Dopaminergic changes in striatal pathway competition modify specific cognitive decision parameters

Kyle Dunovan (kdunovan@andrew.cmu.edu)
Dept. Psychology & Center for the Neural Basis of Cognition,
Carnegie Mellon University, USA

Catalina Vich
Dept. de Matemtiques i Informtica,
Universitat de les Illes Balears, Spain

Matthew Clapp
Dept. of Biomedical Engineering
University of South Carolina, USA

Jonathan Rubin
Dept. of Mathematics,
University of Pittsburgh, USA

Timothy Verstynen (timothyv@andrew.cmu.edu)
Dept. Psychology & Center for the Neural Basis of Cognition,
Carnegie Mellon University, USA

Abstract

Cortico-basal-ganglia-thalamic (CBGT) networks are critical for adaptive decision-making, yet it remains unclear how their circuit-level properties manifest as cognitive processes. Using a multilevel, biologically plausible modeling approach we simulate CBGT networks to illustrate how (1) dopamine (DA) signals modify the strength of striatal direct (D) and indirect (I) pathways in accordance with a simple reinforcement learning model and (2) learning induced asymmetries in D/I efficacy map to specific cognitive parameters. Simulations of corticostriatal synapses show that DA feedback leads to asymmetrical weights for D and I pathways within a given action channel. The ratio of these weights (w_D/w_I) effectively encodes the action's expected value (Q). We then simulated the full CBGT network in the context of a 2-choice value-based decision task, varying the corticostriatal weighting schemes (i.e., w_D/w_I) for one action channel. The response times from these simulations were fit with a drift-diffusion model (DDM). As w_D/w_I increased, both drift rate (*v*) and boundary height (*a*) parameters changed in the DDM model, with *v* associated with differences in between-channel D pathway activity, while *a* modulated with overall I pathway activity. These simulations show how microscale plasticity at corticostriatal synapses can alter specific macroscale properties of cognitive decision processes.

Keywords: basal ganglia; dopamine; reinforcement learning; drift diffusion models

Background

DA error signals favor the D pathway over the I pathway for rewarding actions with the opposite tendency for aversive ones, effectively encoding the values of alternative actions. It remains unclear how changes in action value influence the mechanisms of the action selection process itself. We used a multi-level modeling approach (Figure 1) to investigate how dopaminergic feedback alters the circuit-level dynamics of CBGT pathways, and ultimately, emergent decision computations.

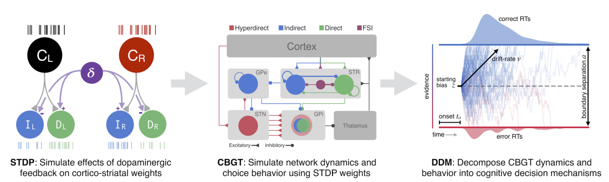

Figure 1: *Multi-level modeling schematic. Left: Spike-timing dependent plasticity (STDP) model of dopaminergic effects on D and I cortico-MSN synapses. Middle: Cortico-basal ganglia-thalamic (CBGT) pathways. Right: Drift-diffusion model (DDM) of two-alternative choice behavior.*

Methods & Results

Spike-timing dependent plasticity (STDP) model

We first simulated the performance of a two-alternative forced choice task in a spiking CBGT network. Performance of an action yields a reward, the size and possibly the sign of which depend on the action selected. The spiking model is augmented with phenomenological representations of (1) action values, $Q_i(t)$ for action $i\{L,R\}$, which are updated based on a difference scheme suggested previously (Mikhael & Bogacz, 2016), and (2) immediate dopamine release, DA(t), the level of which is derived from comparing the reward level r_i resulting from the performance of action *i* to the maximum value believed to be attainable based on past experience. Mathematically, the performance of action *i* at times t_{j-1} yields the following updates at time t_j: $Q_i(t_j) = Q_i(t_{j-1}) + (r_i - Q(t_{j-1}))$ (Eq.1) and $DA(t_j) = r_i - max(Q_L(t_{j-1}), Q_R(t_{j-1}))$ (Eq.2), with exponential decay of DA between updates.

The synaptic conductance $g_k(t)$ to the kth striatal neuron from its cortical input signal is augmented by $g_k(t) \leftarrow g_k(t) + w_k(t)$ if a cortical spike occurs at time *t* (Baladron, Nambu, & Hamker, 2017). Weight $w_k(t)$ evolves as $w_k'(t) = \alpha_w E_k(t) f(DA(t))(w_{max} - w_k(t))$ where $f(DA(t))$ is either DA or $DA/(c + |DA|)$ if the kth neuron is I or D pathway, respectively. The sign of α_w may depend on whether the striatal neuron has a D1 or D2 dopamine receptor, $E_k(t)$ denotes an eligibility based on comparison of previous cortical and striatal spike timing (Gurney, Humphries, & Redgrave, 2015; Baladron et al., 2017), and the sign of DA is determined by (Eqs. 1-2). Time courses of these components, for a simplified simulation of a population of striatal neurons whose synchrony drives selection between two actions with distinct reward outcomes according to a phenomenological rule, show learning of action values and corresponding synaptic weight trajectories (Fig. 2).

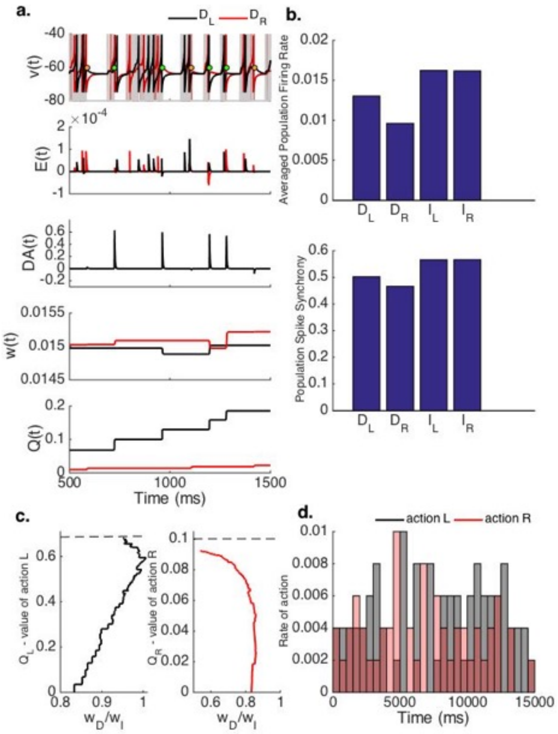

Figure 2: *Corticostriatal synaptic plasticity in the two-alternative forced choice task. A) Striatal spikes yield increases in eligibility and possibly action performance (green dots: larger reward; red dots: smaller reward), resulting in changes in DA, weights, and action value estimates Q. B) Greater value of action L drives greater synchrony in D pathway neurons for action L. C) Values Q converge to reward levels for both actions, while weight ratios w_D/w_I also evolve. D) Action L, with a higher reward value, is performed more than action R.*

Spiking CBGT Network

To determine how increased asymmetries in the cortical influence on D and I pathways impact the decision process by which a single action is selected, the full CBGT network was used to simulate responses in a simple two-alternative forced choice decision task (see (Wei, Rubin, & Wang, 2015)). The relative asymmetry in D and I pathway efficacy was determined based on the relative corticalstriatal synaptic weights from the STDP simulations following a low (0.65), medium (0.75), or high (0.85) reward probability for leftward decisions. On each trial, the action selected and response time (RT) of the decision were determined when the thalamic population of one of the action channels exceeded a threshold of 30Hz. Three experiments (N=2500 trials/experiment) were performed with either low, medium, or high weight ratios for cortical inputs to D and I neurons (see STDP model), capturing the post-training effects of DA under different reward-to-penalty regimes (Fig. 3).

Drift Diffusion Model (DDM) Fits

The RT distributions generated from the spiking CBGT network were then fit to a DDM model (Wiecki, Sofer, & Frank, 2013), in each case leaving either a single or a pair of parameters free to be fit across D/I conditions. Based on fits of the DDM to the distribution of correct and error RTs produced by the network in each condition, we found that increasing reward probability led to an increase in drift-rate (*v*) of evidence accumulation toward the correct (upper; left) decision threshold, as well as a greater separation of correct and incorrect (lower, right) boundaries (*a*). While this combination of drift-rate and boundary height effects provided the best account of the simulated behavior across conditions, we found that *v* and *z* (the initial bias towards one decision boundary or the other) provided a significant improvement in fit beyond that afforded by drift-rate alone (e.g., the best-fitting single parameter model). We next refit the *v,a* and *v,z* models to the same simulated behavioral dataset as in the previous rounds, with the addition of different trialwise measures of striatal activity included as regressors for one of the two free parameters in the DDM. For each regression DDM (N=24), one of the summary measures shown in Figure 4B-C was regressed on *v*, and another regressed on either *a*, or *z*, with separate regression weights estimated for each level of reward probability. The relative goodness-of-fit afforded by all 24 regression models is visualized in Figure 4 (lower panel), revealing model II, with the difference in left and right D pathways mapping onto *v* and average activity of all I pathways mapping onto *a*, as the clear winner with an overall DIC=18860.37, and DIC=9716.17 compared to a null model.

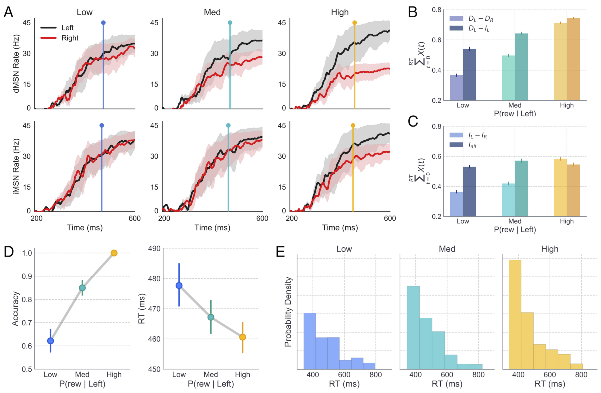

Figure 3: *Striatal pathway dynamics and behavioral effects of reward probability in a spiking CBGT network. A) Time courses show the average population firing rates for left (black) and right (red) D (top) and I (bottom) MSNs. B-C) Summary statistics of D and I population firing rates were extracted on each trial and included as trial-wise regressors on parameters of the DDM, allowing specific hypotheses to be tested about the mapping between neural and cognitive mechanisms. In B) lighter colored bars show the summed difference between D firing rates in the left and right action channels and where the darker colored bars show the summed difference between D and I firing rates in the left action channel. In C) lighter colored bars show the difference between I firing rates in the left and right action channels and darker colored bars show the average I firing rate (across left and right channels). D) Average accuracy (probability of choosing left) and RT (left choices only) of CBGT choices across levels of reward probability. E) RT distributions for correct choices across levels of reward probability.*

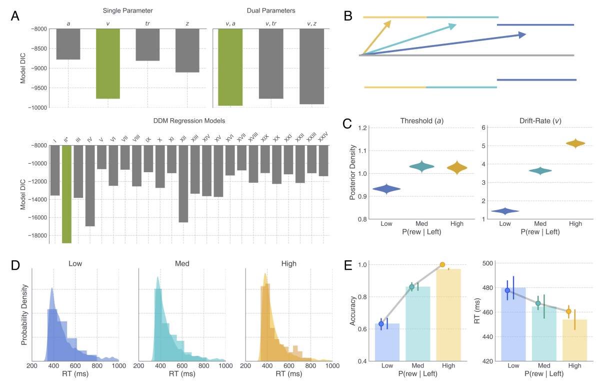

Figure 4: *Model comparison and parameter estimates and behavioral predictions of best-fitting DDM. A) Deviance information criterion (DIC) scores for single- and dual-parameter DDMs (top) and for all DDM regression models (bottom).The DIC score of the best-fitting model at each stage is plotted in green. B) Schematic of DDM showing the change in drift-rate and boundary height across low (blue), medium (green), and high (yellow) reward probability conditions. C) Posterior distributions showing the estimated weights for neural regressors on threshold (a), which was estimated on each trial as a function of the average I firing rate across left and right action channels, and drift-rate (v), which was estimated on each trial as a function of the the difference between D firing rates in the left and right channels. D) Histograms and kernel density estimates showing the CBGT-simulated and DDM-predicted RT distributions, respectively. E) Point plots showing the CBGT networks average accuracy and RT across reward conditions overlaid on bars showing the DDM-predicted averages.*

Discussion

Derived from a multilevel modeling approach, these results show how dopaminergic plasticity at local corticostriatal synapses can alter the relative balance of D and I pathway competition that, in turn, is reflected as changes in specific parameters of cognitive algorithms of decision making. Asymmetries in the efficacy of D, but not I, pathways maps to the speed of information accumulation during decision making while the overall efficacy of all I MSNs sets the level of information needed to gate a decision. These results provide key insights into how plasticity in implementation-level mechanisms in CBGT pathways can impact cognitive processes that can be measured at the behavioral level.

References

Baladron, J., Nambu, A., & Hamker, F. H. (2017). The subthalamic nucleus-external globus pallidus loop biases exploratory decisions towards known alternatives: a neuro-computational study. *European Journal of Neuroscience*, 1–14. Retrieved from http://dx.doi.org/10.1111/ejn.13666 doi: 10.1111/ejn.13666

Gurney, K. N., Humphries, M. D., & Redgrave, P. (2015, 01). A new framework for cortico-striatal plasticity: Behavioural theory meets in vitro data at the reinforcement-action interface. *PLOS Biology*, *13*(1), 1-25. Retrieved from https://doi.org/10.1371/journal.pbio.1002034 doi: 10.1371/journal.pbio.1002034

Mikhael, J. G., & Bogacz, R. (2016, 09). Learning reward uncertainty in the basal ganglia. *PLOS Computational Biology*, *12*(9), 1-28. Retrieved from https://doi.org/10.1371/journal.pcbi.1005062 doi: 10.1371/journal.pcbi.1005062

Wei, W., Rubin, J. E., & Wang, X.-J. (2015). Role of the indirect pathway of the basal ganglia in perceptual decision making. *Journal of Neuroscience*, *35*(9), 4052–4064.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). Hddm: hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in neuroinformatics*, *7*, 14.