7.2: Basal Ganglia, Action Selection and Reinforcement Learning

Figure 7.2: Parallel circuits through the basal ganglia for different regions of the frontal cortex -- each region of frontal cortex has a corresponding basal ganglia circuit, for controlling action selection/initiation in that frontal area. Motor loop: SMA = supplementary motor area -- the associated striatum (putamen) also receives from premotor cortex (PM), and primary motor (M1) and somatosensory (S1) areas -- everything needed to properly contextualize motor actions. Oculomotor loop: FEF = frontal eye fields, also receives from dorsolateral PFC (DLPFC), and posterior parietal cortex (PPC) -- appropriate context for programming eye movements. Prefrontal loop: DLPFC also controlled by posterior parietal cortex, and premotor cortex. Orbitofrontal loop: OFC = orbitofrontal cortex, also receives from inferotemporal cortex (IT), and anterior cingulate cortex (ACC). Cingulate loop: ACC also modulated by hippocampus (HIP), entorhinal cortex (EC), and IT.

The basal ganglia performs its action selection function over a wide range of frontal cortical areas, by virtue of a sequence of parallel loops of connectivity (Figure 7.2). These areas include motor (skeletal muscle control) and oculomotor (eye movement control), but also prefrontal cortex, orbitofrontal cortex, and anterior cingulate cortex, which are not directly motor control areas. Thus, we need to generalize our notion of action selection to include cognitive action selection -- more abstract forms of selection that operate in higher-level cognitive areas of prefrontal cortex. For example, the basal ganglia can control the selection of large-scale action plans and strategies in its connections to the prefrontal cortex. The orbitofrontal cortex is important for encoding the reward value associated with different possible stimulus outcomes, so the basal ganglia connection here is important for driving the updating of these representations as a function of contingencies in the environment. The anterior cingulate cortex is important for encoding the costs of motor actions (time, effort, uncertainty), and basal ganglia similarly can help control updating of these costs as different actions are considered. We can summarize the role of basal ganglia in these more abstract frontal areas as
controlling working memory updating, as is discussed further in the Executive Function chapter.

Interestingly, the additional inputs that converge into the basal ganglia for a given area all make good sense. Motor control needs to know about the current somatosensory state, as well as inputs from the slightly higher-level motor control area known as premotor cortex. Orbitofrontal cortex is all about encoding the reward value of stimuli, and thus needs to get input from IT cortex, which provides the identity of relevant objects in the environment.

Figure 7.3: Gating mechanisms of the direct pathway in the oculomotor circuit. An eye saccade movement is made when the superior colliculus (SC) neurons coding for the saccade direction exhibit burst firing. The SC receives excitatory input from elsewhere (e.g., frontal cortex) indicating planned eye movements. However, the SC is under tonic inhibitory regulation from the output of the basal ganglia (in this circuit, it is the substantia nigra pars reticulata (SNr), equivalent to the GPi for other movements). SNr neurons fire at high tonic rates in the absence of input, and prevent the SC from initiating a burst. Neurons in the caudate nucleus (part of the striatum), upstream of the SNr, are normally silent but fire when detecting the appropriate conditions under which to initiate the eye movement (e.g., when it is predictive of reward). Caudate neurons inhibit the SNr, causing a pause in tonic firing, and disinhibit the SC. This disinhibition acts as a gating mechanism because the Caudate does not directly elicit SC firing but instead allows SC to burst fire in particular SC neurons that also receive excitatory input about the planned movement. Not shown here are indirect pathway Caudate neurons which would have the opposite effect, increasing SNr activity and preventing gating of particular movements. From Hikosaka, 2007.

Figure 7.4: Biology of the basal ganglia system, with two cases shown: a) Dopamine burst activity that drives the direct "Go" pathway neurons in the striatum, which then inhibit the tonic activation in the globus pallidus internal segment (GPI), which releases specific nuclei in the thalamus from this inhibition, allowing them to complete a bidirectional excitatory circuit with the frontal cortex, resulting in the initiation of a motor action. The increased Go activity during dopamine bursts results in potentiation of corticostriatal synapses, and hence learning to select actions that tend to result in positive outcomes. b) Dopamine dip (pause in tonic dopamine neuron firing), leading to preferential activity of indirect "NoGo" pathway neurons in the striatum, which inhibit the external segment globus pallidus neurons (GPe), which are otherwise tonically active, and inhibiting the GPI. Increased NoGo activity thus results in disinhibition of GPe, making it more active and thus inhibiting the thalamus, preventing initiation of the corresponding motor action. The dopamine dip results in potentiation of corticostriatal NoGo synapses, and hence learning to avoid selection actions that tend to result in negative outcomes. From Frank, 2005.
Zooming in on any one of these loops, the critical elements of the basal ganglia system are diagrammed in Figure 7.4, with two important activation patterns shown. First, the basal ganglia system involves the following subregions:

- **The striatum**, which is the major input region, consisting of the caudate and putamen subdivisions (as shown in Figure 7.2). The striatum is anatomically subdivided into many small clusters of neurons, with two major types of clusters: patch/striosomes and matrix/matrisomes. The matrix clusters contain direct (Go) and indirect (NoGo) pathway medium spiny neurons, which together make up 95% of striatal cells, both of which receive excitatory inputs from all over the cortex but are inhibitory on their downstream targets in the globus pallidus as described next. The patch cells project to the dopaminergic system, and thus appear to play a more indirect role in modulating learning signals. There are also a relatively few widely spaced tonically active neurons (TAN's), which release acetylcholine as a neurotransmitter and appear to play a modulatory role, and inhibitory interneurons, which likely perform the same kind of dynamic gain control that they play in the cortex.

- **The globus pallidus, internal segment (GPI)**, which is a much smaller structure than the striatum, and contains neurons that are constantly (tonically) active even with no additional input. These neurons send inhibition to specific nuclei in the thalamus. When the direct/Go pathway striatum neurons fire, they inhibit these GPI neurons, and thus disinhibit the thalamus, resulting ultimately in the initiation of a specific motor or cognitive action (depending on which circuit is involved). Note that in other fronto-basal ganglia circuits, the role of the GPI is taken up by the substantia nigra pars reticulata (SNr), which is situated identically to the GPI anatomically, but receives from other areas of striatum and projects to outputs regulating other actions (e.g., eye movements in the superior colliculus).

- **The globus pallidus, external segment (GPe)**, which is also small, and contains tonically active neurons that send focused inhibitory projections to corresponding GPI neurons. When the indirect/NoGo pathway neurons in the striatum fire, they inhibit the GPe neurons, and thus disinhibit the GPi neurons, causing them to provide even greater inhibition onto the thalamus. This blocks the initiation of specific actions coded by the population of active NoGo neurons.

- **The thalamus**, specifically the medial dorsal (MD), ventral anterior (VA), and ventrolateral (VL) nuclei (as shown in Figure 7.2). When the thalamic neurons get disinhibited by Go pathway firing, they can fire, but only when driven by top-down excitatory input from the frontal cortex. In this way, the basal ganglia serve as a gate on the thalamocortical circuit -- Go firing opens the gate, while NoGo firing closes it, but the contents of the information that go through the gate (e.g., the specifics of the motor action plan) depend on the thalamocortical system. In the oculomotor circuit (as shown in Figure 7.3, the role of the thalamus is taken up by the superior colliculus, the burst firing of which initiates eye saccades).

- **The substantia nigra pars compacta (SNC)** has neurons that release the neuromodulator dopamine, and specifically innervate the striatum. Interestingly, there are two different kinds of dopamine receptors in the striatum. D1 receptors are prevalent in Go pathway neurons, and dopamine has an excitatory effect on neurons with D1 receptors (particularly those neurons that are receiving convergent glutamatergic excitatory input from cortex). In contrast, D2 receptors are prevalent in NoGo pathway neurons, and dopamine has an inhibitory effect via the D2 receptors. Thus, when a burst of dopamine hits the striatum, it further excites active Go units and inhibits NoGo units. This change in activity results in activity-dependent plasticity, and thus leads to an increased propensity for initiating motor and cognitive actions. In contrast, when a dip in dopamine firing occurs, Go neurons are less excited, while NoGo neurons are disinhibited, and thus those NoGo neurons receiving excitatory input from cortex (representing the current state and action) will become more excited due to the dopamine dip. Again, this change in activity results in potentiation of synapses, such that this specific population of NoGo neurons will be more likely to become active in future encounters of this sensory state and candidate motor action. Both of these effects of dopamine bursts and dips make perfect sense: dopamine bursts are associated with positive reward prediction errors (when rewards are better than expected), and thus reinforce selection of actions that lead to good results. Conversely, dopamine dips are associated with negative reward prediction errors (worse than expected) and thus lead to avoidance (NoGo) of those actions that tend to result in these bad results. Also, tonic levels of dopamine can influence the relative balance of activity of these pathways, so that even if learning has already occurred, changes in dopamine can affect whether action selection is influenced primarily by learned Go vs learned NoGo values -- roughly speaking, the higher the dopamine, the more risky the choices (insensitivity to negative outcomes).
• The **subthalamic nucleus** is also a major component of the basal ganglia (not pictured in the figure), which acts as the third **hyperdirect pathway**, so named because it receives input directly from frontal cortex and sends excitatory projections directly to BG output (GPI), bypassing the striatum altogether. These STN-GPI projections are diffuse, meaning that a single STN neuron projects broadly to many GPI neurons, and as such the STN is thought to provide a global **NoGo** function that prevents gating of any motor or cognitive action (technically, it raises the threshold for gating). This area has been shown in models and empirical data to become more active with increasing demands for response inhibition or when there is conflict between alternative cortical action plans, so that the STN buys more time for striatal gating to settle on the best action (Frank, 2006).

This is a fairly complex circuit, and it probably takes a few iterations through it to really understand how all the parts fit together. The bottom line should nevertheless be easier to understand: the basal ganglia learn to select rewarding actions (including more abstract cognitive actions), via a disinhibitory gating relationship with different areas of frontal cortex. Moreover, the general depiction above, motivated by computational considerations and a lot of detailed anatomical, physiological, and pharmacological data, has been overwhelmingly been supported by empirical data across species. For example, in mice, Kravitz et al., 2010 showed that selective stimulation of D1 striatal neurons resulted in inhibition of BG output nuclei and disinhibition of motor actions, whereas selective stimulation of D2 striatal neurons resulted in excitation of output nuclei and suppression of motor actions. A follow up paper in 2012 showed that transient stimulations of these pathways after movements causes the mouse to be more likely (go unit stimulation) or less likely (nogo unit stimulation) to repeat that same movement in the future, consistent with a learning effect. Zalocusky et al., 2016 showed that when a mouse experiences a negative reward prediction error (i.e. they expect a reward but don't get one), the D2 neurons respond by increasing their activity levels, and the extent of this is related to their subsequent avoidance of the action in favor of a safer option leading to certain reward. Shen et al., 2008 showed evidence for the model prediction that D1 and D2 receptors oppositely modulate synaptic plasticity in the two pathways. Hikida et al., 2010 showed that selective blockade of neurotransmission along the Go pathway resulted in impairments in learning to select rewarding actions but no deficits in avoiding punishing actions, and exactly the opposite pattern of impairments was observed after blockade of the NoGo pathway. In humans, striatal dopamine depletions associated with Parkinson's disease result in impaired "Go learning" in probabilistic reinforcement learning tasks, but enhanced "NoGo learning", with the opposite pattern of findings elicited by medications that increase striatal dopamine (Frank, Seeberger & O'Reilly 2004 and many others since). Even individual differences in young healthy human performance in Go vs NoGo learning are associated with genetic variants that affect striatal D1 and D2 receptor function and D1 vs D2 receptor expression in PET studies (Cox et al, 2015; for review of this body of literature, see Frank & Fossella, 2011).

The division of labor between frontal cortex and basal ganglia is such that the frontal cortex entertains many different possible actions, by virtue of rich patterns of connectivity from other cortical areas providing high-level summaries of the current environment, which then activate a range of different possible actions, and the basal ganglia then selects the best (most likely to be rewarding) of these actions to actually execute. In more anthropomorphic terms, the frontal cortex is the fuzzy creative type, with a million ideas, but no ability to focus on the real world, and it has a hard time narrowing things down to the point of actually doing anything: kind of a dreamer. Meanwhile, the basal ganglia is a real take-charge type who always has the bottom line in mind, and can make the tough decisions and get things done. We need both of these personalities in our heads (although people clearly differ in how much of each they have), and the neural systems that support these different modes of behavior are clearly different. This is presumably why there are two separable systems (frontal cortex and basal ganglia) that nevertheless work very closely together to solve the overall action selection problem.
Exploration of the Basal Ganglia

Open [BG](https://med.libretexts.org/Bookshelves/Pharmacology_and_Neuroscience/Book%3A_Computational_Cognitive_Neuroscience…) for an exploration of a basic model of go vs. nogo action selection and learning dynamics in the basal ganglia. This model also allows you to investigate the effects of Parkinson's disease and dopaminergic medications.