CORTEX 64 (2015) 123–135

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Review

Learning robust cortico-cortical associations with the basal ganglia: An integrative review

Sébastien Hélie [a,*], Shawn W. Ell [b] and F. Gregory Ashby [c]

[a] Purdue University, USA
[b] University of Maine, USA
[c] University of California, Santa Barbara, USA

ARTICLE INFO

Article history:
Received 22 April 2014
Reviewed 25 June 2011
Revised 8 August 2014
Accepted 13 October 2014
Action editor Sarah MacPherson
Published online 27 October 2014

Keywords:
Basal ganglia
Prefrontal cortex
Premotor cortex
Supplementary motor area
Fronto-striatal connectivity

ABSTRACT

This article focuses on the interaction between the basal ganglia (BG) and prefrontal cortex (PFC). The BG are a group of nuclei at the base of the forebrain that are highly connected with cortex. A century of research suggests that the role of the BG is not exclusively motor, and that the BG also play an important role in learning and memory. In this review article, we argue that one important role of the BG is to train connections between posterior cortical areas and frontal cortical regions that are responsible for automatic behavior after extensive training. According to this view, one effect of BG trial-and-error learning is to activate the correct frontal areas shortly after posterior associative cortex activation, thus allowing for Hebbian learning of robust, fast, and efficient cortico-cortical processing. This hypothesized process is general, and the content of the learned associations depends on the specific areas involved (e.g., associations involving premotor areas would be more closely related to behavior than associations involving the PFC). We review experiments aimed at pinpointing the function of the BG and the frontal cortex and show that these results are consistent with the view that the BG is a general purpose trainer for cortico-cortical connections. We conclude with a discussion of some implications of the integrative framework and how this can help better understand the role of the BG in many different tasks.

1. Introduction

The basal ganglia (BG) are a group of nuclei at the base of the forebrain that are highly connected with cortex. Early on, the BG were assigned a role in motor functions, whereas cortex was assigned a more 'cognitive' role (e.g., Vogt, 1911; Wilson, 1912). However, a subsequent century of research suggests that the role of the BG is not exclusively motor, and that the BG also play an important role in learning and memory (Helie, Chakravarthy, & Moustafa, 2013; Packard & Knowlton, 2002). For example, Wise, Murray, and Gerfen (1996) have argued that 'rules' are represented in frontal cortex and that the BG are a context detector that are involved in changing the appropriate rule in cortex as a function of context – e.g., to

* *Corresponding author*. Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA.
E-mail address: shelie@purdue.edu (S. Hélie).
http://dx.doi.org/10.1016/j.cortex.2014.10.011
0010-9452/

disengage from the current behavior by switching to a new rule. In this review article, we argue that the role of the BG is not restricted to learning contexts using trial-and-error learning, but also to train connections between posterior cortical areas and frontal cortical regions that are responsible for automatic behavior after extensive training (Ashby, Ennis, & Spiering, 2007). According to this view, one effect of BG trial-and-error learning is to activate the correct frontal areas (e.g., rules or responses) shortly after posterior associative cortex activation (e.g., the context or stimulus), thus allowing for Hebbian learning of robust, fast, and efficient cortical–cortical processing. This hypothesized process is general, and the content of the learned associations depends on the specific areas involved (e.g., associations involving premotor areas would be more closely related to behavior than associations involving the prefrontal cortex; PFC).

In this article, we review experiments aimed at pinpointing the function of the BG and the frontal cortex and show that these results are consistent with the view that the BG is a general purpose trainer for cortico-cortical connections. The remainder of this article is organized as follows. First, we review relevant anatomy of the BG and frontal cortex. Next, we detail how the anatomy and connectivity of the BG and frontal cortex supports the learning of cortico-cortical associations. Following this presentation, we review and re-interpret data collected with the goal of elucidating the conditions under which the BG can be used to train cortico-cortical connections. This presentation is followed by possible alternative accounts of the data, including the view that automaticity is characterized by a progressive transfer between anterior and posterior striatum. We conclude with a discussion of the theoretical impact of this integrative role of the BG and frontal cortex and propose future experiments that would provide a direct test of the new integrative framework.

2. Anatomy

2.1. *Frontal cortex*

2.1.1. *PFC*

The PFC is a central hub in the brain that is connected to all other cortical areas except for the primary sensory areas and the primary motor area (Miller & Cohen, 2001). While there are many ways to identify the PFC, in this article we define the PFC as the cortical region receiving afferent connections from the medial dorsal nucleus (MDN) of the thalamus (Fuster, 2008). A schematic of the frontal cortex, and its subdivisions, is shown in Fig. 1.

The PFC is typically separated into lateral PFC (Brodmann areas 8, 9, 10, 44, 45, and 46), medial PFC (areas 8m, 9m, 10m, 11m, and 12), and orbitofrontal PFC (areas 10, 11, and 47). The principal neurons in the PFC are pyramidal neurons and stellate neurons. The pyramidal neurons are excitatory (glutamatergic) whereas the stellate cells are inhibitory (GABAergic) interneurons. The pyramidal neurons are connected with other PFC neurons, but also with neurons in other brain regions (e.g., MDN of the thalamus, associative sensory areas). In contrast, the stellate neurons mostly synapse within the PFC. Note that most subdivisions of the PFC are heavily interconnected. Hence, the PFC as a whole can be considered as a single, unified brain area. The separation is mostly functional and based on connectivity with other brain areas (although cytoarchitectonic differences are also present).

2.1.2. *Premotor areas (area 6)*

The premotor areas (area 6) are caudally adjacent to area 4 (primary motor cortex) and rostrally adjacent to areas 8, 9, and 44 of the PFC. The cytoarchitecture of area 6 allows for the identification of at least three different structures (Barbas & Pandya, 1987). First, area 6 can be divided into a ventral and a dorsal area. The ventral premotor area (PMv) has an emergent layer IV that separates layers III and V, whereas layers III and V merge to form a prominent central band in the dorsal premotor areas. In addition, the pyramidal neurons in the dorsal areas are generally smaller than in the PMv, and the myelin content is less dense. These architectonic features suggest that the dorsal premotor areas are more similar to motor cortex, whereas PMv is more closely related to sensory cortex (Barbas & Pandya, 1987).

The dorsal premotor areas can be further subdivided into a lateral and a medial part. The medial part is called the supplementary motor area (SMA) and contains prominent medium-size pyramidal neurons in layers III and V that are more compact and darkly stained when compared with the lateral part of dorsal premotor area (PMd). In addition, the SMA is often split into a rostral and a caudal region using the vertical commissure anterior (VCA) (Nachev, Kennard, & Husain, 2008). The portion of SMA caudal to the VCA is simply called the SMA and produces movement when a current is

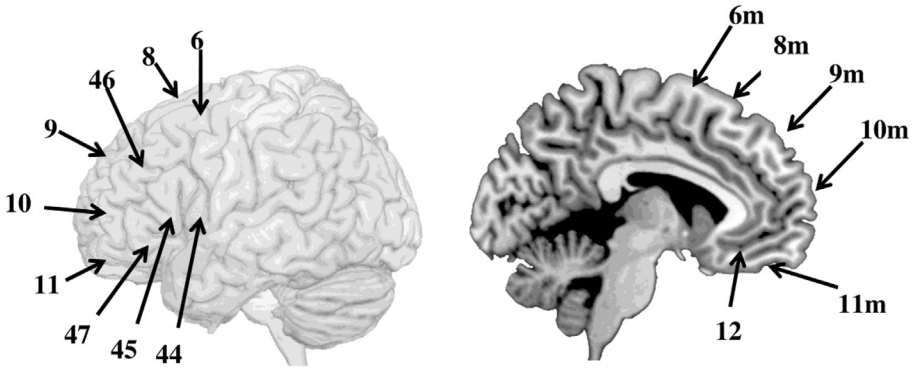

Fig. 1 – Subdivisions of frontal cortex. Numbers refer to Brodmann areas.

injected. In contrast, the region rostral to the VCA, called the pre-SMA, does not project to any premotor or motor areas (including the SMA), and instead sends all its cortical projections to the PFC (Dum & Strick, 1991). As a result, whereas SMA is similar to other premotor regions, pre-SMA is more similar to PFC (Akkal, Dum, & Strick, 2007).

2.1.3. *Frontal connectivity*

Frontal cortex receives its sensory input from higher-level associative cortical areas (e.g., posterior parietal cortex, inferotemporal cortex) and can produce actions through the premotor areas. Most of frontal cortex has excitatory efferent and afferent connections with other cortical areas (Fuster, 2008). Thus, frontal cortex is in a position to influence, and be influenced by, almost all other cortical regions.

First, lateral PFC is bidirectionally connected with all sensory association areas of cortex. The visual information reaching the PFC is highly processed and integrated (especially in humans, via extensive crosstalk between the ventral and dorsal visual pathways). For instance, occipital, temporal, parietal and other areas of PFC that respond to sensory information all converge onto the frontal eye fields (area 8). Sensory pathways (from primary to associative areas) are often thought to represent information hierarchically and every level of the hierarchy projects to the PFC (except for primary visual cortex). Projections from sensory areas suggest that lateral PFC is mostly concerned with visual, auditory, and somatic processing. In particular, lateral PFC has important efferent connections to posterior parietal cortex.

Lateral PFC is also bidirectionally connected with lateral premotor areas (i.e., PMd and PMv). Connections between dorsolateral PFC and lateral premotor areas are stronger than those between ventrolateral PFC and lateral premotor areas, and most connections between lateral PFC and the lateral premotor areas target PMv (Lu, Preston, & Strick, 1994). These in turn project to primary motor cortex and then to the spinal cord to produce movement. As in sensory areas, each area downstream of lateral PFC represents a more effector-specific version of the action and projects back to lateral PFC (except for primary motor cortex). This extensive reciprocal connectivity suggests that lateral PFC may be implicated in goal-directed behavior (Fuster, 2008).

In contrast to lateral PFC, medial and orbitofrontal PFC are mostly connected to the auditory association areas and the associative visual areas of the ventral stream (e.g., inferotemporal cortex). More specifically, medial PFC receives limited projections from visual and somatic areas, but substantial connections from auditory association areas. In contrast, the orbitofrontal PFC receives direct inputs from gustatory and olfactory areas, in addition to activation from other associative sensory areas (i.e., visual, auditory, and somatosensory). Many of these sensory projections overlap in ventral and medial PFC.

At the subcortical level, the PFC has prominent projections to the BG (described below) and the thalamus. PFC is bidirectionally connected with several thalamic nuclei (e.g., ventrolateral, ventral anterior, medial dorsal), but over 80% of its connections are with the MDN (Barbas, Henion, & Dermon, 1991). These connections are topographically organized, with more rostral parts of PFC being connected with more lateral parts of each of these thalamic nuclei (Fuster, 2008). In contrast, premotor areas are mostly connected with the ventrolateral and ventral anterior nuclei of the thalamus.

Frontal cortex also receives neuromodulatory projections from the basal forebrain, the locus coeruleus, the ventral tegmental area (VTA), and the raphe nuclei. These projections supply the frontal cortex with neuromodulators such as acetylcholine, norepinephrine, dopamine (DA), and serotonin. Interestingly, the PFC is able to regulate the impact of at least some of these neuromodulators. For example, the PFC modulates the mesocortical DA system via projections to the VTA (Karreman & Moghaddam, 1996).

2.2. *BG*

The BG include the striatum (caudate nucleus, putamen, nucleus accumbens), the globus pallidus (GP), subthalamic nucleus (STN), substantia nigra (SN), VTA, and olfactory tubercle. These structures (except the olfactory tubercle) and some of their most prominent projections are illustrated in Fig. 2. The major input structures within the BG, which receive inputs from all of cortex (except V1) are the striatum and STN. The primary output structures are the internal segment of the GP (GPi) and the SN pars reticulata (SNr), which project widely to cortex via the thalamus. Compared to cortex, the organization of the striatum is simple. It contains two general classes of neurons. Most common (~95%) are the GABAergic medium spiny neurons (MSN), which receive input from cortex and send their axons to BG output structures (Bolam, Powell, Wu, & Smith, 1985). The remaining 5% of striatal neurons are aspiny interneurons (Kemp & Powell, 1971), which can be further divided into large aspiny neurons and medium aspiny neurons.

Based on both structural and functional evidence, the striatum is often divided into ventral and dorsal components. The ventral striatum includes the nucleus accumbens, ventromedial portions of the caudate and putamen, and the olfactory tubercle. The dorsal striatum includes the remainder of the caudate and putamen. The caudate and putamen

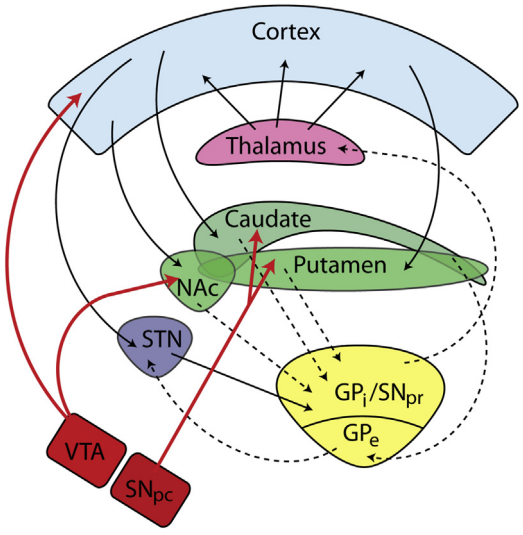

Fig. 2 – Major components of the BG. Full lines are excitatory while dashed lines are inhibitory.

themselves are often divided into the associative striatum, which includes all of the caudate and the anterior putamen, and the motor striatum, which includes the posterior putamen (Parent & Hazrati, 1995). In rodents, the dorsomedial and dorsolateral striatum are homologous with the anterior and motor striatum, respectively (Joel & Weiner, 2000).

2.2.1. *Afferent connections to the striatum*

Virtually all of the neocortex sends excitatory (glutamatergic) projections to the striatum (Reiner, 2010). These projections arise almost exclusively from pyramidal neurons in layers III and V of neocortex. These cortical inputs are massively convergent with estimates ranging from 5,000 to 30,000 cortical neurons converging on a single striatal MSN (Bolam et al., 2006; Kincaid, Zheng, & Wilson, 1998; Wilson, 1995). A slightly more controversial factor is the extent to which individual cortico-striatal projections diverge. Any given cortico-striatal pyramidal neuron synapses within the dendritic field of multiple MSNs. The degree of divergence, however, is thought to be extremely small relative to the degree of cortico-striatal convergence with each cortico-striatal neuron innervating approximately 10–100 MSNs (Wickens & Arbuthnott, 2010).

Classically, cortico-striatal organization was thought to follow a fairly strict spatial topography (Kemp & Powell, 1970). For instance, ventral striatum receives input predominantly from orbitofrontal cortex, ventromedial PFC, and anterior cingulate cortex (ACC). As one moves caudally within the striatum, inputs from areas 9, 46, and 8 become more prevalent (Calzavara, Mailly, & Haber, 2007; Haber, Kim, Mailly, & Calzavara, 2006; Selemon & Goldman-Rakic, 1985), followed by inputs from premotor regions (area 6), with the most caudal motor and somatosensory cortical regions projecting preferentially to the caudal putamen (Alexander & DeLong, 1985; Flaherty & Graybiel, 1994). These topographic projections continue rostrally and ventrally through parietal and temporal cortices as well as other extrastriate visual areas (Kemp & Powell, 1970; Selemon & Goldman-Rakic, 1985; Van Hoesen, Yeterian, & Lavizzo-Mourey, 1981; Webster, Bachevalier, & Ungerleider, 1993; Yeterian & Pandya, 1993, 1995, 1998; Yeterian & Van Hoesen, 1978).

Recent tracing studies have confirmed the presence of focused, convergent cortico-striatal projections that are both topographically and functionally organized (Calzavara et al., 2007; Haber et al., 2006). In addition, these studies have found more diffuse projections suggesting some degree of divergence between distant striatal targets. This divergence may provide a possible means by which information from functionally distinct cortical regions (or loops) are integrated (Haber, 2003).

2.2.2. *BG efferent connections*

The GABAergic projection neurons of the GPi and SNr are the primary sources of output from the BG. The vast majority of BG output targets frontal cortex via the thalamus (Gerfen & Bolam, 2010). The fact that the striatal and GPi output neurons are both GABAergic (see Fig. 2) is rather unusual. However, the GPi neurons have a high tonic firing rate, whereas the striatal MSNs have a low tonic rate. Thus, under tonic conditions, GPi inhibits its thalamic targets, which prevents thalamus from exciting cortex. Striatal activity disinhibits thalamus by reducing GPi output. The effect of this disinhibition on thalamic firing is unclear. One possibility is that it allows thalamus to fire to other excitatory inputs. For example, in the case of the MDN, there could be excitatory inputs from PFC that exert control during periods of striatal activity. According to this model, striatal activity does not cause cortical activation, it merely allows other brain regions (e.g., PFC) to exert cortical control. A second possibility, however, follows from the observation that many thalamic neurons fire a rebound burst following a release from inhibition (e.g., Kim, Sanchez-Vives, & McCormick, 1997). This phenomenon provides a potential mechanism via which striatal activity could cause a cortical response.

The DAergic neurons of the SNc/VTA send projections throughout the brain including, but not limited to PFC, ACC, thalamus, hippocampus, striatum, external segment of the GP, and STN. The SNc/VTA receives projections from the GP and SNr as well as several subcortical structures including the amygdala, the pedunculopontine tegmental nucleus (PPTN), and the lateral habenular nucleus (Gerfen & Bolam, 2010). The nigrostriatal DAergic system typically refers to SNc DAergic inputs to the dorsal striatum, whereas the mesocorticolimbic DAergic system typically refers to the VTA DAergic inputs to ventral striatum and frontal cortex.

2.2.3. *Cortical-striatal-thalamic-cortical loops*

It has been proposed that many cortical targets of striatal projections (via the thalamus) ultimately are in regions where cortico-striatal afferents originate, thereby providing the anatomical basis for a number of functionally distinct cortico-striatal loops (e.g., motor, oculomotor, prefrontal, orbitofrontal, cingulate) (Alexander, DeLong, & Strick, 1986). For example, the motor loop connects the SMA (and other cortical motor regions) with the putamen, the oculomotor loop connects the frontal eye fields with the body of the caudate nucleus, the prefrontal loop connects the dorsolateral PFC with the dorsolateral caudate nucleus, the orbitofrontal loop connects the lateral orbitofrontal cortex with the ventromedial caudate nucleus, and the cingulate loop connects the ACC with the ventral striatum. Although these loops are not strictly 'closed' (e.g., dorsolateral PFC provides input to the oculomotor and prefrontal loops), each of these loops was originally thought to be largely functionally independent.

It is highly unlikely, however, that these loops are truly functionally independent because there are numerous points at which information could be integrated across loops. For instance, BG output can also target cortical regions other than the source of input, thereby providing a powerful mechanism for integrating information from functionally distinct cortical regions. In light of the high likelihood for crosstalk among loops and functional views of cortico-striatal input, several researchers have proposed revisions to the Alexander et al. (1986) model (e.g., Haber, 2003; Parent & Hazrati, 1995; Seger, 2006). These revisions generally suggest a smaller number of functionally defined loops that span the caudate and putamen (e.g., affective, cognitive, sensorimotor). The revised loops have received considerable support in monkeys (e.g., Calzavara et al., 2007) and humans (Di Martino et al., 2008; Draganski et al., 2008; Postuma & Dagher, 2006). The contributions of these loops to cognition have been the focus of

several recent reviews (e.g., Chudasama & Robbins, 2006; Ell, Helie, & Hutchison, 2012; Helie et al., 2013; Seger & Miller, 2010).

3. The BG as a general training machine for cortico-cortical connections

Doya (2000) suggested that cortical learning and BG learning were best described by different learning algorithms, namely Hebbian learning and reinforcement learning (RL) (respectively). In Hebbian learning, long-term potentiation (LTP) occurs at synapses with strongly correlated pre- and post-synaptic activity whereas long-term depression occurs at synapses with weakly correlated pre- and post-synaptic synapses. RL is similar to Hebbian learning with the additional requirement of a learning (or reinforcement) signal. DA has been argued to be the RL signal in the BG (e.g., Montague, Dayan, & Sejnowski, 1996).

Much evidence supports the hypothesis that LTP at cortico-striatal synapses requires strong pre- and post-synaptic activation and DA levels above baseline (e.g., Arbuthnott, Ingham, & Wickens, 2000). However, if strong pre- and post-synaptic activation is present but DA release is below baseline, then LTD is observed (Ronesi & Lovinger, 2005). As a result, many researchers have proposed that DA is critical for striatal-based RL (e.g., Houk, Adams, & Barto, 1995).

A necessary feature of any RL signal is high temporal resolution. Consider an instrumental learning task in which the goal is to learn a number of stimulus–response associations. Following a correct response, DA must be released into the relevant synapses quickly, before the critical traces disappear. But after the correct synapses have been strengthened, it is also essential that excess DA be quickly cleared from the synapse. If it is not, and the response on the next trial is an error, then the residual DA will strengthen inappropriate synapses – namely, those responsible for producing the incorrect response. This would undo the beneficial learning that occurred following correct responses, and prevent discrimination learning. Within the striatum, DA is quickly cleared from synapses by dopamine active transporter (DAT) and, as a result, the temporal resolution of DA in striatum is high enough for DA to serve as an effective RL signal. Unlike the striatum however, DAT concentrations in frontal cortex are low (e.g., Seamans & Robbins, 2009). As a result, cortical DA levels change slowly. For example, the delivery of a single food pellet to a hungry rat increases DA levels in PFC above baseline for approximately 30 min (Feenstra & Botterblom, 1996). Thus, the first rewarded behavior in a training session is likely to cause frontal cortical DA levels to rise, and the absence of DAT will cause DA levels in frontal cortex to remain high throughout the training session. As a result, all synapses that are activated during the session are likely to be strengthened, regardless of whether the associated behavior is appropriate or not. Thus, although DA may facilitate LTP in frontal cortex, it appears to operate too slowly to serve as a frontal-cortical RL signal (Lapish, Kroener, Durstewitz, Lavin, & Seamans, 2007).

Instead, it is thought that cortical LTP/LTD follows Hebbian learning rules (Feldman, 2009), which means that cortical LTP can occur regardless of whether the resulting behavior was rewarded. This implies that the BG may be necessary in cases where learning depends, in part, upon errors, whereas frontal cortical contributions to learning may be restricted to correct responses. For these reasons, the BG, but not frontal cortex, may be suitable for the learning of skilled behaviors that depend upon trial-and-error learning.

A more controversial question, however, is whether the BG are the long-term store for such learning. In fact, a number of results are problematic for this hypothesis. Section 4 describes some of these results, but one example is that people with Parkinson's disease, who have DA reductions and striatal dysfunction, are impaired in procedural-learning tasks (Soliveri, Brown, Jahanshahi, Caraceni, & Marsden, 1997; Thomas-Ollivier et al., 1999) but relatively normal in producing automatic behaviors (Asmus, Huber, Gasser, & Schöls, 2008). Based on evidence such as this, Ashby et al. (2007) proposed that automatic behaviors are stored in cortex, but learned in the BG. According to this account, a critical function of the BG is to train purely cortical representations of automatic behaviors. The idea is that, via RL, the BG learns to activate the correct post-synaptic target in frontal cortex, which allows the appropriate cortico-cortical synapses to be strengthened via Hebbian learning (because the product of pre- and post-synaptic activations will be greatest at the correct synapse). Once the cortico-cortical synapses have been built, the BG is no longer required to produce the appropriate behavior. In the following section, we argue that this learning process is general and underlies many functions that depend upon interactions between the BG and frontal cortex.

4. Functions

4.1. *Working memory (WM)*

Classically, WM maintenance has been associated with closed loops within the PFC (Buchsbaum & D'Esposito, 2008; Wager & Smith, 2003), while manipulation of information has been associated with both the BG (Cools, Sheridan, Jacobs, & D'Esposito, 2007) and lateral PFC (Osaka & Osaka, 2007; Wager & Smith, 2003). This view is best summarized by an influential model proposed by Goldman-Rakic and colleagues (Wilson, Scalaidhe, & Goldman-Rakic, 1993), which assumes that object memory follows a ventral pathway whereas spatial memory follows a dorsal pathway (similar to the famous visual pathways hypothesis). While this model is supported by much data collected from monkeys (Goldman-Rakic, 1987), more recent human neuroimaging data suggest that this ventral—dorsal separation may not be an appropriate reflection of human WM (D'Esposito et al., 1998). Specifically, D'Esposito and colleagues (Buchsbaum & D'Esposito, 2008; Cools et al., 2007) argued that WM maintenance involves the ventrolateral PFC, whereas a primary function of dorsolateral PFC is to manipulate information (Osaka & Osaka, 2007), more specifically by selecting or focusing attention on specific items. As a result, the ventrolateral PFC is often sufficient for WM maintenance in low-load conditions (e.g., as measured by forward memory span). However, the dorsolateral PFC gets recruited in high-load conditions, especially when the load exceeds WM capacity. The BG are recruited to facilitate the entry of new information into WM and/or to switch attention between different WM

items (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Frank, Loughry, & O'Reilly, 2001; Monchi, Taylor, & Dagher, 2000; Taylor & Taylor, 2000). This is because cortical DA is thought to stabilize representations, whereas BG DA is used to quickly update the representations in a task-relevant manner (Cools et al., 2007; Hélie, Paul, & Ashby, 2012a, 2012b).

In addition, there is mounting evidence that the BG may also help mediate WM maintenance—not just manipulation. For instance, Ashby, Ell, Valentin, and Casale (2005) reviewed a variety of evidence supporting the hypothesis that the thalamus, BG, and posterior cortex all play critical roles in a widely distributed WM circuit. First, lesions to the MDN have been reported to impair WM in both animals and humans (Van der Werf, Witter, Uylings, & Jolles, 2000). Second, patients with BG lesions have WM deficits (Voytek & Knight, 2010). Third, increased thalamic and BG activation have been reported in some neuroimaging studies of WM (Cairo, Liddle, Woodward, & Ngan, 2004; Callicott et al., 1999; Chang, Crottaz-Herbette, & Menon, 2007; Jonides et al., 1997). Finally, single-cell recording data from a number of brain regions show an increased activation (relative to baseline) during the entire delay period in WM tasks. Specifically, some neurons in the caudate nucleus, GP, and MDN of the thalamus do not respond to the initial presentation of the target. Instead, their firing rates change at the beginning of the delay period. Finding neurons in the thalamus and BG with sustained delay-related activity that is correlated with the activity of WM units in PFC supports the view that the BG also play a role in WM maintenance (Ashby et al., 2005; Schroll, Vitay, & Hamker, 2012).

All of these regions were included in the FROntal-STriatal (FROST) model of WM proposed by Ashby and colleagues (Ashby et al., 2005), a neurocomputational model of WM that postulates that both frontal-posterior cortical loops and frontal-BG-thalamic loops are required for maintenance. The frontal-posterior cortical loops are driven by bottom-up visuospatial input whereas the frontal-BG-thalamic loops support the maintenance of this information in WM. FROST does not incorporate any mechanism for learning, but researchers have recently argued for an integration between RL and WM (e.g., Collins & Frank, 2012). Thus, one possibility is that BG-mediated RL facilitates cortico-cortical Hebbian learning between the frontal and posterior cortical components of FROST. Based on the aforementioned work investigating the development of automaticity (e.g., Ashby et al., 2007), the frontal-BG-thalamic loops may no longer be required after extensive training. As a consequence, the visuospatial information would be sufficient to initiate WM for an item, and few WM resources would be required for maintenance. This process could account for a form of long-term stimulus/context-WM item associative learning that, after extensive training, would allow for the memorized items to be used without depleting WM resources. More research is required to directly test this hypothesis.

4.2. *Rule-guided behavior*

The lateral PFC facilitates executive function via its role in the flexible and temporary maintenance of task-relevant information in WM (Wallis, 2007). This information could be stimulus–outcome associations (Wallis & Miller, 2003), rules prescribing the appropriate response to a given stimulus (Ashby et al., 1998), summary representations for a collection of related stimuli (Freedman, 2007), or sets of context-dependent rules/representations (Fuster, 2008). Interestingly, several researchers have suggested that such task-relevant representations follow a rostrocaudal organization that depends upon their abstractness or complexity (Badre, 2008; Bunge & Zelazo, 2006; Christoff, Keramatian, Gordon, Smith, & Mädler, 2009). For example, Bunge and Zelazo (2006) suggest that abstract rules are represented more rostrally, whereas concrete rules are represented more caudally. Hence, univalent rules (i.e., relatively concrete rules where a specific stimulus is always associated with the same response) are usually represented in ventrolateral PFC whereas bivalent rules (i.e., relatively abstract rules where a specific stimulus can be associated with a different response depending on the context) and sets are usually represented in dorsolateral PFC.

Although this account of rule maintenance and rule application focuses on lateral PFC, there is mounting evidence that rule learning is mediated by the interaction between the PFC and the BG (Badre & Frank, 2012; Bunge, 2004; Seger & Miller, 2010). For instance, disruption of PFC-BG networks due to lesion or disease impairs rule learning (Ell, 2013; Ell, Marchant, & Ivry, 2006; Ell, Weinstein, & Ivry, 2010). Moreover, Pasupathy and Miller (2005) showed that activation in both the PFC and the striatum changes when training monkeys in an associative rule-learning task, but that learning-related changes occur earlier in the striatum than in the PFC. Similar results were found by Helie, Roeder, and Ashby (2010) in an fMRI study of (univalent) rule-based categorization. Early in training, BOLD activation in the head of the caudate nucleus was positively correlated with categorization accuracy, and this correlation progressively decreased and became highly negative with extensive training. In contrast, the correlation of BOLD activation in ventrolateral PFC was initially low and increased with practice. Muhammad, Wallis, and Miller (2006) reported similar evidence in a study that recorded from single neurons in the PFC, head of the caudate, and premotor cortex while monkeys were making rule-based responses. In agreement with Helie, Roeder, et al. (2010), they found many neurons in the PFC and caudate that fired selectively to a particular rule. However, after training the animals for a year, they also found many premotor neurons that were rule selective and the degree of rule selectivity was ordered from premotor cortex to PFC to the caudate. Importantly, the premotor rule-selective neurons responded on average about 100 msec before the PFC rule-selective neurons and about 80 msec before the caudate neurons. Thus, after categorization had become automatic, the PFC, although still active, was not mediating response selection. Instead, the single-unit data suggested that the automatic representation included regions of premotor cortex. Similar results have been reported for learning arbitrary stimulus–response associations – that is, the PFC seems to play a critical role in initial learning, but not in the expression of well-learned associations (Puig & Miller, 2012).

Together, these results support the hypothesis that rewarded associations are first learned by the BG, which in turn trains slower-learning mechanisms in the PFC and premotor cortex that encode more abstract rules (i.e., bottom-up

learning; Helie, Proulx, & Lefebvre, 2011; Sun, Merrill, & Peterson, 2001). This bottom-up learning mechanism is further supported by Antzoulatos and Miller (2011), who showed that striatal activation is more related to stimulus–response association learning whereas PFC activation is more closely related to the abstraction of categories. Helie, Roeder, et al. (2010) further suggested that rules are progressively re-coded more caudally in the PFC as they become more 'concrete' with extensive training until they reach the premotor cortex (consistent with Muhammad, Wallis, & Miller, 2006; see also Badre, Kayser, & D'Esposito, 2010).

Although the previous discussion has focused on interactions between lateral PFC and the dorsal striatum, other PFC-BG networks perform a similar function when learning different types of rule-based representations. For instance, orbitofrontal PFC and ventral striatum are involved in the learning (and reversal) of stimulus–outcome associations (Cools, Clark, & Robbins, 2004; Dias, Robbins, & Roberts, 1996; Wallis & Miller, 2003). Stimulus-outcome associations may lead to approach/avoidance behavior that, once automatized, may be mediated solely by cortico-cortical representations linking the conditioned stimulus to the appropriate behavior.

4.3. *Cognitive skill learning*

Although much cognitive activity is based on executive functioning and explicit reasoning, it is well-established that at least some cognitive skills seem to be less dependent upon such controlled processing (Helie, Waldschmidt, & Ashby, 2010). Cognitive skill learning has been defined in various ways, but in this article cognitive skill learning refers to knowledge representations that are incrementally acquired, often in the absence of conscience awareness. These knowledge representations are generally associative, often arising from the gradual accumulation of a number of abstract stimulus/stimuli–response associations.

One example occurs in the information-integration (II) category-learning task, in which stimuli are assigned to categories in such a way that accuracy is maximized only if information from two or more noncommensurable stimulus dimensions is integrated at some pre-decisional stage (Ashby et al., 1998). Typically, the optimal strategy in II tasks is difficult or impossible to describe verbally, which makes it difficult to discover via logical reasoning. Nevertheless, many studies have shown that with enough practice, people reliably learn such categories (e.g., Ashby & Maddox, 2005).

Evidence suggests that II tasks recruit a categorization system that gradually learns to associate high-level visual cortical representations with abstract representations of the motor response in frontal cortex (Ashby et al., 1998, 2007). Critically, this system is thought to depend on BG-based RL. Consistent with this model, learning in II tasks is impaired by many manipulations that would be predicted to disrupt successful RL learning in this network, and it is spared by manipulations intended to disrupt more controlled processing (see Ashby & Maddox, 2005 for a review). For example, delaying corrective feedback (Maddox, Ashby, & Bohil, 2003; Maddox & Ing, 2005) or disrupting the consistency of the category-response mapping (Ashby, Ell, & Waldron, 2003; Maddox, Bohil, & Ing, 2004) impair learning in II tasks. A variety of other cognitive skill learning tasks appear to be dependent upon RL in intact fronto-striatal networks (e.g., Beauchamp, Dagher, Aston, & Doyon, 2003; Knowlton, Mangels, & Squire, 1996; Nagy et al., 2007).

Waldschmidt and Ashby (2011) used fMRI to examine changes in neural activation from initial learning all the way until automaticity developed in II categorization. In this experiment, each human subject trained for more than 10,000 trials on the same II task, with four sessions performed inside the MRI scanner. The results showed that although BG activation (i.e., sensorimotor striatum) predicted categorization performance in early training sessions, only cortical activation was related to categorization performance after extensive training (i.e., the SMA and pre-SMA). Thus, these results are also consistent with the hypothesis that a major role of the BG is to train cortico-cortical connections. Interestingly, although the BG appear to be unnecessary once automaticity has developed, applying such knowledge in a novel context seems to re-recruit the BG (e.g., Wan et al., 2012).

Similar but more direct results have been reported in instrumental conditioning paradigms, which are thought to recruit similar BG circuits as II category learning (Ashby & Crossley, 2011; Ashby et al., 2007). For example, Carelli, Wolske, and West (1997) trained rats for many days to lever press to a tone. The animals learned to lever press reliably within just a few sessions, and their mean response time gradually improved over the course of training. Throughout the extended training period, Carelli et al. (1997) recorded from single units in the striatum. During Session 4, many of these striatal units fired a burst just before the lever press. However, during Sessions 5 and 6, the same striatal units still fired bursts, but now these bursts came after the response had been made, and therefore they could play no role in response selection. In even later sessions, presumably after automaticity was well established, the striatum ceased responding altogether; that is, neither the tone nor the response elicited any activity from the same striatal units that apparently controlled the response earlier in training. Similarly, Choi, Balsam, and Horvitz (2005) reported that DA D1 antagonists disrupted the expression of learned instrumental behaviors early in training, but not after the rats had been over trained in the task.

4.4. *Motor skill learning*

Skills are acquired by experience, not knowledge (Hikosaka, Nakamura, Sakai, & Nakahara, 2002). As mentioned in the introduction of this article, one of the first roles assigned to the BG was motor skill learning and production. Accordingly, almost all models of BG function have included motor skill learning. For instance, Hikosaka et al. (2002) proposed that connectivity between the sensorimotor striatum and premotor cortex allows for the learning of motor sequences, whereas connectivity between the PFC and the associative striatum allows for learning spatial sequences. This model is supported by Miyachi, Hikosaka, Miyashita, Kárádi, and Rand (1997), who showed that muscimol (a GABA agonist) injections into the associative striatum disrupt learning of new sequences (which should initially be spatial) but not

performance on well-learned sequences (which should be motor). In contrast, muscimol injections into the sensorimotor striatum impair the performance of well-learned sequences as well.

Whereas the research of Hikosaka and his colleagues suggest that the BG are important for both motor sequence learning and production (a transfer from anterior to posterior striatum), results from Desmurget and Turner (2010) suggest that the BG contributes to motor execution, but not to motor sequencing or the storage of overlearned sequential skills. Specifically, Desmurget and Turner trained two monkeys in a sequential movement task for an extended period of time. After this training, muscimol was injected in the sensorimotor territory of the GPi. Following the injections, the animals had impaired movements (e.g., dysmetria and slowing of individual movements), but were unimpaired in producing either random or overlearned sequences. As a result, Helie, Roeder, Vucovich, Runger, and Ashby (2014) proposed a neurocomputational model of sequence learning where the BG is used to learn stimulus–response associations and train cortico-cortical connections within the SMA using Hebbian learning to allow for automatic sequence production. Although the BG are required for initial trial-and-error learning in the model, the cortico-cortical connections are sufficient to produce the sequence once it has been automatized. This interpretation is consistent with Poldrack et al. (2005), who observed a decrease in BG activation after extended training in the serial reaction time task. The Helie et al. model has been used to account for behavioral sequence-learning data (Matsuzaka, Picard, & Strick, 2007), transcranial magnetic stimulation (TMS) data (Verwey, Lammens, & van Honk, 2002), single-cell recordings (Shima & Tanji, 2000), and the Desmurget and Turner (2010) data. The model also makes the intriguing prediction that the same muscimol injection used by Desmurget and Turner would impair early sequence learning and production (before the sequence is transferred into cortex). However, this new prediction remains to be tested. Finally, the model could also potentially be used to account for the data collected by Hikosaka and colleagues by assuming that their animals had not practiced enough to produce the sequences automatically.

5. Alternative accounts

The best known alternative to the hypothesis proposed here is that the development of automaticity is mediated by a gradual transfer of control from the associative striatum to the sensorimotor striatum (Belin, Jonkman, Dickinson, Robbins, & Everitt, 2009; Yin & Knowlton, 2006). Much of the evidence in support of this view comes from animal learning studies of habit versus goal-directed behavior, where a variety of evidence implicates the associative striatum in flexible goal-directed behavior and the sensorimotor striatum in rigid habitual behavior. For example, Miyachi, Hikosaka, and Lu (2002) recorded from single units in the striatum while monkeys were learning sequences of arm movements. They reported that neurons in the associative striatum were active during initial skill acquisition and that this activity decreased with extended training. In contrast, most striatal neurons that responded strongly after over-learning the sequence were in the sensorimotor striatum. Similarly, as noted earlier, Miyachi et al. (1997) reported that temporary inactivation of the associative striatum disrupted learning of new motor sequences, but had less effect on the execution of previously-acquired sequences, whereas temporary inactivation of the sensorimotor striatum did not interfere with the learning of new motor sequences, but it did disrupt the execution of previously-acquired sequences. Several fMRI studies have also reported that the associative striatum is active during initial skill acquisition and that its activity decreases with extended training (Lehéricy et al., 2005; Poldrack et al., 2005; Wu, Kansaku, & Hallett, 2004).

Other supporting evidence comes from studies showing that lesions of the associative striatum can speed the transfer of control from the associative to the sensorimotor striatum. For example, rats with lesions of the associative striatum show habit performance even during early stages of training when goal-directed behavior would normally be seen (Yin, Ostlund, Knowlton, & Balleine, 2005). In rats with lesions of the sensorimotor striatum, extended training of a behavior does not lead to a shift from goal-directed action to an S-R habit. In these rats, devaluing the food reward reduces lever-pressing rates even after extended training, suggesting that automaticity is severely delayed or else never occurs (Yin, Knowlton, & Balleine, 2004).

On the other hand, we have reviewed many results that seem inconsistent with the hypothesis that the sensorimotor striatum mediates automatic responding. Perhaps most damaging are results showing that striatal dysfunction (Asmus et al., 2008) or even disconnecting the striatum from motor cortex (Desmurget & Turner, 2010; Doupe, Perkel, Reiner, & Stern, 2005) does not disrupt the expression of automatic behaviors. If the sensorimotor striatum is not mediating automatic responding then what other hypothesis could account for the many results that have been interpreted as evidence supporting this theory? Two alternatives seem plausible. One possibility is that the transfer of control from the associative striatum to the sensorimotor striatum is more accurately described as a switch from a PFC-based declarative memory system to a striatal-based procedural memory system. According to this hypothesis, goal-directed instrumental responding recruits similar neural circuitry as declarative memory tasks, whereas instrumental responding that is under the control of habit learning (i.e., which immediately precedes habitual behavior) depends on much the same circuitry as procedural learning. Some evidence supports this similarity hypothesis. For example, rule-based learning that depends on declarative memory and goal-directed instrumental responding are both flexible with regard to their sensitivity to the nature and timing of feedback (Balleine, 2001; Dickinson & Balleine, 1994; Maddox et al., 2003; Maddox & Ing, 2005). Similarly, procedural learning and habit-based instrumental responding both require extensive practice to acquire, and are both slow to adapt to changing response-outcome contingencies (Ashby et al., 2003; Dickinson, Nicholas, & Adams, 1983; Frankland et al., 2004; Holman, 1975; Maddox et al., 2004). Finally, the striatal region most often implicated in rule-based learning is the head of the caudate nucleus, which is in the associative striatum,

whereas neuroimaging evidence points to the posterior putamen (which is in the sensorimotor striatum) as the striatal region most important for procedural learning (Waldschmidt & Ashby, 2011).

According to this hypothesis, the associative striatum and the sensorimotor striatum are each part of a separate learning system. If so, then the results showing early activation in the associative striatum and later activation in the sensorimotor striatum reflect an early preference for the system that includes the associative striatum over the system that includes the sensorimotor striatum, rather than an early stage of learning that must necessarily precede a later stage. This prediction is supported by results showing that lesions of the associative striatum hasten habit formation (Yin et al., 2005). If the associative striatum played some role that was necessary for habit formation then its lesion should prevent habit formation.

An alternative, but not necessarily contradictory hypothesis is that activation in the sensorimotor striatum after extended overtraining occurs mostly *after* response selection is complete. Most natural environments include sequential regularities. Rewards are often maximized, not by a series of independent and unrelated actions, but by a structured sequence of behaviors in which one movement reliably follows another. Thus, a critical component of skill learning is to form response–response associations. All of premotor cortex projects directly into the sensorimotor striatum, so one possibility is that a significant component of the activation seen in the sensorimotor striatum after extended training is driven by input from premotor cortex that is mediating the learning of motor–motor associations.

There is also evidence supporting this hypothesis. In particular, sensorimotor activation after overtraining is most often reported in studies that used sequence-learning tasks. For example, a variety of fMRI studies have reported that sensorimotor striatal activation increases with extended training in sequence-learning tasks that require button presses (Lehéricy et al., 2005; Poldrack et al., 2005; Wu et al., 2004). In contrast, different results have been reported for tasks that require a single motor response, such as a reinforced lever-press, head-movement, or directed locomotion in a T-maze (Barnes, Kubota, Hu, Jin, & Graybiel, 2005; Carelli et al., 1997; Tang, Pawlak, Prokopenko, & West, 2007). These studies reported that extended training led to a decrease in the number of neurons within the sensorimotor striatum showing task-related activation.

6. Conclusion

This article reviewed the anatomy of the BG and frontal cortex and proposed a general unifying framework to understand the function of frontal-BG connectivity. Specifically, we argue that because learning at cortico-striatal synapses is feedback-driven (RL) while learning at cortico-cortical synapses is not (Hebbian learning), the BG are required for initial trial-and-error learning to ensure that the correct behavior is produced. Meanwhile, cortex learns to reproduce previous behaviors, regardless of whether or not those behaviors are appropriate. BG learning allows for the production of the appropriate behavior, so that what is learned by cortico-cortical synapses is appropriate and useful. Once the cortico-cortical synapses are built, the BG are no longer required to produce the behavior.

Evidence for the role of the BG and frontal cortex in many different functions has been reviewed and re-interpreted according to the new framework. New papers on the role of the BG are published every week, so this review cannot be exhaustive. However, the results are generally compatible with the new interpretation and, as more details are learned about the exact functioning and connectivity of the BG with cortex, some differences in the biological mechanisms supporting these cognitive functions will likely emerge. Yet, the proposed framework highlights a very important commonality that has been observed repeatedly in BG research—namely that BG-cortical loops provide a sufficient circuitry to allow for fast RL-based BG trial-and-error learning that can trigger slow Hebbian learning in cortex for the storage and expression of automatic skills.

Importantly, this theory has possible applications for rehabilitation of patients with neuropsychological conditions. For instance, patients with BG dysfunction (e.g., Parkinson's disease, Huntington's disease, BG lesion) should be impaired in early trial-and-error learning (depending on the BG) but not as much on cortico-cortical learning (depending on cortex). Thus, if techniques less dependent upon trial-and-error learning can be adapted (e.g., errorless learning – Terrace, 1964), then extended training should result in the reinforcement of appropriate cortico-cortical connections. In contrast, patients with frontal lesions should have problems building some cortico-cortical connections (depending on the locus of the lesion) but their initial trial-and-error learning should be mostly spared.

It is our hope that providing an integrative function for the BG may facilitate progress in understanding the role of the BG in many different cognitive and motor tasks. However, most of the reviewed experiments were not specifically designed to test the proposed framework. It is thus possible that alternative explanations could account for all these results. Hence, future research should focus on directly testing the integrative framework using longitudinal designs in tandem with neurophysiological and neuroimaging techniques. One possible experiment is to directly verify that (1) posterior cortex is activated shortly before frontal cortex and (2) that the resulting plasticity is causally related to automaticity.[1] This data would provide for a direct test of the proposed framework. We hope that this review will contribute to the generation such research programs.

Acknowledgments

The authors contributed equally to this work. This research was supported in part by AFOSR grant FA9550-12-1-0355 and by a grant from the U.S. Army Research Office through the Institute for Collaborative Biotechnologies under Grant No. W911NF-07-1-0072.

[1] We thank an anonymous reviewer for suggesting this experiment.

REFERENCES

Akkal, D., Dum, R. P., & Strick, P. L. (2007). Supplementary motor area and presupplementary motor area: targets of basal ganglia and cerebellar output. *Journal of Neuroscience, 27*, 10659–10673.

Alexander, G. E., & DeLong, M. R. (1985). Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. *Journal of Neurophysiology, 53*, 1417–1430.

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.

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

Arbuthnott, G. W., Ingham, C. A., & Wickens, J. R. (2000). Dopamine and synaptic plasticity in the neostriatum. *Journal of Anatomy, 196*, 587–596.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*, 442–481.

Ashby, F. G., & Crossley, M. J. (2011). A computational model of how cholinergic interneurons protect striatal-dependent learning. *Journal of Cognitive Neuroscience, 23*, 1549–1566.

Ashby, F. G., Ell, S. W., Valentin, V., & Casale, M. B. (2005). FROST: a distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience, 17*, 1728–1743.

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition, 31*, 1114–1125.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review, 114*, 632–656.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56*, 149–178.

Asmus, F., Huber, H., Gasser, T., & Schöls, L. (2008). Kick and rush: paradoxical kinesia in Parkinson disease. *Neurology, 71*, 695.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences, 12*, 193–200.

Badre, D., & Frank, M. J. (2012). Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from FMRI. *Cerebral Cortex, 22*, 527–536.

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron, 66*, 315–326.

Balleine, B. W. (2001). Incentive processes in instrumental conditioning. In R. R. Mowrer, & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 307–366). Mahwah, NJ: Lawrence Erlbaum Associates.

Barbas, H., Henion, T. H. H., & Dermon, C. R. (1991). Diverse thalamic projections to the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology, 313*, 65–94.

Barbas, H., & Pandya, D. N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *Journal of Comparative Neurology, 256*, 211–228.

Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature, 437*, 1158–1161.

Beauchamp, M. H., Dagher, A., Aston, J. A., & Doyon, J. (2003). Dynamic functional changes associated with cognitive skill learning of an adapted version of the Tower of London task. *NeuroImage, 20*, 1649–1660.

Belin, D., Jonkman, S., Dickinson, A., Robbins, T. W., & Everitt, B. J. (2009). Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behavioral Brain Research, 199*, 89–102.

Bolam, J. P., Bergman, H., Graybiel, A. M., Kimura, M., Plenz, D., Seung, S. S., et al. (2006). Microcircuits, molecules, and motivated behavior—microcircuits in the striatum. In S. Grillner, & A. M. Graybiel (Eds.), *Microcircuits: The interface between neurons and global brain function* (pp. 165–190). MIT Press.

Bolam, J. P., Powell, J. F., Wu, J. Y., & Smith, A. D. (1985). Glutamate decarboxylase-immunoreactive structures in the rat neostriatum: a correlated light and electron microscopic study including a combination of Golgi impregnation with immunocytochemistry. *Journal of Comparative Neurology, 237*, 1–20.

Buchsbaum, B., & D'Esposito, M. (2008). Short-term and working memory systems. In H. Eichenbaum (Ed.), *Memory systems of learning and memory: A comprehensive reference* (Vol. 3, pp. 237–260). Oxford: Elsevier.

Bunge, S. A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cognitive, Affective & Behavioral Neuroscience, 4*, 564–579.

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15*, 118–121.

Cairo, T. A., Liddle, P. F., Woodward, T. S., & Ngan, E. T. (2004). The influence of working memory load on phase specific patterns of cortical activity. *Brain Research. Cognitive Brain Research, 21*, 377–387.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., et al. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex, 9*, 20–46.

Calzavara, R., Mailly, P., & Haber, S. N. (2007). Relationship between the corticostriatal terminals from areas 9 and 46, and those from area 8A, dorsal and rostral premotor cortex and area 24c: an anatomical substrate for cognition to action. *European Journal of Neuroscience, 26*, 2005–2024.

Carelli, R. M., Wolske, M., & West, M. O. (1997). Loss of lever press related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *Journal of Neuroscience, 17*, 1804–1814.

Chang, C., Crottaz-Herbette, S., & Menon, V. (2007). Temporal dynamics of basal ganglia response and connectivity during verbal working memory. *NeuroImage, 34*, 1253–1269.

Choi, W. Y., Balsam, P. D., & Horvitz, J. C. (2005). Extended habit training reduces dopamine mediation of appetitive response expression. *Journal of Neuroscience, 25*, 6729–6733.

Christoff, K., Keramatian, K., Gordon, A. M., Smith, R., & Mädler, B. (2009). Prefrontal organization of cognitive control according to levels of abstraction. *Brain Research, 1286*, 94–105.

Chudasama, Y., & Robbins, T. W. (2006). Functions of fronto-striatal systems in cognition: comparative neuropsychopharmacological studies in rats, monkeys, and humans. *Biological Psychology, 73*, 19–38.

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience, 35*, 1024–1035.

Cools, R., Clark, L., & Robbins, T. W. (2004). Differential responses in human striatum and prefrontal cortex to changes in object and rule relevance. *Journal of Neuroscience, 24*, 1129–1135.

Cools, R., Sheridan, M., Jacobs, E., & D'Esposito, M. (2007). Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *Journal of Neuroscience, 27*, 5506–5514.

Desmurget, M., & Turner, R. S. (2010). Motor sequences and the basal ganglia: kinematics, not habits. *Journal of Neuroscience, 30*, 7685–7690.

D'Esposito, M., Aguirre, G. K., Zarahn, E., Ballard, D., Shin, R. K., & Lease, J. (1998). Functional MRI studies of spatial and nonspatial working memory. *Cognitive Brain Research, 7*, 1–13.

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Learning & Behavior, 22*, 1–18.
Dickinson, A., Nicholas, D., & Adams, C. D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 35B*, 35–51.
Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M., Uddin, L. Q., Shehzad, Z., et al. (2008). Functional connectivity of human striatum: a resting state FMRI study. *Cerebral Cortex, 18*, 2735–2747.
Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*, 69–72.
Doupe, A. J., Perkel, D. J., Reiner, A., & Stern, E. A. (2005). Birdbrains could teach basal ganglia research a new song. *Trends in Neurosciences, 28*, 353–363.
Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10*, 732–739.
Draganski, B., Kherif, F., Kloppel, S., Cook, P. A., Alexander, D. C., Parker, G. J., et al. (2008). Evidence for segregated and integrative connectivity patterns in the human Basal Ganglia. *Journal of Neuroscience, 28*, 7143–7152.
Dum, R. P., & Strick, P. L. (1991). The origin of corticospinal projections from the premotor areas in the frontal lobe. *Journal of Neuroscience, 11*, 667–689.
Ell, S. W. (2013). Targeted training of the decision rule benefits rule-guided behavior in Parkinson's disease. *Cognitive, Affective, and Behavioral Neuroscience, 13*, 830–846.
Ell, S. W., Hélie, S., & Hutchinson, S. (2012). Contributions of the putamen to cognitive function. In A. Costa, & E. Villalba (Eds.), *Horizon in neuroscience* (Vol. 7, pp. 29–52). Nova Publishers.
Ell, S. W., Marchant, N. L., & Ivry, R. B. (2006). Focal putamen lesions impair learning in rule-based, but not information-integration categorization tasks. *Neuropsychologia, 44*, 1737–1751.
Ell, S. W., Weinstein, A., & Ivry, R. B. (2010). Rule-based categorization deficits in focal basal ganglia lesion and Parkinson's disease patients. *Neuropsychologia, 48*, 2974–2986.
Feenstra, M. G., & Botterblom, M. H. (1996). Rapid sampling of extracellular dopamine in the rat prefrontal cortex during food consumption, handling and exposure to novelty. *Brain Research, 742*, 17–24.
Feldman, D. E. (2009). Synaptic mechanisms for plasticity in neocortex. *Annual Review of Neuroscience, 32*, 33–55.
Flaherty, A. W., & Graybiel, A. M. (1994). Input-output organization of the sensorimotor striatum in the squirrel monkey. *Journal of Neuroscience, 14*, 599–610.
Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory. A computational model. *Cognitive, Affective, & Behavioral Neuroscience, 1*, 137–160.
Frankland, P., Wang, Y., Rosner, B., Shimizu, T., Balleine, B., Dykens, E., et al. (2004). Sensorimotor gating abnormalities in young males with fragile X syndrome and Fmr1-knockout mice. *Molecular Psychiatry, 9*, 417–425.
Freedman, D. (2007). Exploring the roles of the frontal, temporal, and parietal lobes in visual categorization. In S. A. Bunge, & J. D. Wallis (Eds.), *Neuroscience of rule-guided behavior* (pp. 391–418). Oxford University Press.
Fuster, J. M. (2008). *The prefrontal cortex* (4th ed.). Singapore, CN: Academic Press.
Gerfen, C. R., & Bolam, J. P. (2010). The neuroanatomical organization of the basal ganglia. In H. Steiner, & K. Y. Tseng (Eds.), *Handbook of basal ganglia structure and function* (pp. 3–28). Elsevier.
Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology: The nervous system* (Vol. 5, pp. 373–417). Bethesda, MD: American Physiological Society.
Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy, 26*, 317–330.
Haber, S. N., Kim, K. S., Mailly, P., & Calzavara, R. (2006). Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *Journal of Neuroscience, 26*, 8368–8376.
Helie, S., Chakravarthy, S., & Moustafa, A. A. (2013). Exploring the cognitive and motor functions of the basal ganglia: an integrative review of computational cognitive neuroscience models. *Frontiers in Computational Neuroscience, 7*, 174.
Hélie, S., Paul, E. J., & Ashby, F. G. (2012a). A neurocomputational account of cognitive deficits in Parkinson's disease. *Neuropsychologia, 50*, 2290–2302.
Hélie, S., Paul, E. J., & Ashby, F. G. (2012b). Simulating the effect of dopamine imbalance on cognition: from positive affect to Parkinson's disease. *Neural Networks, 32*, 74–85.
Helie, S., Proulx, R., & Lefebvre, B. (2011). Bottom-up learning of explicit knowledge using a Bayesian algorithm and a new Hebbian learning rule. *Neural Networks, 24*, 219–232.
Helie, S., Roeder, J., Vucovich, L., Runger, D., & Ashby, F. G. (July 2014). Automatic sequence production in the brain: a neurocomputational model. In *47th Annual Meeting of the Society for Mathematical Psychology. Quebec City, Quebec, Canada*.
Helie, S., Roeder, J. L., & Ashby, F. G. (2010a). Evidence for cortical automaticity in rule-based categorization. *Journal of Neuroscience, 30*, 14225–14234.
Helie, S., Waldschmidt, J. G., & Ashby, F. G. (2010b). Automaticity in rule-based and information-integration categorization. *Attention, Perception, & Psychophysics, 72*, 1013–1031.
Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology, 12*, 217–222.
Holman, E. W. (1975). Some conditions for the dissociation of consummatory and instrumental behavior in rats. *Learning & Motivation, 6*, 358–366.
Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, MA: MIT Press.
Joel, D., & Weiner, I. (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience, 96*, 451–474.
Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9*, 462–475.
Karreman, M., & Moghaddam, B. (1996). The prefrontal cortex regulates the basal release of dopamine in the limbic striatum: an effect mediated by ventral tegmental area. *Journal of Neurochemistry, 66*, 589–598.
Kemp, J. M., & Powell, T. P. (1971). The structure of the caudate nucleus of the cat: light and electron microscopy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 262*, 383–401.
Kemp, J. M., & Powell, T. P. S. (1970). The cortico-striate projections in the monkey. *Brain, 93*, 525–546.
Kim, U., Sanchez-Vives, M. V., & McCormick, D. A. (1997). Functional dynamics of GABAergic inhibition in the thalamus. *Science, 278*, 130–134.
Kincaid, A. E., Zheng, T., & Wilson, C. J. (1998). Connectivity and convergence of single corticostriatal axons. *Journal of Neuroscience, 18*, 4722–4731.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273*, 1399–1402.
Lapish, C. C., Kroener, S., Durstewitz, D., Lavin, A., & Seamans, J. K. (2007). The ability of the mesocortical dopamine system to operate in distinct temporal modes. *Psychopharmacology, 191*, 609–625.
Lehéricy, S., Benali, H., Van de Moortele, P. F., Pélégrini-Issac, M., Waechter, T., Ugurbil, K., et al. (2005). Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 12566–12571.
Lu, M.-T., Preston, J. B., & Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *The Journal of Comparative Neurology, 341*, 375–392.
Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 29*, 650–662.
Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural-learning-based system in perceptual category learning. *Psychonomic Bulletin & Review, 11*, 945–952.
Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31*, 100–107.
Matsuzaka, Y., Picard, N., & Strick, P. L. (2007). Skill representation in the primary motor cortex after long term practice. *Journal of Neurophysiology, 97*, 1819–1832.
Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.
Miyachi, S., Hikosaka, O., & Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Experimental Brain Research, 146*, 122–126.
Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z., & Rand, M. K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. *Experimental Brain Research, 115*, 1–5.
Monchi, O., Taylor, J. G., & Dagher, A. (2000). A neural model of working memory processes in normal subjects, Parkinson's disease, and schizophrenia for fMRI design and predictions. *Neural Networks, 13*, 953–973.
Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*, 1936–1947.
Muhammad, R., Wallis, J. D., & Miller, E. K. (2006). A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *Journal of Cognitive Neuroscience, 18*, 974–989.
Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews: Neuroscience, 9*, 856–869.
Nagy, H., Keri, S., Myers, C. E., Benedek, G., Shohamy, D., & Gluck, M. A. (2007). Cognitive sequence learning in Parkinson's disease and amnestic mild cognitive impairment: dissociation between sequential and non-sequential learning of associations. *Neuropsychologia, 45*, 1386–1392.
Osaka, M., & Osaka, N. (2007). Neural bases of focusing attention in working memory: an fMRI study based on group differences. In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 99–117). New York: Oxford University Press.
Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience, 25*, 563–593.
Parent, A., & Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Research Reviews, 20*, 91–127.
Pasupathy, A., & Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature, 433*, 873–876.
Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., et al. (2005). The neural correlates of motor skill automaticity. *Journal of Neuroscience, 25*, 5356–5364.
Postuma, R. B., & Dagher, A. (2006). Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cerebral Cortex, 16*, 1508–1521.
Puig, M., & Miller, E. K. (2012). The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron, 74*, 874–886.
Reiner, A. (2010). Organization of corticostriatal projection neuron types. In H. Steiner, & K. Y. Tseng (Eds.), *Handbook of basal ganglia structure and function* (pp. 323–340). Elsevier.
Ronesi, J., & Lovinger, D. M. (2005). Induction of striatal long-term synaptic depression by moderate frequency activation of cortical afferents in rats. *Journal of Physiology, 562*, 245–256.
Schroll, H., Vitay, J., & Hamker, F. H. (2012). Working memory and response selection: a computational account of interactions among cortico-basalganglio-thalamic loops. *Neural Networks, 26*, 59–74.
Seamans, J. K., & Robbins, T. W. (2009). Dopamine modulation of prefrontal cortex and cognitive function. In K. A. Neve (Ed.), *The dopamine receptors* (2nd ed.). (pp. 373–398). Springer.
Seger, C. A. (2006). The basal ganglia in human learning. *Neuroscientist, 12*, 205–290.
Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual Review of Neuroscience, 33*, 203–219.
Selemon, L. D., & Goldman-Rakic, P. S. (1985). Longitudinal topography and interdigitation of cortico-striatal projections in the rhesus monkey. *Journal of Neuroscience, 5*, 776–794.
Shima, K., & Tanji, J. (2000). Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *Journal of Neurophysiology, 84*, 2148–2160.
Soliveri, P., Brown, R. G., Jahanshahi, M., Caraceni, T., & Marsden, C. D. (1997). Learning manual pursuit tracking skills in patients with Parkinson's disease. *Brain, 120*, 1325–1337.
Sun, R., Merrill, E., & Peterson, T. (2001). From implicit to explicit knowledge: a bottom-up model of skill learning. *Cognitive Science, 25*, 203–244.
Tang, C., Pawlak, A. P., Prokopenko, V., & West, M. O. (2007). Changes in activity of the striatum during formation of a motor habit. *European Journal of Neuroscience, 25*, 1212–1227.
Taylor, J. G., & Taylor, N. R. (2000). Analysis of recurrent cortico-basal ganglia-thalamic loops for working memory. *Biological Cybernetics, 82*, 415–432.
Terrace, H. S. (1964). Wavelength generalization after discrimination with and without errors. *Science, 144*, 78–80.
Thomas-Ollivier, V., Reymann, J. M., Le Moal, S., Schueck, S., Lieury, A., & Allain, H. (1999). Procedural memory in recent onset Parkinson's disease. *Dementia and Geriatric Cognitive Disorders, 10*, 172–180.
Van Hoesen, G. W., Yeterian, E. H., & Lavizzo-Mourey, R. (1981). Widespread corticostriate projections from temporal cortex of the rhesus monkey. *Journal of Comparative Neurology, 199*, 205–219.
Van der Werf, Y. D., Witter, M. P., Uylings, H. B. M., & Jolles, J. (2000). Neuropsychology of infarctions in the thalamus: a review. *Neuropsychologia, 38*, 613–627.
Verwey, W. B., Lammens, R., & van Honk, J. (2002). On the role of the SMA in the discrete sequence production task: a TMS study. *Neuropsychologia, 40*, 1268–1276.

Vogt, C. (1911). Quelques considerations generales sur le syndrome du corps strie. *Journal fur Psychologie und Neurologie, 18*, 479–488.
Voytek, B., & Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. *Proceedings of the National Academy of Sciences, 107*, 18167–18172.
Wager, T., & Smith, E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience, 3*, 255–274.
Waldschmidt, J. G., & Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *NeuroImage, 56*, 1791–1802.
Wallis, J. D. (2007). Single neuron activity underlying behavior-guiding rules. In S. A. Bunge, & J. D. Wallis (Eds.), *Neuroscience of rule-guided behavior* (pp. 23–44). Oxford University Press.
Wallis, J. D., & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience, 18*, 2069–2081.
Wan, X., Takano, D., Asamizuya, T., Suzuki, C., Ueno, K., Cheng, K., et al. (2012). Developing intuition: neural correlates of cognitive skill learning in caudate nucleus. *Journal of Neuroscience, 32*, 17492–17501.
Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1993). Subcortical connections of inferior temporal areas TE and TEO in macaque monkeys. *Journal of Comparative Neurology, 335*, 73–91.
Wickens, J., & Arbuthnott, G. W. (2010). Gating of the cortical input to the striatum. In H. Steiner, & K. Y. Tseng (Eds.), *Handbook of basal ganglia structure and function* (pp. 341–352). Elsevier.
Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 29–50). MIT Press.
Wilson, F. A., Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260*, 1955–1958.
Wilson, S. A. K. (1912). Progressive lenticular degeneration: a familiar nervous disease associated with cirrhosis of the liver. *Brain, 34*, 295–509.
Wise, S., Murray, E., & Gerfen, C. (1996). The frontal cortex-basal ganglia system in primates. *Critical Reviews in Neurobiology, 10*, 317–356.
Wu, T., Kansaku, K., & Hallett, M. (2004). How self-initiated memorized movements become automatic: a functional MRI study. *Journal of Neurophysiology, 91*, 1690–1698.
Yeterian, E. H., & Pandya, D. N. (1993). Striatal connections of the parietal association cortices in rhesus monkeys. *Journal of Comparative Neurology, 332*, 175–197.
Yeterian, E. H., & Pandya, D. N. (1995). Corticostriatal connections of extrastriate visual areas in rhesus monkeys. *Journal of Comparative Neurology, 352*, 436–457.
Yeterian, E. H., & Pandya, D. N. (1998). Corticostriatal connections of the superior temporal region in rhesus monkeys. *Journal of Comparative Neurology, 399*, 384–402.
Yeterian, E. H., & Van Hoesen, G. W. (1978). Cortico-striate projections in the rhesus monkey: the organization of certain cortico-caudate connections. *Brain Research, 139*, 43–63.
Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience, 7*, 464–476.
Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience, 19*, 181–189.
Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience, 22*, 513–523.