In particular, the models that performed

best (with or without altered connectivity) always showed a w_{D1}/w_{D2} ratio lower than about 0.6, but this ratio constraint was not a sufficient condition for improved performance.

How do the optimal weights w_{D1} and w_{D2} observed here compare to physiological observations? Two approaches can be used in order to experimentally determinate the functional connectivity between neurons. First, one can record the electrophysiological response of the postsynaptic neuron following artificial stimulation of the presynaptic neuron (Chuhma, Tanaka, Hen, & Rayport, 2011). However, such an approach is technically challenging, and the dopamine effect on striatal neurons depends on the overall state of the network (Surmeier et al., 2007). For instance, striatal neurons may be up-state or down-state, or the drive of cortical input may vary. Therefore, there is to our knowledge no study that directly assesses the functional connectivity of dopaminergic neurons in the striatum.

The second approach consists in quantitatively determining the functional contribution of each of the steps involved in synaptic transmission that contribute to the complete response. There are three such contributions: the proportion of D1 and D2 receptors expressed, their binding affinity, and amplification of intracellular signal transduction following receptor activation. Generally, quantifying D1 and D2 sensitivity has proved a challenging question, and the answer likely differs depending on the timescale and temporal profile of the dopaminergic signal and state of the postsynaptic neuron (Yapo et al., 2017; Richfield et al., 1989; Cumming, 2011; Skinbjerg et al., 2012; Kenakin, 2013; Marcott et al., 2014). Here we briefly review, in the context of the current model, some of the data pertaining to each of the three possible mechanisms underlying dopamine sensitivity.

Perhaps the least challenging contribution to estimate is that of the proportion of D1 and D2 receptors expressed. D1 receptors are estimated to be three to five times more common than D2 receptors in the striatum (Liu, Goel, & Kaeser, 2021), thereby enhancing relative sensitivity in the D1 pathway. Regarding binding affinity, previous work evaluated 20% and 90% the proportion of D1 and D2 receptors in a high-affinity state against a selective agonist (Richfield et al., 1989). However, these estimates can suffer from methodological limitations (see Cumming, 2011; Skinbjerg et al., 2012 for reviews). Therefore, while the reported results on affinity states appear consistent with our result, their full interpretation remains unclear.

Concerning intracellular signal transduction, the D1 and D2 pathways respectively lead to a phosphorylation and dephosphorylation of DARPP-32, a molecule that is sometimes considered as the final effector of D1 and D2 intracellular pathways (Svenningsson et al., 2004; Lindskog, Kim, Wikström, Blackwell, & Kotaleski, 2006; Surmeier et al., 2007; Nishi, Snyder, & Greengard, 1997; Nishi et al., 2000). The opposite effect of D1 and D2 signaling pathways on DARPP-32 is in line with the functionally opposite effect of D1 and D2. Critically, an increase of six-fold in DARPP-32 phosphorylation can be observed for D1 activation, while a two-fold decrease is observed following D2 activation (Nishi et al., 1997, 2000). However,

measurements of intracellular signaling have sometimes provided conflicting accounts, adding to the uncertainty on the D1 and D2 functional weights (Yapo et al., 2017; Marcott et al., 2014; Watts et al., 1998; Watts & Neve, 1996; Kim et al., 2004; Dumartin et al., 2000).

Taken together, the larger number of D1 receptors expressed and the stronger intracellular amplification they produce could result in greater D1 sensitivity, while the larger proportion of high affinity-state D2 receptors suggests the opposite. Our model prediction, based on the action selection hypothesis extended to include dopaminergic control of selection regime, is that there is an overall increased sensitivity to dopamine in the control (D2) pathway compared to that in the selection (D1) pathway. Given the discussion above, this is not inconsistent with the current state of physiological knowledge. Testing our prediction requires an integrated interpretation of this knowledge in terms of functional sensitivity. This, in turn, will require a direct experimental means of assessing functional dopamine sensitivity at cortico-striatal synapses. In lieu of this, our prediction could be tested against those from a detailed, low-level model of the relevant synaptic mechanisms.

Our model, like all others, presents several limitations. One assumption in our model is that selection occurs via a BG-wide, off-center on-surround network (see Figure 2). However, we know that the striatum (as well as other BG nuclei) contains a complex GABAergic microcircuitry (Burke, Rotstein, & Alvarez, 2017) that may contribute to a selection function, a possibility that several computational models have explored (Bariselli et al., 2019; Tomkins, Vasilaki, Beste, Gurney, & Humphries, 2014). This additional striatal function does not preclude our mechanisms but should work synergistically to enhance it. It remains the subject of future work to see precisely how this might take place. Second, we have focused on the role of dopamine that enables switching between hard and soft selection regimes. However, a large body of evidence shows that dopamine performs a plethora of other functions in the BG (Lerner, Holloway, & Seiler, 2021). For instance, it has been hypothesized to propagate information about belief states (Gershman & Uchida, 2019), statistical properties of reward (distributional encoding; Lowet, Zheng, Matias, Drugowitsch, & Uchida, 2010), novelty (Redgrave & Gurney, 2006), or to contribute to timing of movement onset (Howe & Dombeck, 2016; da Silva, Tecuapetla, Paixão, & Costa, 2018). It remains to be shown that the weighting factors we predict here are consistent with optimal performance of these other functions of dopamine. Dopaminergic modulation is also considered constant here, while in vivo dopamine release shows high temporal variability (Liu et al., 2021), with the presence of phasic and tonic release (Wang, Toyoshima, Kunimatsu, Yamada, & Matsumoto, 2021) or dopamine ramps (Berke, 2018) that are not included in the GPR model. The spatial organization of individual channels might also have an impact on the role of dopaminergic diffusion and therefore channel-dependent changes in dopamine levels (Liu et al., 2021; Bariselli et al., 2019). Furthermore, the architecture of the BG in the GPR model is a simplification

of biological reality. For instance, recent work clearly demonstrates a complex local striatal circuitry (Burke et al., 2017), with GABA-ergic interneurons, collateral transmission between medium spiny neurons, or cholinergic modulation that may work in synergy with dopaminergic modulation (Howe et al., 2019; Burke et al., 2017). Striatal neurons also express up- and down-states that alter their physiological response (Arbuthnott & Wickens, 2007). While the exact purpose of this local circuitry is not yet fully understood, it cannot be ruled out that it may alter the effect dopaminergic modulation has on BG function.

The "original" GPR model we used here corresponds to an updated version of the first proposed GPR model (Gurney et al., 2001a, 2001b) that includes more recently characterized anatomical features of the GPe (Suryanarayana et al., 2019). Interestingly, performing the same set of simulations with the GPR model proposed in Gurney et al. (2001a, 2001b) yields a result opposite to that of the GPR model based on Suryanarayana et al. (2019). Specifically, we observe that $w_{D1} = 1.95$ and $w_{D2} = 0.82$ are functionally optimal. The full set of simulation results is available in a previous version of this study, available online (Codol, Gribble, & Gurney, 2020). Considering that the Suryanarayana et al. (2019) model contains a more exhaustive architecture matching physiological reality, we give greater weight to the findings it yields than to those of the previous version of the GPR model. However, it is critical to assess whether this change in simulation result is indeed due to the model's updated architecture and not just because the GPR model is inherently oversensitive to small changes. This motivated the simulations and analyses presented in section 4.3, where random perturbations to each connectivity weights were performed to assess the overall robustness of the GPR model. The main results remained identical to that of the unperturbed model, suggesting that the changes in results we see with the Suryanarayana et al. (2019) model compared to the Gurney et al. (2001a, 2001b) model are likely due to the inclusion of the GPe intrinsic connectivity rather than an inherent lack of robustness in the GPR model.

Dopamine sensitivity at the level of striatum is physiologically grounded in many complex and interacting mechanisms whose overall effect is intractable just now. In this study, we used a high-level computational model to predict the optimal sensitivity of the D1 and D2 receptors for the purpose of switching between hard and soft selection regimes. We have shown that selection should benefit from differential rather than identical sensitivity therein and that this benefit would be with D2-favorable emphasis in such sensitivity.

Appendix A: Model Formulation

A.1 General Structure. Each nucleus was connected to the other nuclei according to the architecture in Figure 1a. Any projection from a source nucleus to a target nucleus was associated with a fixed connectivity weight

w_{target}^{source}, which determined the contribution of the source nucleus's output y_i^{source} into the target nucleus input u_i for a given channel i. The input for the target nucleus u_i was then used to define an activity function a_i over time using an ordinary differential equation:

$$\frac{da_i}{dt} = -k(a_i - u_i), \tag{A.1}$$

where $k = 25$ is a decay constant and t is time. Since at equilibrium $da_i/dt = 0$, the equation converges to $a_i = u_i$ for a fixed input value over time. The activity a_i is then passed through a piece-wise output function to avoid runaway activity, giving the nucleus's output y_i:

$$y_i = \begin{cases} 0 & \text{if } a_i < \epsilon \\ (a_i - \epsilon) & \text{if } \epsilon \leq a_i \leq 1 + \epsilon \\ 1 & \text{if } a_i > 1 + \epsilon \end{cases} \tag{A.2}$$

where ϵ is a constant threshold whose value depends on the nucleus. The parameter values employed for each connectivity weight and constant threshold ϵ are listed in Table 2. A rationale of most parameters of the original model can be found in Gurney et al. (2001a, 2001b), and Suryanarayana et al. (2019).

A.2 Striatum. Although the mathematical implementation of the striatum has been detailed, it is repeated here for convenience. In the original model (Suryanarayana et al., 2019), the striatum D1 and striatum D2 receive the same salience input c_i from the cortex, which is then modulated by the dopaminergic input λ and weighted by a shared striatal weight w_s^{cortex}. All three GPe neural populations also send inhibitory projections, with $y_i^{ark}, y_i^{out}, y_i^{inn}$ the projection of the GPe's arkypallidal, outer, and inner neural population to the striatum, respectively. Their input weights are $w_{D1}^{ark}, w_{D1}^{out}, w_{D1}^{inn}$, for the striatum D1 and $w_{D2}^{ark}, w_{D2}^{out}, w_{D2}^{inn}$ for the striatum D2:

$$u_i^{D1} = w_s^{cortex}(1 + \lambda)\, c_i + w_{D1}^{out}\, y_i^{out} + w_{D1}^{inn}\, y_i^{inn} + w_{D1}^{ark} \sum_i y_i^{ark}, \tag{A.3}$$

$$u_i^{D2} = w_s^{cortex}(1 - \lambda)\, c_i + w_{D2}^{out}\, y_i^{out} + w_{D2}^{inn}\, y_i^{inn} + w_{D2}^{ark} \sum_i y_i^{ark}. \tag{A.4}$$

The output of the striatum D1 and striatum D2 is denoted y_i^{D1} and y_i^{D2}, respectively, after applying equations A.1 and A.2. They share the same activation threshold ϵ^s for A.2. Note that for $\lambda = 0$ (no dopamine), the dopaminergic modulation $(1 \pm \lambda)$ reduces to 1. Also note the sum over all channels for the y_i^{ark} projection, indicating a diffuse projection from the arkypallidal

Table 2: Parameter Values for the GPR Model.

Parameter	Value	Description
w_s^{cortex}	1	cortex → striatum weight (both D1 and D2)
w_{stn}^{cortex}	1	cortex → STN weight
w_{D1}^{ark}	−0.25	arkypallidal GPe → striatum D1 weight
w_{D1}^{out}	0.5	outer GPe → striatum D1 weight
w_{D1}^{inn}	0.25	inner GPe → striatum D1 weight
w_{D2}^{ark}	−0.25	arkypallidal GPe → striatum D2 weight
w_{D2}^{out}	0.5	outer GPe → striatum D2 weight
w_{D2}^{inn}	0.25	inner GPe → striatum D2 weight
w_{stn}^{out}	−0.8	outer GPe → STN weight
w_{stn}^{inn}	−0.8	inner GPe → STN weight
w_{ark}^{D2}	−.9	striatum D2 → arkypallidal GPe weight
w_{ark}^{stn}	0.8	STN → arkypallidal GPe weight
w_{ark}^{ark}	−0.75	arkypallidal GPe → arkypallidal GPe weight
w_{ark}^{out}	−0.75	outer GPe → arkypallidal GPe weight
w_{ark}^{inn}	−0.75	inner GPe → arkypallidal GPe weight
w_{out}^{D2}	−.9	striatum D2 → outer GPe weight
w_{out}^{stn}	0.8	STN → outer GPe weight
w_{out}^{out}	−0.75	outer GPe → outer GPe weight
w_{inn}^{D2}	−.9	striatum D2 → inner GPe weight
w_{inn}^{stn}	0.8	STN → inner GPe weight
w_{inn}^{out}	−0.3	outer GPe → inner GPe weight
w_{inn}^{inn}	−0.75	inner GPe → inner GPe weight
w_{GPi}^{D1}	−1	striatum D1 → GPi/SNr weight
w_{GPi}^{stn}	0.9	STN → GPi/SNr weight
w_{GPi}^{out}	−1	outer GPe → GPi/SNr weight
w_{GPi}^{inn}	−0.2	inner GPe → GPi/SNr weight
θ	0	Selection threshold
θ_d	0.01032	Distortion threshold
ϵ^s	0.2	Output threshold for striatum D1 and D2
ϵ^{stn}	−0.25	Output threshold for STN
ϵ^{GPe}	−0.2	Output threshold for all GPe subpopulations
ϵ^{GPi}	−0.2	Output threshold for GPi/SNr

neurons (see Figure 1a). Finally, equations A.3 and A.4 are altered in the augmented model, as indicated in equations 2.3 and 2.4.

A.3 Other Nuclei. Equations A.5 to A.9 define the input to the STN, arkypallidal GPe, outer GPe, inner GPe, and GPi/SNr, in that order:

$$u_i^{stn} = w_{stn}^{cortex}\, c_i + w_{stn}^{out}\, y_i^{out} + w_{stn}^{inn}\, y_i^{inn}, \tag{A.5}$$

$$u_i^{ark} = w_{ark}^{D2}\, y_i^{D2} + w_{ark}^{out}\, y_i^{out} + w_{ark}^{inn}\, y_i^{inn} + w_{ark}^{ark}\, y_i^{ark} + w_{ark}^{stn} \sum_i y_i^{stn}, \tag{A.6}$$

$$u_i^{out} = w_{out}^{D2}\, y_i^{D2} + w_{out}^{out} y_i^{out} + w_{out}^{stn} \sum_i y_i^{stn}, \tag{A.7}$$

$$u_i^{inn} = w_{inn}^{D2}\, y_i^{D2} + w_{inn}^{out}\, y_i^{out} + w_{inn}^{inn}\, y_i^{inn} + w_{inn}^{stn} \sum_i y_i^{stn}, \tag{A.8}$$

$$u_i^{GPi} = w_{GPi}^{D1}\, y_i^{D1} + w_{GPi}^{out}\, y_i^{out} + w_{GPi}^{inn}\, y_i^{inn} + w_{GPi}^{stn} \sum_i y_i^{stn}. \tag{A.9}$$

Their outputs are denoted y_i^{stn}, y_i^{ark}, y_i^{out}, y_i^{inn}, and Y_i. Their output function thresholds are denoted ϵ_{stn} for the STN, ϵ_{GPe} for all three GPe subpopulations and ϵ_{GPi} for the GPi/SNr. Importantly, Y_i is also the final output of the model for channel i.

A.4 Parameter Values. All the models, including the original one, contain $N = 6$ channels. However, in the simulations, only two of them receive salience input, that is, we always have $c_{3-6} = 0$. This allows keeping the set of potential outcomes simple while still exposing the effect of channel-to-channel competition. The role of the remaining four channels is to model the tonic inhibition from quiescent channels via the diffuse projections of the STN. Previous work on the GRP model also contains four quiescent channels for the same reason (Gurney et al., 2004; Suryanarayana et al., 2019).

The parameter R_w ranges from 1 to 10 because when $\lambda = 0$, we have $R_w = 1$, and spanning this parameter to 10 allows varying dopamine levels enough for the model to fully and efficiently express both hard and soft selection regimes (Gurney et al., 2004). Moreover, since the dopaminergic modulation of the striatum D2 is formalized as $(1 - \lambda)$, we want to avoid $\lambda > 1$ because it would shut down the control pathway. Since we have $\lambda = (R_w - 1)/(R_w + 1)$, $R_w = 10$ gives us $\lambda = 9/11$ at maximum. Critically, this implies that one of the new parameters added to the augmented model, w_{D2}, cannot evaluate at 11/9 or more, because when dopamine modulation reaches its maximum value during simulations ($R_w = 10$), the control pathway would be shut down, or even inverted, which is biologically implausible.

References

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*, 357–381. 10.1146/annurev.ne.09.030186.002041, PubMed: 3085570

Arbuthnott, G. W., & Wickens, J. (2007). Space, time and dopamine. *Trends in Neurosciences*, *30*(2), 62–69. 10.1016/j.tins.2006.12.003, PubMed: 17173981

Bariselli, S., Fobbs, W. C., Creed, M. C., & Kravitz, A. V. (2019). A competitive model for striatal action selection. *Brain Research*, *1713*, 70–79. 10.1016/j.brainres.2018.10.009, PubMed: 30300636

Berke, J. D. (2018). What does dopamine mean? *Nature Neuroscience*, *21*(6), 787–793. 10.1038/s41593-018-0152-y, PubMed: 29760524

Blanco, N. J., & Sloutsky, V. M. (2021). Systematic exploration and uncertainty dominate young children's choices. *Developmental Science*, *24*(2), e13026. 10.1111/desc.13026, PubMed: 32767496

Bogacz, R. (2020). Dopamine role in learning and action inference. *eLife*, *9*, e53262. 10.7554/eLife.53262

Burke, D. A., Rotstein, H. G., & Alvarez, V. A. (2017). Striatal local circuitry: A new framework for lateral inhibition. *Neuron*, *96*(2), 267–284. 10.1016/j.neuron.2017.09.019, PubMed: 29024654

Chevalier, G., & Deniau, J. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends in Neurosciences*, *13*(7), 277–280. 10.1016/0166-2236(90)90109-N, PubMed: 1695403

Chuhma, N., Tanaka, K. F., Hen, R., & Rayport, S. (2011). Functional connectome of the striatal medium spiny neuron. *Journal of Neuroscience*, *31*(4), 1183–1192. 10.1523/JNEUROSCI.3833-10.2011, PubMed: 21273403

Codol, O., Gribble, P. L., & Gurney, K. N. (2020). *Differential dopamine receptor-dependent sensitivity improves action selection in the basal ganglia*. bioXiv. 10.1101/2020.11.12.380451

Coizet, V., Graham, J. H., Moss, J., Bolam, J. P., Savasta, M., McHaffie, . . . Overton, P. G. (2009). Short-latency visual input to the subthalamic nucleus is provided by the midbrain superior colliculus. *Journal of Neuroscience*, *29*(17), 5701–5709. 10.1523/JNEUROSCI.0247-09.2009, PubMed: 19403836

Costa, R. M. (2007). Plastic corticostriatal circuits for action learning. *Annals of the New York Academy of Sciences*, *1104*(1), 172–191. 10.1196/annals.1390.015, PubMed: 17435119

Costa, R. M., Lin, S.-C., Sotnikova, T. D., Cyr, M., Gainetdinov, R. R., Caron, M. G., & Nicolelis, M. A. (2006). Rapid alterations in corticostriatal ensemble coordination during acute dopamine-dependent motor dysfunction. *Neuron*, *52*(2), 359–369. 10.1016/j.neuron.2006.07.030, PubMed: 17046697

Cumming, P. (2011). Absolute abundances and affinity states of dopamine receptors in mammalian brain: A review. *Synapse*, *65*(9), 892–909. 10.1002/syn.20916, PubMed: 21308799

da Silva, J. A., Tecuapetla, F., Paixão, V., & Costa, R. M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature*, *554*(7691), 244–248. 10.1038/nature25457, PubMed: 29420469

Dumartin, B., Jaber, M., Gonon, F., Caron, M. G., Giros, B., & Bloch, B. (2000). Dopamine tone regulates D1 receptor trafficking and delivery in striatal neurons in dopamine transporter-deficient mice. In *Proceedings of the National Academy of Sciences*, *97*(4), 1879–1884. 10.1073/pnas.97.4.1879

Ericsson, J., Stephenson-Jones, M., Pérez-Fernández, J., Robertson, B., Silberberg, G., & Grillner, S. (2013). Dopamine differentially modulates the excitability of striatal neurons of the direct and indirect pathways in lamprey. *Journal of Neuroscience*, *33*(18), 8045–8054. 10.1523/JNEUROSCI.5881-12.2013, PubMed: 23637194

Feger, J., Bevan, M., & Crossman, A. R. (1994). The projections from the parafascicular thalamic nucleus to the subthalamic nucleus and the striatum arise from separate neuronal populations: A comparison with the corticostriatal and

corticosubthalamic efferents in a retrograde fluorescent double-labelling study. *Neuroscience*, *60*(1), 125–132. 10.1016/0306-4522(94)90208-9, PubMed: 8052406

Friend, D. M., & Kravitz, A. V. (2014). Working together: Basal ganglia pathways in action selection. *Trends in Neurosciences*, *37*(6), 301–303. 10.1016/j.tins.2014.04.004, PubMed: 24816402

Gershman, S. J., & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience*, *20*(1), 703-714. 10.1038/s41583-019-0220-7, PubMed: 31570826

Girard, B., Lienard, J., Gutierrez, C. E., Delord, B., & Doya, K. (2020). A biologically constrained spiking neural network model of the primate basal ganglia with overlapping pathways exhibits action selection. *European Journal of Neuroscience*, ejn.14869.

Gizer, I. R., Ficks, C., & Waldman, I. D. (2009). Candidate gene studies of ADHD: A meta-analytic review. *Human Genetics*, *126*(1), 51–90. 10.1007/s00439-009-0694-x, PubMed: 19506906

Graybiel, A. M., Ragsdale Jr., C. W., Yoneoka, E. S., & Elde, R. P. (1981). An immunohistochemical study of enkephalins and other neuropeptides in the striatum of the cat with evidence that the opiate peptides are arranged to form mosaic patterns in register with the striosomal compartments visible by acetylcholinesterase staining. *Neuroscience*, *6*, 377–397. 10.1016/0306-4522(81)90131-7, PubMed: 6164013

Groves, P. M., Linder, J. C., & Young, S. J. (1994). 5-hydroxydopamine-labeled dopaminergic axons: Three-dimensional reconstructions of axons, synapses and postsynaptic targets in rat neostriatum. *Neuroscience*, *58*(3), 593–604. 10.1016/0306-4522(94)90084-1, PubMed: 8170539

Gurney, K., Humphries, M., Wood, R., Prescott, T., & Redgrave, P. (2004). Testing computational hypotheses of brain systems function: A case study with the basal ganglia. *Network: Computation in Neural Systems*, *15*(4), 263–290. 10.1088/0954-898X_15_4_003

Gurney, K., Prescott, T. J., & Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics*, *84*(6), 401–410. 10.1007/PL00007984

Gurney, K., Prescott, T. J., & Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics*, *84*(6), 411–423. 10.1007/PL00007985

Howe, M. W., & Dombeck, D. A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*, *535*(7613), 505–510. 10.1038/nature18942, PubMed: 27398617

Howe, M., Ridouh, I., Allegra Mascaro, A. L., Larios, A., Azcorra, M., & Dombeck, D. A. (2019). Coordination of rapid cholinergic and dopaminergic signaling in striatum during spontaneous movement. *eLife*, *8*, e44903. 10.7554/eLife.44903

Humphries, M. D., Stewart, R. D., & Gurney, K. N. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *Journal of Neuroscience*, *26*(50), 12921–12942. 10.1523/JNEUROSCI.3486-06.2006, PubMed: 17167083

Kenakin, T. (2013). New concepts in pharmacological efficacy at 7TM receptors: IUPHAR Review 2. *British Journal of Pharmacology*, *168*(3), 554–575. 10.1111/j.1476-5381.2012.02223.x, PubMed: 22994528

Kim, O.-J., Gardner, B. R., Williams, D. B., Marinec, P. S., Cabrera, D. M., Peters, J. D., . . . Sibley, D. R. (2004). The role of phosphorylation in D1 dopamine receptor desensitization: Evidence for a novel mechanism of arrestin association. *Journal of Biological Chemistry*, *279*(9), 7999–8010. 10.1074/jbc.M308281200, PubMed: 14660631

Kwak, S., & Jung, M. W. (2019). Distinct roles of striatal direct and indirect pathways in value-based decision makin8g. *eLife*, e46050.

Lanciego, J. L., Gonzalo, N., Castle, M., Sanchez-Escobar, C., Aymerich, M. S., & Obeso, J. A. (2004). Thalamic innervation of striatal and subthalamic neurons projecting to the rat entopeduncular nucleus. *European Journal of Neuroscience*, *19*(5), 1267–1277. 10.1111/j.1460-9568.2004.03244.x, PubMed: 15016084

Lerner, T. N., Holloway, A. L., & Seiler, J. L. (2021). Dopamine, updated: Reward prediction error and beyond. *Current Opinion in Neurobiology*, *67*, 123–130. 10.1016/j.conb.2020.10.012, PubMed: 33197709

Lindskog, M., Kim, M., Wikström, M. A., Blackwell, K. T., & Kotaleski, J. H. (2006). Transient calcium and dopamine increase PKA activity and DARPP-32 phosphorylation. *PLOS Computational Biology*, *2*(9), e119. 10.1371/journal.pcbi.0020119, PubMed: 16965177

Liu, C., Goel, P., & Kaeser, P. S. (2021). Spatial and temporal scales of dopamine transmission. *Nature Reviews Neuroscience*, *22*(6), 345–358. 10.1038/s41583-021-00455-7, PubMed: 33837376

Lowet, A. S., Zheng, Q., Matias, S., Drugowitsch, J., & Uchida, N. (2020). Distributional reinforcement learning in the brain. *Trends in Neurosciences*, *43*(12), 980–997. 10.1016/j.tins.2020.09.004, PubMed: 33092893

Mallet, N., Micklem, B. R., Henny, P., Brown, M. T., Williams, C., Bolam, J. P., . . . Magill, P. J. (2012). Dichotomous organization of the external globus pallidus. *Neuron*, *74*(6), 1075–1086. 10.1016/j.neuron.2012.04.027, PubMed: 22726837

Marcott, P. F., Mamaligas, A. A., & Ford, C. P. (2014). Phasic dopamine release drives rapid activation of striatal D2-receptors. *Neuron*, *84*(1), 164–176. 10.1016/j.neuron.2014.08.058, PubMed: 25242218

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*(4), 381–425. 10.1016/S0301-0082(96)00042-1, PubMed: 9004351

Mink, J. W., & Thach, W. T. (1993). Basal ganglia intrinsic circuits and their role in behavior. *Current Opinion in Neurobiology*, *3*(6), 950–957. 10.1016/0959-4388(93)90167-W, PubMed: 8124079

Monakow, K. H., Akert, K., & Kunzle, H. (1978). Projections of the precentral motor cortex and other cortical areas of the frontal lobe to the subthalamic nucleus in the monkey. *Experimental Brain Research*, *33*(3–4), 395–403. 10.1007/BF00235561, PubMed: 83239

Nakano, K., Kayahara, T., Tsutsumi, T., & Ushiro, H. (2000). Neural circuits and functional organization of the striatum. *Journal of Neurology*, *247*(5), V1–V15. 10.1007/PL00007778, PubMed: 11081799

Nishi, A., Bibb, J. A., Snyder, G. L., Higashi, H., Nairn, A. C., & Greengard, P. (2000). Amplification of dopaminergic signaling by a positive feedback loop. In *Proceedings of the National Academy of Sciences*, *97*(23), 12840–12845. 10.1073/pnas.220410397

Nishi, A., Snyder, G. L., & Greengard, P. (1997). Bidirectional regulation of DARPP-32 phosphorylation by dopamine. *Journal of Neuroscience*, *17*(21), 8147–8155. 10.1523/JNEUROSCI.17-21-08147.1997, PubMed: 9334390

Parent, A., & Hazrati, L.-N. (1993). Anatomical aspects of information processing in primate basal ganglia. *Trends in Neurosciences*, *16*(3), 111–116. 10.1016/0166-2236(93)90135-9, PubMed: 7681234

Redgrave, P., & Gurney, K. N. (2006). The short-latency dopamine signal: A role in discovering novel actions? *Nature Reviews Neuroscience*, *7*(12), 967–975. 10.1038/nrn2022, PubMed: 17115078

Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience*, *89*(4), 1009–1023. 10.1016/S0306-4522(98)00319-4, PubMed: 10362291

Richfield, E., Penney, J., & Young, A. (1989). Anatomical and affinity state comparisons between dopamine-D1 and Dopamine-D2 receptors in the rat central nervous-system. *Neuroscience*, *30*(3), 767–777. 10.1016/0306-4522(89)90168-1, PubMed: 2528080

Ringo, J. (1991). Neuronal interconnection as a function of brain size. *Brain Behavior and Evolution*, *38*(1), 1–6. 10.1159/000114375, PubMed: 1657274

Sadek, A. R., Magill, P. J., & Bolam, J. P. (2007). A single-cell analysis of intrinsic connectivity in the rat globus pallidus. *Journal of Neuroscience*, *27*(24), 6352–6362. 10.1523/JNEUROSCI.0953-07.2007, PubMed: 17567796

Sharpe, M. J., Stalnaker, T., Schuck, N. W., Killcross, S., Schoenbaum, G., & Niv, Y. (2019). An integrated model of action selection: Distinct modes of cortical control of striatal decision making. *Annual Review of Psychology*, *70*(1), 53–76. 10.1146/annurev-psych-010418-102824, PubMed: 30260745

Shink, E., Bevan, M. D., Bolam, J. P., & Smith, Y. (1996). The subthalamic nucleus and the external pallidum: Two tightly interconnected structures that control the output of the basal ganglia in the monkey. *Neuroscience*, *73*(2), 335–357. 10.1016/0306-4522(96)00022-X, PubMed: 8783253

Skinbjerg, M., Sibley, D. R., Javitch, J. A., & Abi-Dargham, A. (2012). Imaging the high-affinity state of the dopamine D2 receptor in vivo: Fact or fiction? *Biochemical Pharmacology*, *83*(2), 193–198. 10.1016/j.bcp.2011.09.008, PubMed: 21945484

Smith, Y., Bennett, B., Bolam, J., Parent, A., & Sadikot, A. (1994). Synaptic relationships between dopaminergic afferents and cortical or thalamic input in the sensorimotor territory of the striatum in monkey. *Journal of Comparative Neurology*, *344*(1), 1–19. 10.1002/cne.903440102, PubMed: 7914894

Smith, Y., Bevan, M. D., Shink, E., & Bolam, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience*, *86*(2), 353–387. 10.1016/S0306-4522(97)00608-8, PubMed: 9881853

Surmeier, D. J., Ding, J., Day, M., Wang, Z., & Shen, W. (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends in Neurosciences*, *30*(5), 228–235. 10.1016/j.tins.2007.03.008, PubMed: 17408758

Suryanarayana, S. M., Hellgren-Kotaleski, J., Grillner, & S., Gurney, K. N. (2019). Role for globus pallidus externa revealed in a computational model of action selection in the basal ganglia. *Neural Networks*, *109*, 113–136. 10.1016/j.neunet.2018.10.003, PubMed: 30414556

Svenningsson, P., Nishi, A., Fisone, G., Girault, J.-A., Nairn, A. C., & Greengard, P. (2004). DARPP-32: An integrator of neurotransmission. *Annual Review of Pharmacology and Toxicology*, *44*(1), 269–296. 10.1146/annurev.pharmtox.44.101802.121415, PubMed: 14744247

Tomkins, A., Vasilaki, E., Beste, C., Gurney, K. N., & Humphries, M. D. (2014). Transient and steady-state selection in the striatal microcircuit. *Frontiers in Computational Neuroscience*, *7*(192). 10.3389/fncom.2013.00192, PubMed: 24478684

Wang, Y., Toyoshima, O., Kunimatsu, J., Yamada, H., & Matsumoto, M. (2021). Tonic firing mode of midbrain dopamine neurons continuously tracks reward values changing moment-by-moment. *eLife*, *10*, e63166.

Watts, V. J., & Neve, K. A. (1996). Sensitization of endogenous and recombinant adenylate cyclase by activation of D2 dopamine receptors. *Molecular Pharmacology*, *50*(4), 966–976. 8863843

Watts, V. J., Wiens, B. L., Cumbay, M. G., Vu, M. N., Neve, R. L., & Neve, K. A. (1998). Selective activation of G-alpha-o by D2L dopamine receptors in NS20Y neuroblastoma cells. *Journal of Neuroscience*, *18*(21), 8692–8699. 10.1523/JNEUROSCI.18-21-08692.1998, PubMed: 9786976

Wickens, J. R., Horvitz, J. C., Costa, R. M., & Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. *Journal of Neuroscience*, *27*(31), 8181–8183. 10.1523/JNEUROSCI.1671-07.2007, PubMed: 17670964

Yapo, C., Nair, A. G., Clement, L., Castro, L. R., Hellgren Kotaleski, J., & Vincent, P. (2017). Detection of phasic dopamine by D1 and D2 striatal medium spiny neurons: Comparison of D1 and D2 responsiveness to transient dopamine. *Journal of Physiology*, *595*(24), 7451–7475. 10.1113/JP274475, PubMed: 28782235

Received October 15, 2021; accepted April 1, 2022.