5 Integrating what & how/where with instrumental and Pavlovian learning

A biologically based computational model

Wolfgang M. Pauli, Hisham E. Atallah, and Randall C. O'Reilly

Introduction

Skill learning and instrumental conditioning have a long association with the behaviorist research program and are often thought of in terms of rote learning of simple stimulus–response (S–R) associations based on external rewards. These forms of learning have also been strongly associated with the striatum of the basal ganglia, with different striatal regions apparently making various different contributions (Atallah, Frank, & O'Reilly, 2004; Featherstone & McDonald, 2004; Graybiel, 1998; Kantak, Green-Jordan, Valencia, Kremin, & Eichenbaum, 2001; O'Doherty, 2004; Packard, Hirsh, & White, 1989; Packard & McGaugh, 1992; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999; Williams & Eskandar, 2006). In the influential multiple memory systems framework of Squire (1992) for example, the basal ganglia are associated with procedural "habit" learning. However, both aspects of this dogma are currently being challenged: in the early phase of learning instrumental conditioning actually requires considerable high-level cognitive function to discover which actions lead to reward delivery in novel situations, and this early-phase high-level cognitive function depends critically on the basal ganglia, whereas the cortex is more likely the site of more rote longer-term habit learning (Atallah et al., 2004; Frank, 2005; Houk & Wise, 1995; Loh, Pasupathy, Miller, & Deco, 2008; Pasupathy & Miller, 2005). Here, we describe a biologically based computational model that demonstrates how different regions of the basal ganglia and prefrontal cortex contribute to this early learning.

We focus our model on a specific set of recent data demonstrating striking dissociations in the contributions of the ventral striatum (VS) and dorsolateral striatum (DS) in a two-alternative forced-choice (2AFC) task (Atallah, Lopez-Paniagua, Rudy, & O'Reilly, 2006). Systematic inactivation of VS or DS during or after task acquisition (see details below) lead Atallah et al. (2006) to the conclusion that the VS is essential for task acquisition but not for performance after task acquisition, and that the DS is important for the

Perception, attention, and action:
International Perspectives on Psychological Science (Volume1). Peter A. Frensch and Ralf Schwarzer (Eds). 2010.
Published by Psychology Press on behalf of the International Union of Psychological Science.

expression of learned behavior. Atallah et al. (2006) argued that the VS acts as a director system that controls an actor system that depends on the DS as an extension to the often-discussed actor–critic architecture (Barto, Sutton, & Anderson, 1983; Dayan, Niv, Seymour, & Daw, 2006; O'Doherty et al., 2004; Sutton & Barto, 1998). Our *WHIP* model (What-How Instrumental & Pavlovian) captures this director–actor distinction (though with some important modifications), and in making the mechanisms explicit, it provides insight into the complex interactions between neural systems that support the learning of new behavioral tasks such as this one. Specifically, the model shows that novel behavior emerges from the interaction of separable "What" and "How/Where" processing streams in the ventral and dorsal cortical pathways (Goodale & Milner, 1992) and their associated striatal areas, together with the recruitment and modification of a set of basic response strategies (e.g., *approach* and *avoid*). At the most abstract level, novel behavior in this model emerges through combinations of existing neural subcomponents, which are then shaped through reinforcement contingencies to adapt to the situation at hand.

It would be relatively simple to capture in an abstract model the basic director–actor distinction. For example, one could have the director system learn as the actor typically does in standard actor–critic models, and this director simply sends its outputs through an actor system, which is responsible for driving behavior. In the absence of director inputs, behavior is random. This would capture the core qualitative finding from Atallah et al. (2006), but it would raise more questions than it answers: Why have a separate actor system in the first place? Why is the ventral striatum a director to the dorsal striatum's actor? What are the striatal areas doing as compared to the frontal cortex and other brain areas? How does this fit within a larger picture of the differential contributions of prefrontal cortex, basal ganglia, and other brain areas on other kinds of learning tasks?

To explore potential answers to these more complex questions, we focus on the division of labor between a ventral "What" pathway and a dorsal "How/Where" pathway within an established framework called PBWM (prefrontal-cortex, basal ganglia working memory model), which has been used in a variety of contexts for modeling prefrontal cortex and basal ganglia contributions to working memory, action selection, and other cognitive functions (Hazy, Frank, & O'Reilly, 2006, 2007; O'Reilly & Frank, 2006). In the following sections, we describe the core principles behind this model and their specific application to the WHIP model.

The what–how instrumental Pavlovian model

Dynamic gating in frontal areas by basal ganglia

The central tenet of the PBWM model is that basal ganglia (BG) provide an adaptive, dynamic gating signal for controlling the active maintenance,

updating, and output of information in frontal cortex. The basal ganglia (BG) layers are interconnected with frontal cortex through a series of parallel loops (Alexander, DeLong, & Strick, 1986). These loops enable the basal ganglia to exert a gating-like modulation of representations in frontal areas (see Figure 5.1). This kind of gating mechanism is consistent with a wide range of empirical data, and similar implementations of dynamic gating were included in previous computational models (e.g., (e.g. Berns & Sejnowski, 1998; Cisek, 2007; Dominey, Arbib, & Joseph, 1995; Gurney, Prescott, & Redgrave, 2001; Houk et al., 2007; Houk & Wise, 1995; Humphries, Stewart, & Gurney, 2006; Wickens, Kotter, & Alexander, 1995). The PBWM framework used in the present model includes a biologically plausible implementation of this gating mechanism (Hazy et al., 2006, 2007; O'Reilly & Frank, 2006).

The current model uses one important feature that has been added since the original publication (O'Reilly & Frank, 2006): *output gating*. The original

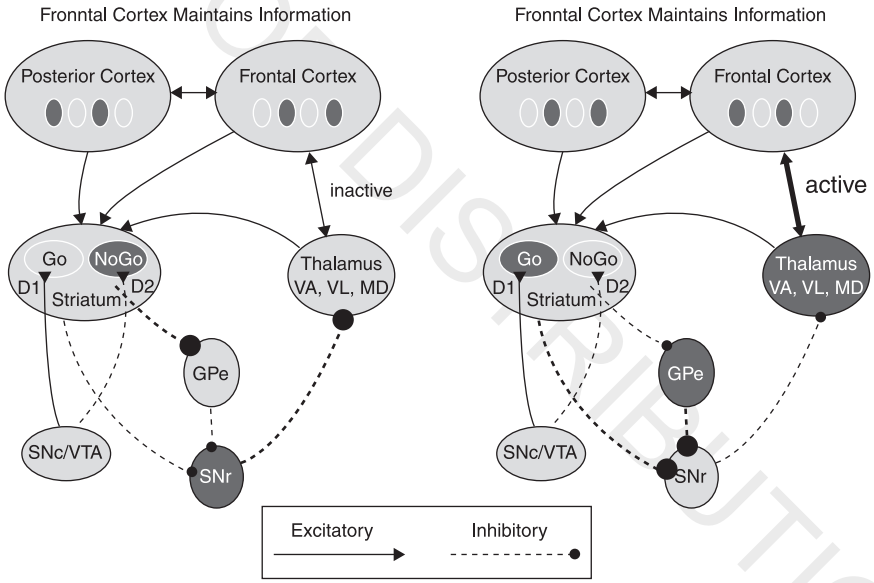

Figure 5.1 The BG are interconnected with frontal cortex through a series of parallel loops each of the form shown. Working backward from the thalamus, which is bidirectionally excitatory with frontal cortex, the SNr (substantia nigra pars reticulata) is tonically active and inhibiting this excitatory circuit. When direct pathway Go neurons in DS fire, they inhibit the SNr and thus disinhibit frontal cortex producing a gating-like modulation that, we argue, triggers the update of working memory representations in PFC. The indirect pathway NoGo neurons of DS counteract this effect by inhibiting the inhibitory GPe (globus pallidus, external segment). The STN (subthalamic nucleus) provides an additional dynamic background of inhibition (NoGo) by exciting the SNr (Frank, 2006, Frank et al., 2007, Hazy et al., 2007).

PBWM model only included a gating effect on the maintenance of information in simulated PFC areas. However, considerable evidence summarized in Brown, Bullock, and Grossberg (2004) suggests that the BG can also modulate the output (deep) layers of frontal areas, and that the neurons that perform this modulation are distinct from those that modulate the maintenance of information, which takes place in the superficial PFC layers (II & III). In the case of maintenance gating, a Go signal from the BG allows the rapid updating of working memory representations in frontal cortex by interrupting the maintenance currents of currently active representations. By contrast, output gating is thought to result from extra BG-driven activation enabling the deep-layer neurons in frontal cortex to get over their firing thresholds, and thus communicate information to other cortical areas and subcortical targets. This is the same mechanism leading to motor gating, which is typically described as the main function of BG interactions with motor cortex (Frank, 2005; Gurney et al., 2001; Houk, Adams, & Barto, 1995; Mink, 1996). In both forms of gating, striatal NoGo neurons of the inhibitory pathway counteract Go signals. In addition, the subthalamic nucleus (STN) provides a global dynamic background of inhibition (NoGo) by exciting the SNr (Frank, 2006; Frank, Samanta, Moustafa, & Sherman, 2007; Hazy et al., 2007). (For a graphical depiction and more details, see Figure 5.1).

Learning what and when to gate

For the above gating mechanism to work successfully, the BG have to learn when to gate what information "in" frontal areas for active maintenance and "out" of frontal areas in the case of output gating. This learning is dopamine-based and allows each striatal projection neuron (medium spiny neuron, MSN) to develop its own unique pattern of input weights that determine its actions. Dopamine release in the striatum of our model is determined by two different mechanisms: projections from the dopaminergic neurons of the SNc/VTA (Substantia Nigra pars compacta; ventral tegmental area), captured by the PVLV model (Primary Value, Learned Value; Hazy et al., 2007; O'Reilly & Frank, 2006; O'Reilly, Frank, Hazy, & Watz, 2007), and projections from the basolateral complex of amygdala (BLA).

The midbrain dopamine neurons in the SNc/VTA of the mammalian brain are driven by inputs from the central nucleus of the amygdala (CNA), the lateral hypothalamus (LH), and the patch-like neurons of the ventral striatum (Ahn & Phillips, 2003; Floresco, West, Ash, Moore, & Grace, 2003; Fudge & Haber, 2000; Joel & Weiner, 2000; Rouillard & Freeman, 1997; Semba & Fibiger, 1992). The contributions of these inputs are described by the PVLV model (Hazy et al., 2007; O'Reilly & Frank, 2006; O'Reilly et al., 2007) as follows. The LH (PVe layer) delivers primary reward information, and striosome/patc neurons of the ventral striatum/nucleus accumbens (PVi layer) learn to expect such rewards, and thereby block the dopamine spike that would otherwise occur to them. This is the Primary Value (PV) system of

PVLV. The Learned Value (LV) system, involving the CNA, is important for learning reward associations for conditioned stimuli (CSs), which can then drive dopamine firing at the time of CS onset. These two interacting systems provide a good account of the extant neural recording data from the SNc (Schultz, 1998; Schultz, Apicella, & Ljungberg, 1993). The PVLV algorithm is an alternative to the temporal-differences (TD) algorithm (Sutton, 1988; Sutton & Barto, 1998) but is more directly related to the underlying biology (Hazy et al., 2007; O'Reilly et al., 2007).

The functional contribution of the PVLV system is to provide positive dopamine bursts for successful behavior and CSs associated therewith, and negative dopamine dips for unsuccessful behavior and associated CSs. The positive dopamine bursts cause Go pathway neurons in the striatum to become more active (due to a preponderance of dopamine D1 receptors, which are excitatory) and NoGo pathway neurons to become less active (from D2 receptors, which are inhibitory) (Frank, 2005; Frank, Seeberger, & O'Reilly, 2004). The opposite case holds for negative dopamine dips. This bidirectional plasticity shapes the gating firing in ways that lead to successful learning of complex working memory tasks in the PBWM model (Hazy et al., 2006, 2007; O'Reilly & Frank, 2006).

In addition to SNc phasic dopamine, the BLA can influence learning in the VS. The BLA (like the CNA) learns to associate stimuli with positive/negative valence representations, with individual neurons having stable valence coding to which stimuli get mapped (Murray, 2007; Schoenbaum, Chiba, & Gallagher, 1999). In contrast to the CNA, the BLA sends dense glutamatergic projections to the VS that synapse onto MSNs there. It has also been shown that BLA activation can produce localized dopamine release in the VS due to collateral projections that synapse directly onto terminal boutons of midbrain DA cells. This later effect can occur without benefit of midbrain DA cell firing (Floresco, Yang, Phillips, & Blaha, 1998; Johnson, Aylward, Hussain, & Totterdell, 1995). In our model, this BLA contribution to learning is important for enabling the VS to learn even when the dorsal pathway lesion disrupts the behavioral contingency on PVLV-mediated reward values. These differences between BLA- and PVLV-driven DA release in the model have important functional implications that are discussed below.

Division of labor between cortical pathways

The WHIP model (Figure 5.2) includes a highly simplified version of the what and how/where dissociation of ventral and dorsal processing streams (Goodale & Milner, 1992; Ungerleider & Mishkin, 1982), which then carries over into the VS and DS and prefrontal cortex areas (Toni, Rowe, Stephan, & Passingham, 2002).[1] Ventral areas are concerned with processing stimulus information and their reward associations (Vogels, 1999), whereas dorsal areas are concerned with spatial location information, and more generally in

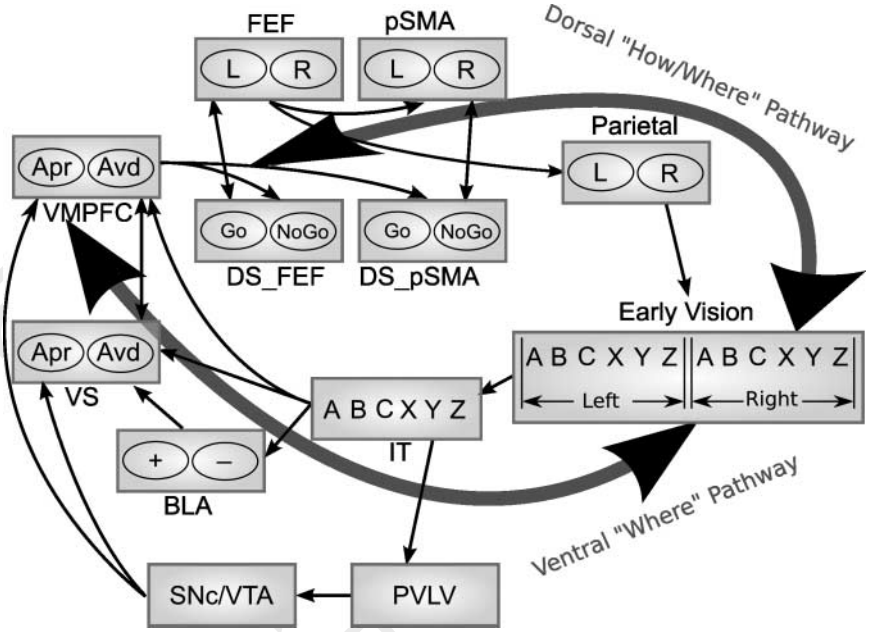

Figure 5.2 Abstract version of the model, showing flow of activation from initial sensory input (Early Vision) along two separate pathways: the ventral "What" pathway extracts identity information and associates that with reward values and overall action plans (approach vs. avoid). The dorsal "How/Where" pathway extracts location and other perception-for-action information (only location is relevant in this model) and is necessary for executing appropriate motor responses (e.g., knowing which side of the Y maze to approach). Abbreviations: BLA, basolateral complex of amygdala; DS, dorsal striatum; FEF, Frontal Eye Field; IT, inferiotemporal cortex; pSMA, pre-supplementary motor area; PVLV, primary value learned value algorithm; SNc/VTA, substantia nigra pars compacta/ventral tegmental area; VMPFC, ventromedial prefrontal cortex; VS, ventral striatum;

perception-for-action (Goodale & Milner, 1992), and in dorsal frontal areas, actually organizing responses.

At a general level, the WHIP model simulates the 2AFC task used by Atallah et al. (2006) in terms of a stimulus sampling phase followed by a decision to either approach the sampled stimulus, or avoid it and switch to sampling the other stimulus. Thus, the model samples one of the stimuli (e.g., A) at a time, and if it decides to approach, then it receives the reward associated with that stimulus (a reward for approaching A, and no reward for approaching B). If it decides to switch, then on the next trial it will likely sample the other stimulus, and either decide to approach, or to switch again.

The various systems in the **ventral/what** pathway in the WHIP model are involved in:

- encoding the sampled stimulus, but *not* its location (in a simulated inferotemporal (IT) cortex)
- learning whether this stimulus is associated with reward (in simulated VS (patches/PVi) and amygdala, including the basal-lateral amygdala (BLA) and central nucleus of the amygdala (CNA))
- learning what kind of generalized instrumental behavior (approach or avoid) to take in relation to the sampled stimulus (in simulated VS areas interacting with ventromedial prefrontal areas (VMPFC).

In contrast, the **dorsal/how/where** pathway systems are involved in:

- deciding which spatial location to sample – via interactions between simulated parietal and frontal eye field (FEF) areas.
- initiating and guiding the approach or avoid motor programs – in simulated pre-supplementary motor areas (pSMA).

Across both pathways, a common division of labor between the basal ganglia and associated frontal cortical areas is implemented, consistent with the PBWM model. The basal ganglia is important for rapid initial task learning driven by strong dopaminergic modulation, and it serves to provide *Go* and *NoGo* signals to frontal areas that control the initiation of motor plans or the updating of information in working memory. Thus, this model correctly predicts that striatal areas are critical early in training, but cortical learning can slowly consolidate habit-like task behavior that is increasingly independent of the basal ganglia.

There is considerable evidence in support of these functional specializations associated with the corresponding brain areas, as we review below. Given such a division of labor, accurate performance of this task requires the integration of ventral pathway information about which stimulus to approach and which to avoid, with dorsal pathway information about where the stimuli are, and the actual motor plans to carry out those motor programs. In this sense, the ventral system plays the director to the dorsal system's actor. But the actor system is contributing critical new information to the problem, in the form of knowledge of where the sampled stimulus is actually located and how to properly approach it. Thus, when the dorsal system is inactivated, the system still knows which stimulus is the good one, but it cannot integrate that knowledge with the stimulus's location representations to produce a sensible response. This is reminiscent of similar dissociations between object identification and action production that are produced with differential lesions to the dorsal and ventral visual pathways (Fiehler, Burke, Bien, Röder, & Rösler, 2008; Goodale & Milner, 1992; Himmelbach & Karnath, 2005; James, Culham, Humphrey, Milner, & Goodale, 2003). Not only does this what/where perspective provide a clear explanation for the differential roles of ventral and dorsal striatum in the current context, it also suggests conditions in which these roles can differ, providing a number of testable predictions enumerated below.

Habit formation in frontal cortex

In contrast to the existing proposal that associates basal ganglia with procedural habit learning and cortex with problem-solving in early phases of learning (e.g., Squire, 1992), the WHIP model demonstrates how the basal ganglia might contribute to early problem-solving (via biasing activation in frontal areas) in a simple choice task, while longer-term habit learning occurs in frontal cortical areas. Very generally, the model exhibits a wave of plasticity propagating through the cortico-striatal loops. First, the midbrain dopamine system and amygdala nuclei learn about the valence of stimuli, then these systems determine plasticity in VS, which then causes habits to be established in VMPFC. Eventually, though not demonstrated in the current simulations, plasticity occurs in areas of the dorsal pathway and the dorso-lateral striatum (see model predictions below).

Habit-like representations are formed in simulated frontal cortical areas of the model, specifically in VMPFC, which is capable of acquiring the direct stimulus-approach/avoid associations as a result of activation patterns driven by the VS. Learning in the VMPFC is determined by the difference in unit activations between when the units have settled to either approach or avoid a stimulus, and after feedback has been received. If it turned out that the decision to either approach or avoid was correct, synaptic weights are updated to make the same decision next time this stimulus is in the focus of attention. In order for the intermediate cortical layers (superficial maintenance layers II/III) to learn whether they indeed held on to the correct decision, they must have the same representations activated as the output-gated representations in deep layers. Bidirectional connectivity among these layers ensures that this is the case (Levitt, Lewis, Yoshioka, & Lund, 1993).

Detailed methods

In the following subsections, we highlight the major implementational features of the WHIP model, with equations and other details available in O'Reilly and Frank (2006). To provide a framework for these features, the main flow of processing in the model is as follows (see Figure 5.3):

- One positive stimulus (associated with reward) and one negative stimulus (not associated with reward) are presented to the sensory input layer at a time, one on the left and the other on the right – a total of 3 positive (A+, B+, C+) and 3 negative (Z–, Y–, Z–) stimuli are simulated. Due to prior maintained activation in the Frontal Eye Field (FEF) layer, one of them is in the focus of attention (let's say Y– in this example).
- The attentional effect is mediated through top-down projections from FEF to the parietal cortex layer, which then produces a selective-attention competition effect that enhances stimuli on the activated side of space (see e.g. Corbetta & Shulman, 2002).

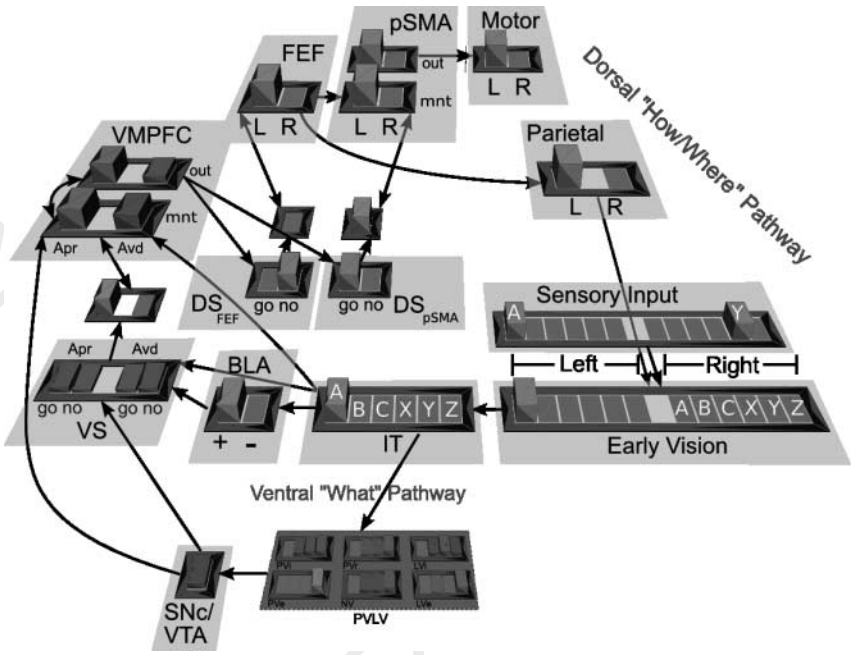

Figure 5.3 Labeled version of the actual model, showing all the relevant layers and their primary connections, and activations corresponding to a successful approach response to the A stimulus. Sensory input represents two items (A+, Y–), and via sustained activation, FEF provides top-down support for one location (left in this case), causing the Early Vision layer to preferentially encode the stimulus on the left side, which is then what IT exclusively represents. This "What" information feeds into ventral striatum (VS), activating an approach response in this case – more Go activation than NoGo for approach (Apr), and the opposite for avoid (Avd). This Go firing causes the VMPFC approach activations to be output-gated, thus activating the dorsal striatum units associated with approach (Go for pSMA, and NoGo for FEF). These dorsal actions then result in physical approach (Go in pSMA) to the current focus of attention (Left), and a maintenance (NoGo) of attention to this location in FEF. Abbreviations as in prior figure.

- The attended stimulus (Y) is then selectively processed by the ventral pathway, which represents only it and not the other stimulus.
- The attended Y– stimulus activates corresponding representations in ventral striatum (VS), via learned associations. In the trained, intact model, because it is not associated with reward, Y– would activate Go pathway units in the VS associated with the *avoid* response, and NoGo pathway units associated with the *approach* response.
- The VS Go–Avoid gating signal activates the associated Avoid representations in VMPFC, which then project to dorsal striatum and PFC areas

(FEF and pSMA) to activate the specific actions associated with the avoid response.

- For FEF, the avoid action is to update the currently maintained state, causing the system to switch its focus of attention to the other stimulus on the other side of space (e.g. Klein, 2000). This is achieved by activating Go pathway units in the dorsal striatum associated with FEF, which trigger updating of active maintenance in FEF.
- For pSMA, the avoid action is to prevent activation of the motor plan to approach the stimulus that is in the current focus of attention (i.e., Y). This is achieved by activating NoGo pathway units for gating the output of pSMA neurons, so that they do not then activate lower-level motor areas to actually carry out the approach plan. Note: the PBWM framework uses exclusively excitatory rather than inhibitory projections from prefrontal areas to striatal MSNs in order to inhibit response execution.
- Correctly avoiding Y– will result in no reward signal, but it allows the system to proceed to the next state of focusing on A+, which, if then approached, would result in reward (see next steps below). However, if the system decides instead to approach Y–, then it will not receive a reward in a location that is otherwise associated with the general expectation of reward – this negative mismatch between outcome and expectation then drives learning in the VS and VMPFC to not approach this stimulus in the future (i.e., reinforcing NoGo over Go for approach to the Y– stimulus).
- Following a correct avoid of Y–, the system on the next time step will be focusing its attention on A+ instead of Y–, as a result of the update in FEF. This A+ stimulus will then propagate through the ventral pathways, and in the trained system will preferentially activate Go firing for VS approach neurons, which will then trigger the approach plan in VMPFC, resulting in the following dorsal actions.
- FEF associated dorsal striatum receives NoGo projections from approach, causing it to *not* update its current state, and therefore maintain its current focus of attention.
- pSMA associated dorsal striatum receives Go projections from approach, causing it to activate its motor plan (approach item in current focus) in more posterior motor areas.
- Approaching the A+ stimulus triggers a dopamine reward, which reinforces the VS approach action (Go over NoGo for approach) as triggered by the A+ stimulus, making it more likely in the future that the system will approach the A+ stimulus.

All of the posterior-cortex sensory processing is done with a simple cartoon model of the what/where ventral/dorsal pathways, with localist units and hand-coded connections. The critical dynamics occur via the influences of the basal ganglia layers on their associated frontal cortex layers and the reward-based learning in these areas. This is all based on the PBWM algorithm,

which is described in several publications (Hazy et al., 2006, 2007; O'Reilly & Frank, 2006), and summarized briefly above.

All components of the model share a common set of simple integrate-and-fire style activation equations with point-neuron geometry, consistent with the Leabra framework in which the model is implemented (O'Reilly, 1998; O'Reilly & Munakata, 2000). Inhibitory dynamics in Leabra are implemented via a k-winners-take-all (kWTA) algorithm that robustly approximates the effects of inhibitory interneurons. Learning generally occurs via a combination of biologically plausible error-driven (O'Reilly, 1996; O'Reilly & Munakata, 2000) and Hebbian learning. Striatal neurons receive simulated dopaminergic inputs that modulate their activation states, which then drive learning. Phasic dopamine is simulated via the Primary Value, Learned Value (PVLV) algorithm (Hazy et al., 2007; O'Reilly et al., 2007), which captures the known properties of midbrain dopaminergic firing in terms of contributions from a collection of interacting brain areas.

Training parameters

The model was trained and tested according to the procedures described by Atallah et al. (2006). We ran 50 replications of each condition, with 4 different inactivation conditions that reflect those run in the Atallah et al. (2006) study. Each condition followed the same training protocol, based on the experimental study, with 4 simulated sessions of training (with the first 3 denoted as Acquisition and the final one as Test), each consisting of 60 trials. We made the task slightly more difficult for the model by having a total of six input stimuli, Atallah et al. (2006) trained rats on only two stimuli. The model was trained with three reward-associated input stimuli (A+, B+, and C+) and three not reward-associated stimuli (X–, Y–, and Z–). At the beginning of every new trial an algorithm randomly selected one of the three reward-associated stimuli to be presented either on the left or the right and one not reward-associated stimulus and presented them to the input layer on the remaining side. At the beginning of a new trial the focus of attention is initialized randomly to either the left or the right stimulus. A trial lasted until the model approached one of the stimuli. The model was rewarded on every correct approach and punished for every incorrect answer. If the model did not approach any of the stimuli in time (15 LeabraSettle cylces), it was punished for avoiding until it finally decided to approach one of the stimuli. The conditions tested were:

- Intact Control: no lesions or other manipulations.
- VS Inactivation, Acquisition: the ventral striatum is lesioned during the first 3 training sessions, then unlesioned for the 4th.
- VS Inactivation, Test: VS is intact for first 3 trials, then lesioned in the 4th.
- DS Inactivation, Acquisition: the dorsal striatum is lesioned during the first 3 training sessions, then unlesioned for the 4th.

- Intact Control, Pavlovian Pretraining: the model had three session of a Pavlovian pretraining before it was trained on the instrumental task.

Parameter fitting

No systematic attempt was made to fit the exact quantitative pattern of the rat behavioral data. Learning rate parameters were adjusted to capture the general timecourse of learning over the four sessions. Activation parameters in the perceptual pathways were set to produce the expected qualitative attentional effects but were not otherwise adjusted from Leabra defaults.

Results

As shown in Figure 5.4 (top), the overall pattern of results of our simulation captures the critical patterns found in the experimental data of Atallah et al. (2006). The intact control network learns to roughly 85% performance over the four sessions. Exactly how this learning is taking place is revealed by the inactivation conditions described in each of the following sections.

Inactivation of VS during acquisition

During acquisition, inactivation of the VS prevented task learning, similar to the impairments found in rats after the injection of Muscimol or AP-5 into the VS (Figure 5.4, top). This shows that the association between stimuli (coded in the IT area of the network) and the approach/avoid responses that are initiated by patterns of Go/NoGo firing in the VS is critical for learning the task in the network. Consistent with this, the network begins to learn in the test session (when the VS is no longer lesioned).

Inactivation of DS during acquisition

The WHIP model exhibits the critical spontaneous recovery of performance in the Test session after the DS has been lesioned throughout the acquisition phase (Figure 5.4, top). Thus, the DS in the model is essential for translating the basic approach/avoid responses into actual action plans that govern behavior, but it is not itself a key locus of learning in this task.

A particularly puzzling aspect of this spontaneous recovery of function after DS inactivation is that learning is apparently unaffected by the significant differences in overall behavior that the system exhibits during the acquisition phase with a DS lesion. If the system is sensitive to the behavioral contingency between action and reward outcome, then it would seem that the DS lesion should produce a strong dissociation between these two factors and thus negatively impact learning. For example, if the VS system signals an approach action to a rewarded stimulus but the DS system fails to execute that action, then the VS is not properly reinforced (and similarly unheeded

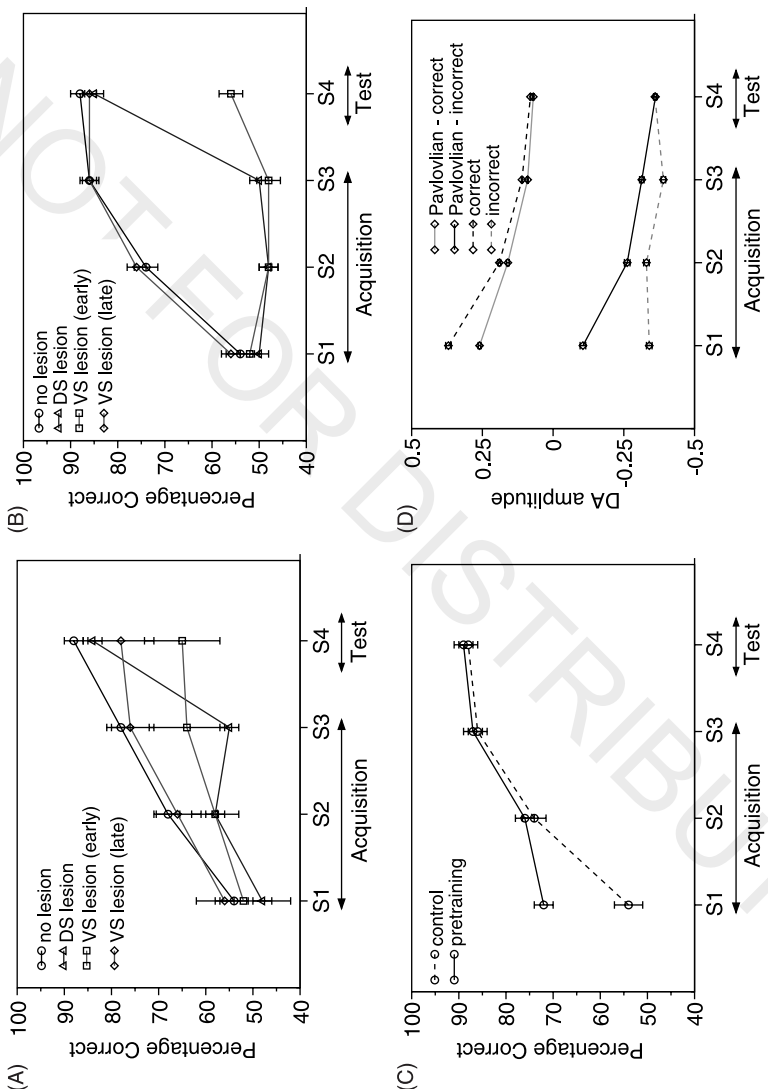

Figure 5.4 *Top:* (A) Mean performance of rats and the WHIP model (B) during Acquisition and Test phase. (C) Mean performance of the WHIP model during acquisition and test phases after Pavlovian pretraining. (D) Mean amplitude of phasic dopamine in correct trials and incorrect trials with Pavlovian pretraining (solid) and without Pavlovian pretraining (dashed). Error bars indicate standard error.

Perception, attention, and action:
International Perspectives on Psychological Science (Volume1). Peter A. Frensch and Ralf Schwarzer (Eds). 2010.
Published by Psychology Press on behalf of the International Union of Psychological Science.

VS avoid signals result in negative reinforcement for this otherwise correct plan). Indeed, this is exactly what we found in our initial model.

The model deals with this inconsistency by having the more Pavlovian-style learning signal provided by the BLA that can provide appropriate learning pressure on the VS approach and avoid patterns irrespective of the actual behavioral action taken. In order for this system to work, it has to be capable of selectively modulating the different VS representations as a function of the valence of the stimulus–reward associations, through learned connections. This is distinct from a global CNA-driven DA signal. With a simulated BLA included in the model, the approach response gets a positive training signal for stimuli associated with reward (e.g., A+), while the avoid response gets a positive training signal for stimuli associated with lower-than-expected reward, and this serves to train the VS even in the absence of appropriate behavioral responses during the DS lesion condition.

VS Inactivation during test: Learning in the VMPFC

When the model VS is lesioned only at the test phase, we see that performance remains at the same level as the prior session. This shows that although the VS is critical for learning of this task, it is not actually essential for task performance after acquisition. In the model, we can see that the VMPFC region learns to associate stimuli directly with approach and avoid responses and does not require the VS once these associations have formed. This is one form of "habit" formation taking place in the cortex, as opposed to the more traditional view of the basal ganglia being the locus of habit learning (Atallah et al., 2004; Frank, 2005; Houk & Wise, 1995; Loh et al., 2008; Pasupathy & Miller, 2005)

Pavlovian control

We ran a Pavlovian control condition similar to that used by Atallah et al. (2006), to determine how much of the model's behavior is governed by purely Pavlovian (as opposed to instrumental) learning. Although we know from the structure of the model that it is critical for learning to affect the approach and avoid plans (which are not activated under the Pavlovian protocol utilized by Atallah et al., 2006), it is nevertheless possible that the PVLV phasic dopamine system or the BLA system could learn in the absence of behavioral contingencies. In comparison to performance of the model without any Pavlovian pretraining and consistent with the Atallah et al. (2006) data, the model demonstrated a slight improvement at the beginning of acquisition but quickly converged to the performance of the control. Further tests revealed that the initial performance increase was caused by the fact that the simulated BLA had learned the Pavlovian CS–US association and to represent the valence of conditioned stimuli. Reduced reward prediction errors in PVLV after Pavlovian pretraining lead to smaller amplitudes of phasic DA from

PVLV and slowed down learning of behavioral contingencies so that in the second Acquisition session the performance of the model with Pavlovian pretraining converged towards the performance of the model without pretraining (see Figure 5.4, bottom).

Discussion

Our computational model demonstrates how a division of labor along multiple dimensions (dorsal "How" vs. ventral "What", Pavlovian vs. Instrumental conditioning, rapid learning in basal-ganglia vs. slow learning in cortex) can lead to the complex pattern of findings reported in Atallah et al. (2006). It does so by leveraging an existing computational framework for understanding the functional roles of the basal ganglia and prefrontal cortex in a range of different task contexts (Hazy et al., 2006, 2007; O'Reilly & Frank, 2006). This framework posits that the basal ganglia play a crucial modulatory role on learning and processing in frontal cortex, and this can explain the importance of learning in the ventral striatum as demonstrated by Atallah et al. (2006).

The differential importance of the ventral striatum relative to dorsal striatum in learning this task arises in the WHIP model because the system is able to leverage pre-existing approach/avoid response plans (coordinated through ventral-dorsal interactions) and thus mainly needs to learn which stimuli to associate with approach and which to associate with avoid. This division of labor is functionally relevant in a complex environment in which learning too much about which specific actions lead to reward decreases the ability to flexibly adapt to changing task requirements. This kind of stimulus-based learning is focused in the ventral pathway that originates in the ventral "what" processing areas of the cortex and continues to ventral striatal areas and associated ventromedial frontal areas. Thus, we re-cast the general director-vs.-actor distinction from Atallah et al. (2006) as a What vs. How/Where distinction in the present model. In this task context, the ventral "What" pathway is in the director's seat, but other tasks may load more heavily on the dorsal How/Where pathway (see below for further elaboration).

More generally, our model demonstrates that the original ventral-What vs. dorsal-How/Where distinction, based on visual processing (Goodale & Milner, 1992; Ungerleider & Mishkin, 1982), may have meaningful implications for ventral vs. dorsal frontal and associated striatal areas. This is consistent with the fact that (in the primate) dorsal frontal and striatal areas are similarly preferentially interconnected with dorsal visual pathway areas (e.g., parietal cortex), and relatively more ventral frontal and striatal areas are most strongly interconnected with ventral visual pathway areas (e.g., IT cortex) (Petrides, 2005; Petrides & Pandya, 1999, 2002; Sakagami & Tsutsui, 1999). In other work, we are exploring how much of the human and primate frontal cognitive neuroscience data this dissociation can account for – initial indications are that it could be substantial.

Critically, the striatal contribution is only during the initial portion of the training, inverting the standard association of habit learning with the basal ganglia and early "problem solving" with the frontal cortex. Indeed, our model shows that the modulatory function of the basal ganglia on the frontal cortex makes it uniquely capable of generating novel patterns of behavior at the start of training. Furthermore the strong phasic modulation from the midbrain dopamine system on learning in the striatum (and direct influence from the basolateral amygdala) enables it to rapidly acquire novel behavioral patterns, which are then more slowly encoded in the frontal cortex (for other consistent data see Loh et al., 2008; Pasupathy & Miller, 2005).

The WHIP model demonstrates that the relationship between instrumental and Pavlovian conditioning can be subtle and complex (e.g., Balleine & Ostlund, 2007). We predict that a model such as WHIP would also be able to account for data from a study on conditioned orienting (COR) by Han, McMahan, Holland, and Gallagher (1997), who found a similar pattern to the Atallah et al. (2006) results whereby reversible dorsal striatal lesions in very nearly the identical region (Han et al., 1997, Fig. 2c) as in Atallah et al. (2006, Supplementary Fig. 1) produced a similar spontaneous recovery effect: during temporary inactivation of a dorsal lateral region of the striatum no evidence of COR was observed, but it suddenly appeared after the lesion was removed (Han et al., 1997, Expt. 2, Fig. 7a). While the COR paradigm is undoubtedly truly Pavlovian, a purely Pavlovian explanation of Atallah et al.'s (2006) data has been ruled out by Pavlovian pre-training. Additionally, Cardinal et al. (2002) reported that in an autoshaping paradigm, post-training lesions of VS lead to a performance impairment, which is contradictory to a purely Pavlovian interpretation of Atallah et al.'s (2006) task.

Similar to the original behavioral data from Atallah et al. (2006), our model did not exhibit strong Pavlovian-instrumental transfer (PIT) from a pure Pavlovian pretraining phase (exposure to stimuli and associated reward without any behavioral contingencies). Nevertheless, it is clear that the behavioral contingencies are significantly degraded because of the inactivation of the dorsal striatum, and yet learning appears to be almost entirely unaffected. Our model resolves this apparent contradiction by including a more specific form of Pavlovian reinforcement on approach/avoid actions mediated by the basolateral amygdala. Interestingly, this characterization nicely fits the difference between the CNA (which drives a global dopamine signal according to the PVLV algorithm) and the BLA (Parkinson, Robbins, & Everitt, 2000) in tests of Pavlovian-instrumental transfer (PIT) (Corbit & Balleine, 2005; Holland & Gallagher, 2003). It was found that CNA was important for global but not for specific forms of PIT, and BLA was the reverse. This middle ground between Pavlovian and instrumental conditioning illustrates that the underlying mechanisms of the two are highly overlapping and interacting, sometimes leaving the impression that these two terms are merely poles of a continuum that exists across various neural systems in the brain.

Predictions from the model

The WHIP model makes a number of predictions beyond those present in the existing data. Here, we concretize a couple of the most important ones:

- BLA inactivation at the same time as DS inactivation should prevent the spontaneous recovery of function when the DS is functional in the test session. Furthermore, this should be a learning effect, such that BLA inactivation only in the test session (i.e., BLA intact during acquisition) should have no effect (i.e., performance comparable to intact control on test).
- The director-like role of the ventral pathway in this task paradigm should not generalize to other tasks that require more substantial amounts of new learning in the dorsal pathway. This is because the ventral pathway in this task can contribute very basic response strategies (approach vs. avoid) that are either learned very early on or are part of the animal's native repertoire of basic behavioral strategies.
- VMPFC inactivation during acquisition should prevent performance improvements during acquisition but lead to a immediate recovery of performance if the VMPFC is re-activated in the test session in the task of Atallah et al. (2006).
- Blocking plasticity in VMPFC during acquisition should impair performance in a final test sessions with an inactivated VS. If this is not the case, habits must have developed in more dorsal areas (DLS and pSMA).
- Atallah et al. (2006) used odors as conditioned stimuli and found that inactivating the odor processing ventral striatum prevented task acquisition. We predict that similar results should be obtained with visual or auditory conditioned stimuli if the corresponding striatal areas are inactivated.

Further target data

There is a vast range of data that this model could potentially be applied to, given that it encompasses many brain areas about which considerable data are available in the context of various conditioning and other related tasks. As just one example, the VMPFC region in our model can be associated with the orbitofrontal brain areas that are critical for reinforcer devaluation and other tasks that involve driving behavior based on stimulus value representations (e.g., Hatfield, Han, Conley, & Holland, 1996; Pickens, Saddoris, Gallagher, & Holland, 2005; Winstanley, Theobald, Cardinal, & Robbins, 2004). Its role in the present model is compatible with its apparent role in these other task contexts, so that it should be possible to account for all of these tasks within the same computational framework. Doing so would help to clarify the functional roles of the components of the WHIP

model and further help to justify the relative complexity of the model by demonstrating the important separable contributions of each region across a range of tasks.

Another important extension of this model framework would be to tasks that require more elaborate dorsal-pathway instrumental action learning, to flesh out the full continuum of instrumental learning that we hypothesize to exist. This model would then be capable of making more precise predictions regarding brain areas critical for different forms of task learning, building on the initial foundation provided by the present model.

In short, there is considerable work to be done to build upon the initial success of this model in accounting for empirical data. This work is under way, and we anticipate it will result in a more thoroughly validated model of a wide range of condition and other learning phenomena.

Note

1 We adopt a primate/human oriented anatomical organization with a focus on visual stimuli, because this is likely more widely known, but the same kinds of functional distinctions are found in rats and other sensory modalities. Relatedly, inactivation of the ventral striatum (which processes odor information) in rats of Atallah et al.'s (2006) study corresponds to an inactivation of the ventromedial striatum, which receives visual information through inferotemporal cortex primates and the WHIP model (Cheng, Saleem, & Tanaka, 1997; Yeterian & Pandya, 1995).

Acknowledgments

Supported by ONR grants N00014–07–1–0651 and N00014–03–1–0428, and NIH grants MH069597 and MH079485.

References

Ahn, S., & Phillips, A. G. (2003). Independent modulation of basal and feeding-evoked dopamine efflux in the nucleus accumbens and medial prefrontal cortex by the central and basolateral amygdalar nuclei in the rat. *Neuroscience*, *116*, 295–305.

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.

Atallah, H. E., Frank, M. J., & O'Reilly, R. C. (2004). Hippocampus, cortex, and basal ganglia: Insights from computational models of complementary learning systems. *Neurobiology of Learning and Memory*, *82*(3), 253–267.

Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., & O'Reilly, R. C. (2006). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nature Neuroscience*, *10*, 126–131.

Balleine, B. W., & Ostlund, S. B. (2007). Still at the choice-point: Action selection and initiation in instrumental conditioning. *Annals of the New York Academy of Sciences*, *1104*, 147–171.

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, & Cybernetics*, *13*(5), 834–846.

Berns, G. S., & Sejnowski, T. J. (1998). A computational model of how the basal ganglia produces sequences. *Journal of Cognitive Neuroscience*, *10*, 108–121.

Brown, J. W., Bullock, D., & Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks*, *17*, 471–510.

Cardinal, R. N., Parkinson, J. A., Lachenal, G., Halkerston, K. M., Rudarakanchana, N., Hall, J., et al. (2002). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behavioral Neuroscience*, *116*, 553–567.

Cheng, K., Saleem, K. S., & Tanaka, K. (1997). Organization of corticostriatal and corticoamygdalar projections arising from the anterior inferotemporal area te of the macaque monkey: A phaseolus vulgaris leucoagglutinin study. *Journal of Neuroscience*, *17*(20), 7902–7925.

Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1485), 1585–1599.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*, 201–215.

Corbit, L. H., & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *Journal of Neuroscience*, *25*(4), 962–970.

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural networks*, *19*(8), 1153–1160.

Dominey, P., Arbib, M., & Joseph, J.-P. (1995). A model of corticostriatal plasticity for learning oculomotor associations and sequences. *Journal of Cognitive Neuroscience*, *7*, 311–336.

Featherstone, R. E., & McDonald, R. J. (2004). Dorsal striatum and stimulus–response learning: Lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a simple discrimination task. *Behavioural Brain Research*, *150*(1–2), 15–23.

Fiehler, K., Burke, M., Bien, S., Röder, B., & Rösler, F. (2008). The human dorsal action control system develops in the absence of vision. *Cerebral Cortex*, *19*(1), 1–12.

Floresco, S. B., West, A. R., Ash, B., Moore, H., & Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience*, *6*, 968–973.

Floresco, S. B., Yang, C. R., Phillips, A. G., & Blaha, C. D. (1998). Basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat. *European Journal of Neuroscience*, *10*(4), 1241–1251.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated parkinsonism. *Journal of Cognitive Neuroscience*, *17*, 51–72.

Frank, M. J. (2006). Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks*, *19*, 1120–1136.

Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007). Hold your

horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, *318*, 1309–1312.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*, *306*, 1940–1943.

Fudge, J. L., & Haber, S. N. (2000). The central nucleus of the amygdala projection to dopamine subpopulations in primates. *Neuroscience*, *97*, 479–494.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory*, *70*(1–2), 119–136.

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

Han, J. S., McMahan, R. W., Holland, P., & Gallagher, M. (1997). The role of an amygdalo-nigrostriatal pathway in associative learning. *Journal of Neuroscience*, *17*(10), 3913–3919.

Hatfield, T., Han, J. S., Conley, M., & Holland, P. (1996). Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. *Journal of Neuroscience*, *16*, 5256–5265.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience*, *139*, 105–118.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2007). Towards an executive without a homunculus: Computational models of the prefrontal cortex/basal ganglia system. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1), 105–118.

Himmelbach, M., & Karnath, H.-O. (2005). Dorsal and ventral stream interaction: Contributions from optic ataxia. *Journal of Cognitive Neuroscience*, *17*(4), 632–640.

Holland, P. C., & Gallagher, M. (2003). Double dissociation of the effects of lesions of basolateral and central amygdala on conditioned stimulus-potentiated feeding and Pavlovian-instrumental transfer. *European Journal of Neuroscience*, *17*(8), 1680–1694.

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. 233–248). Cambridge, MA: MIT Press.

Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., et al. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1485), 1573–1583.

Houk, J. C., & Wise, S. P. (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: Their role in planning and controlling action. *Cerebral Cortex*, *5*, 95–110.

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

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, *126*(11), 2463–2475.

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.

Johnson, L. R., Aylward, R. L., Hussain, Z., & Totterdell, S. (1995). Input from the amygdala to the rat nucleus accumbens: Its relationship with tyrosine hydroxylase immunoreactivity and identified neurons. *Neuroscience*, *61*(4), 851–865.

Kantak, K. M., Green-Jordan, K., Valencia, E., Kremin, T., & Eichenbaum, H. B. (2001). Cognitive task performance after lidocaine-induced inactivation of different sites within the basolateral amygdala and dorsal striatum. *Behavioral Neuroscience*, *115*, 589–601.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Levitt, J. B., Lewis, D. A., Yoshioka, T., & Lund, J. S. (1993). Topography of pyramidal neuron intrinsic connections in macaque monkey prefrontal cortex (areas 9 & 46). *Journal of Comparative Neurology*, *338*, 360–376.

Loh, M., Pasupathy, A., Miller, E. K., & Deco, G. (2008). Neurodynamics of the prefrontal cortex during conditional visuomotor associations. *Journal of Cognitive Neuroscience*, *20*(3), 421–431.

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*, 381–425.

Murray, E. A. (2007). The amygdala, reward and emotion. *Trends in Cognitive Sciences*, *11*(11), 489–497.

O'Doherty, J. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*(5669), 452–454.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*(5), 895–938.

O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences*, *2*(11), 455–462.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283–328.

O'Reilly, R. C., Frank, M. J., Hazy, T. E., & Watz, B. (2007). PVLV: The primary value and learned value pavlovian learning algorithm. *Behavioral Neuroscience*, *121*, 31–49.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

Packard, M. G., Hirsh, R., & White, N. M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience*, *9*, 1465–1472.

Packard, M. G., & McGaugh, J. L. (1992). Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. *Behavioral Neuroscience*, *106*(3), 439–446.

Parkinson, J. A., Robbins, T. W., & Everitt, B. J. (2000). Dissociable roles of the central

and basolateral amygdala in appetitive emotional learning. *European Journal of Neuroscience*, *12*, 405–413.

Pasupathy, A., & Miller, E. K. (2005). Different time courses for learning-related activity in the prefrontal cortex and striatum. *Nature*, *433*, 873–876.

Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 781–795.

Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex: Comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience*, *11*(3), 1011–1036.

Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291–310.

Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal lesions impair use of cue–outcome associations in a devaluation task. *Behavioral Neuroscience*, *119*(1), 317–322.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology*, *13*(4), 564–574.

Rouillard, C., & Freeman, A. S. (1997). Effects of electrical stimulation of the central nucleus of the amygdala on the in vivo electrophysiological activity of rat nigral dopaminergic neurons. *Synapse*, *21*, 348–356.

Sakagami, M., & Tsutsui, K. (1999). The hierarchical organization of decision making in the primate prefrontal cortex. *Neuroscience Research*, *34*, 79.

Schoenbaum, G., Chiba, A. A., & Gallagher, M. (1999). Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *Journal of Neuroscience*, *19*, 1876–1884.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, *13*, 900–913.

Semba, K., & Fibiger, H. (1992). Afferent connections of the laterodorsal and the pedunculopontine tegmental nuclei in the rat: A retro- and antero-grade transport and immunohistochemical study. *Journal of Comparative Neurology*, *323*, 387–410.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

Sutton, R. S. (1988). Learning to predict by the method of temporal diferences. *Machine Learning*, *3*, 9–44.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Toni, I., Rowe, J., Stephan, K. E., & Passingham, R. E. (2002). Changes of corticostriatal effective connectivity during visuomotor learning. *Cerebral Cortex*, *12*(10), 1040–1047.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *The analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Vogels, R. (1999). Categorization of complex visual images by rhesus monkeys. Part I: Behavioural study. *European Journal of Neuroscience*, *11*, 1223–1238.

Wickens, J. R., Kotter, R., & Alexander, M. E. (1995). Effects of local connectivity on striatal function: Simulation and analysis of a model. *Synapse*, *20*, 281–298.

Williams, Z. M., & Eskandar, E. N. (2006). Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nature Neuroscience*, *9*, 562–568.

Winstanley, C. A., Theobald, D. E. H., Cardinal, R. N., & Robbins, T. W. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *Journal of Neuroscience*, *24*(20), 4718–4722.

Yeterian, E. H., & Pandya, D. N. (1995). Corticostriatal connections of extrastriate visual areas in rhesus monkeys. *Journal of Comparative Neurology*, *352*(3), 436–457.