



*Annual Review of Psychology*

# An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making

Melissa J. Sharpe,<sup>1,2,3</sup> Thomas Stalnaker,<sup>1</sup>  
Nicolas W. Schuck,<sup>4</sup> Simon Killcross,<sup>3</sup>  
Geoffrey Schoenbaum,<sup>1,5,6</sup> and Yael Niv<sup>2,7</sup>

<sup>1</sup>Intramural Research Program, National Institute on Drug Abuse, Baltimore, Maryland 21224, USA; email: mjsharpe@princeton.edu, thomas.stalnaker@nih.gov, gscho002@gmail.com

<sup>2</sup>Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey 08544, USA; email: yael@princeton.edu

<sup>3</sup>School of Psychology, UNSW Sydney, New South Wales 2052, Australia; email: s.killcross@unsw.edu.au

<sup>4</sup>Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, 14195 Berlin, Germany; email: nico.schuck@gmail.com

<sup>5</sup>Departments of Anatomy & Neurobiology and Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA

<sup>6</sup>Solomon H. Snyder Department of Neuroscience, Johns Hopkins University, Baltimore, Maryland 21287, USA

<sup>7</sup>Psychology Department, Princeton University, Princeton, New Jersey 08544, USA

Annu. Rev. Psychol. 2019. 70:28.1–28.24

The *Annual Review of Psychology* is online at [psych.annualreviews.org](http://psych.annualreviews.org)

<https://doi.org/10.1146/annurev-psych-010418-102824>

Copyright © 2019 by Annual Reviews.  
All rights reserved

## Keywords

orbitofrontal cortex, prelimbic cortex, state, striatum, dorsolateral prefrontal cortex, decision making

## Abstract

Making decisions in environments with few choice options is easy. We select the action that results in the most valued outcome. Making decisions in more complex environments, where the same action can produce different outcomes in different conditions, is much harder. In such circumstances, we propose that accurate action selection relies on top-down control from the prelimbic and orbitofrontal cortices over striatal activity through distinct thalamostriatal circuits. We suggest that the prelimbic cortex exerts direct



influence over medium spiny neurons in the dorsomedial striatum to represent the state space relevant to the current environment. Conversely, the orbitofrontal cortex is argued to track a subject's position within that state space, likely through modulation of cholinergic interneurons.

## Contents

1. INTRODUCTION .....	28.2
2. THE DORSAL STRIATUM HARBORS OUR CHOICE HISTORY .....	28.2
3. THE HIPPOCAMPUS HARBORS THE ASSOCIATIVE HISTORY OF PAST EXPERIENCE .....	28.4
4. THE ORBITOFRONTAL CORTEX REPRESENTS CURRENT STATE .....	28.4
5. THE PRELIMBIC CORTEX ALLOWS HIGHER-ORDER INFORMATION TO GUIDE CHOICE .....	28.7
6. HOW MIGHT THESE DISTINCT REGIONS WORK IN CONCERT TO INFLUENCE FLEXIBLE BEHAVIOR? .....	28.11
7. CONCLUSIONS .....	28.17

## 1. INTRODUCTION

Choosing an ice cream flavor can be tough. Gone are the days when we had a choice between three standard options—vanilla, strawberry, or chocolate. There has been an ice cream revolution, and we are now faced with choices between hot fudge pistachio cheese cake and chunky monkey. Moreover, even though these same concoctions might be available in many different ice cream shops, their actual flavors often differ widely depending on the particular shop. The Bent Spoon's salted caramel might be the perfect balance between sweet chunks of caramel and a light crunch of salt. But the salted caramel tub at the grocery store might induce a powerful salt-driven thirst. We therefore have to flexibly change our choice depending on the particular location in which we find ourselves and our current appetite for salt.

Of course, decisions often have higher stakes. Drinking a little too much might soon be forgotten if it occurs at a dinner party with friends, but it could cost you your job if it occurs at a work function. And sending a text puts you in no danger—unless you are driving. We are constantly faced with difficult decisions that are best made by considering all of the factors available in our current environment. Thankfully, our brains are wired to integrate a wide array of information to flexibly influence our behavior in different circumstances. In this article, we briefly review research that has explored a role for different neural regions in contributing to action selection. We then propose a novel theory of how these different neural circuits work in concert to allow us to make optimal decisions in complex and changing environments.

## 2. THE DORSAL STRIATUM HARBORS OUR CHOICE HISTORY

We have known for a long time that changes in striatal activity directly influence the ability of humans and other animals to execute an appropriate action. Many disorders that compromise function in striatal circuits produce powerful changes in motor responses, such as Parkinson's disease or dystonia (Coyle & Snyder 1969, Kish et al. 1988, Naumann et al. 1998). Furthermore, dorsal striatal activity in humans, nonhuman primates, and rodents is correlated with the degree



of motivation to work for a particular reward (Palmiter 2008, Volkow et al. 2002), the ability to perform previously learned action sequences (McDonald & White 1993, Miyachi et al. 1997), inhibition of a prepotent response (Eagle & Robbins 2003, Eagle et al. 2011), context-dependent modulation of responding (Fuchs et al. 2006, Nieuwenhuis et al. 2005, Stalnaker et al. 2016), task switching (Baunez & Robbins 1999, Quinlan et al. 2008), extinction of responding (Barnes et al. 2005, Fuchs et al. 2006), reversal learning (Clarke et al. 2008), and the sensitivity of a response to changes in reward value (Tricomi et al. 2009; Yin et al. 2004, 2005, 2006).

Such research has implicated the striatum as a critical interface for integration of information processing from other neural regions to coordinate motor output (Balleine et al. 2007, Barnes et al. 2005, Devan et al. 2011, Graybiel 2005, Liljeholm & O'Doherty 2012). In line with this view, striatal neurons are recruited during learning of instrumental tasks, with neuronal firing developing across time to encode the sequences of responses that are necessary to acquire rewards (Jog et al. 1999, O'Doherty et al. 2004, Thorn et al. 2010). This tuning of striatal activity to reward-relevant features of a task is directly related to choice: The greater the proportion of neurons in dorsal striatum tuned to the relevant features of the choice, the better is choice performance (Jog et al. 1999, Thorn et al. 2010). This can also be seen on a trial-by-trial basis where choice-relevant striatal activity predicts subject performance on any particular trial (Brasted & Wise 2004, Parker et al. 2016, Samejima et al. 2005, Stalnaker et al. 2016, Znamenskiy & Zador 2013). Indeed, optogenetic activation of striatal neurons can bias choice behavior on the current trial in a manner that can be predicted by task-relevant activity encoded in dorsal striatal neurons (Znamenskiy & Zador 2013). Thus, striatal activity is thought to reflect a representation of learned responses, which facilitates effective choice performance in familiar environments on an up-to-the-moment basis.

Different regions of the dorsal striatum receive projections from distinct circuits and contribute to choice behavior in different ways. On the one hand, the dorsomedial striatum (DMS) receives inputs from higher-order cortical areas such as prefrontal and orbital regions (Voorn et al. 2004). Accordingly, damage to or inactivation of the DMS impairs the ability to change responses with fluctuations in the desirability of the goal (Yin et al. 2005), to shift strategies (Ragozzino et al. 2002), and to change responses with a reversal in contingencies (Castañé et al. 2010, Ragozzino et al. 2009). Furthermore, neuronal activity in the DMS reflects complex aspects of choice performance, such as timing, contextual factors, and goal specificity (Emmons et al. 2017, Fuchs et al. 2006, Parker et al. 2016, Stalnaker et al. 2016). As a result, the DMS is widely considered to be the region of the striatum that facilitates flexible choice performance under changing conditions.

On the other hand, the dorsolateral striatum (DLS) receives projections from sensorimotor cortices (Balleine et al. 2007, Voorn et al. 2004) and is implicated in the development of behavioral automaticity. Data show that damage to or inactivation of the DLS does not impair flexibility of choice (Devan et al. 2011, Yin et al. 2004). Instead, these manipulations result in increased sensitivity to changes in goal value even after overtraining, when choice is usually insensitive to goal value (Yin et al. 2006).

Neuronal activity in the DLS evolves across experience to reflect the bracketing of task-relevant features. More specifically, neural coding in the DLS shifts from encoding many aspects of a task to encoding the most relevant features that signal the initiation and end of a sequence of actions or a trial (Barnes et al. 2005, Jog et al. 1999, Thorn et al. 2010). This is unlike encoding in the DMS, which reflects many aspects of a trial relevant to retrieving rewards and does not encode the initiation or end of a trial in the manner seen in the DLS (Thorn et al. 2010). Thus, in contrast to the role of the DMS in behavioral flexibility, these findings implicate the DLS in established behavioral repertoires that facilitate automatic, habitual responding (Balleine et al. 2007, Barnes et al. 2005, Jog et al. 1999, Liljeholm & O'Doherty 2012, Yin et al. 2009).

**State:** term used in computational reinforcement learning (Sutton & Barto 1998) referring to the current position within a particular task that changes by the moment

**Cognitive map:** also called the state space; mental representation of states and the probabilities of transitioning between them, capturing the associative structure of a task

**Context:** physical characteristics of an environment that generally remain static across time (e.g., contextual cues in an experimental chamber)



### 3. THE HIPPOCAMPUS HARBORS THE ASSOCIATIVE HISTORY OF PAST EXPERIENCE

While striatal function supports selection of an appropriate action in a given circumstance, the hippocampus has been widely implicated in learning about the general structure of our environment. This appears to be the case regardless of whether there is a need to influence motor output or reward seeking (McDonald & White 1995, McIntyre et al. 2002, Xu et al. 1998). In particular, the long-held view of the hippocampus has been that it functions to encode a map of spatial environments (O'Keefe & Nadel 1978). This theory has been supported by a plethora of elegant studies that have shown that hippocampal neurons encode detailed maps of space and a subject's current place within those maps (Langston et al. 2010, Morris et al. 1982, O'Keefe & Nadel 1978, O'Keefe & Speakman 1987). In addition, functional inactivation of or damage to hippocampal regions impairs spatial navigation (Morris et al. 1982) and learning about contexts (Corcoran et al. 2005, Matus-Amat et al. 2004); thus, the hippocampus is generally interpreted as influencing the development of a configuration that binds elements of a context together to form a holistic spatial representation (Matus-Amat et al. 2004, Rudy & Sutherland 1989).

It has been argued that the role of the hippocampus in spatial mapping reflects a more general contribution of the hippocampus to relational memory (Eichenbaum 2000, Milner et al. 1968, Redish 1999, Wallenstein et al. 1998, Wikenheiser & Schoenbaum 2016). In line with this broader view of hippocampal function, activity in hippocampal neurons also reflects information about related events in a manner that extends beyond space. For instance, hippocampal activity encodes relationships between discrete stimuli presented in pairs or sequences (Berger et al. 1976, Schendan et al. 2003), as well as upcoming rewards in a manner that is sensitive to the subject's current desire for those rewards (Wikenheiser & Redish 2015). Furthermore, damage to or inactivation of the hippocampus disrupts discrimination learning (Brasted et al. 2003, Davachi & Wagner 2002, Mahut et al. 1982), impairs transitive inference of stimulus relationships presented in sequence (Dusek & Eichenbaum 1997), and reduces behavior dependent on explicit stimulus-reward associations (Miller et al. 2017a). These data support the idea that hippocampal regions contribute to the binding of correlated events to form a broad cognitive map of relationships among stimuli in the environment, which can facilitate learning and action selection in both spatial and nonspatial tasks (Daw et al. 2005, Dusek & Eichenbaum 1997, Eichenbaum 2000, Wikenheiser & Schoenbaum 2016).

### 4. THE ORBITOFRONTAL CORTEX REPRESENTS CURRENT STATE

In contrast to the work on hippocampal and striatal regions, research investigating orbitofrontal function has produced a set of data that has been relatively difficult to interpret within a single theory (Murray et al. 2007, Stalnaker et al. 2015, Wallis 2012). Perhaps part of the reason for this is that damage to, or inactivation of, the orbitofrontal cortex (OFC) produces quite subtle deficits in behavior that can often dissipate following more experience or training. That is, although damage to the OFC has a profound impact on decision making, there is almost nothing that you absolutely cannot do without a functioning OFC, given enough prompting, time, or practice.

For instance, lesions or inactivation of the OFC across species produce deficits in reversal learning (Butter 1969, Rudebeck & Murray 2008, Schoenbaum et al. 2003), but these deficits can be overcome with extended training (Boulougouris et al. 2007, Jang et al. 2015) and are not seen with more restricted OFC damage in primates (Rudebeck et al. 2013). Similarly, OFC neurons have also been found to encode information related to the value of reward-predictive cues (Padoa-Schioppa & Assad 2006, Schultz et al. 2017), but functional inactivation of the OFC does not disrupt value-based choice (Gardner et al. 2017). Furthermore, the OFC has been implicated in



sensitivity to changes in the value of a reward, where lesions or inactivation of the OFC cause subjects to continue to respond to a cue that predicts a devalued reward (Gallagher et al. 1999, Ostlund & Balleine 2007). However, functional inactivation of the OFC does not disrupt goal devaluation in an instrumental setting (Ostlund & Balleine 2007, Parkes et al. 2017; but see Renteria et al. 2018). Thus, pinpointing a function to ascribe to this region has been challenging (Murray et al. 2007, Stalnaker et al. 2015).

To meet the challenge of these perplexing data, Wilson et al. (2014) proposed a theory according to which the OFC represents the current state of the task at any point in time. In this theory, the notion of state captures the underlying structure of the task, incorporating external (observable) information about the state of the environment, as well as any relevant internal (unobservable or hidden) information. For example, the latter can include recent actions or a remembered task instruction that may not be available perceptually (Wilson et al. 2014). The state of the task therefore includes all the information, observable or internal, that is needed to determine how different actions (or inaction) will affect the next state of the task [the terminology of state comes from the computational theory of reinforcement learning; it relates to the Markov property—a description of the task that, at each point in time, includes all the information that is needed to determine the probability of obtaining immediate reward and transitioning to a new state of the task (Sutton & Barto 1998)]. Intuitively, according to the theory, the state space of the task is like a cognitive map of the task contingencies, and at any point in time, the OFC conveys a “you are here” signal within this map of states. If the cognitive map corresponds one to one with external stimuli (that is, the states are fully observable), then this signal is redundant. However, if the map is only partially observable, then knowing your current location becomes critical for appropriate decision making: If you do not know where you are, that is, when the stimuli around you are not sufficient to determine your current location, then you cannot decide in what direction to go. This is when the “you are here” information from the OFC, indicating the current state within the particular cognitive map or state space, becomes invaluable.

This theory therefore suggests that the OFC is important for basing actions on different contingencies across different cognitive spaces. The OFC directs the subject to the relevant set of contingencies, allowing them to make a decision that is appropriate to that particular moment in time. In this manner, the OFC can facilitate learning and choice behavior because it allows a subject to effectively use a complex cognitive map comprising distinct memories in distinct spaces associated with that particular task.

Consider an example where a rat first learns that a particular lever-press response leads to a tasty food pellet (i.e., response 1 leads to outcome 1, or R1–O1), and another lever-press response produces no reward (i.e., R2–nothing). Under these circumstances, subjects will favor making the lever-press response that leads to the rewarding outcome (i.e., R1). However, then consider that these contingencies are reversed. Now the lever-press response that previously produced the reward no longer does (i.e., R1–nothing), and the lever press that previously produced no outcome will result in reward delivery (i.e., R2–O1). In this case, subjects will shift their response toward the other lever to retrieve the reward. A state theory argues that rats will not unlearn the original association. Rather, they will encode the two contingencies as two different states of the task (i.e., state A, in which R1–O1 and R2–nothing are true, and state B, in which R1–nothing and R2–O1 are true) (Gershman et al. 2010, Wilson et al. 2014). This is adaptive because it means that, if the environmental contingencies reversed again, then the rats would be able to quickly infer that state A is again active and adjust their responses accordingly, rather than again unlearning and relearning the reversed contingencies.

Indeed, the state-representation theory of the OFC explains the reversal deficit seen in subjects whose OFC function is compromised as resulting from failure to disambiguate the two states, A





and B, that are externally identical (but can be distinguished based on remembering the outcomes of recent actions) (Wilson et al. 2014). That is, as the states of the task are not explicitly signaled, without the OFC's "you are here" signal, subjects cannot correctly select the response corresponding to the present state within the task. Instead, it is proposed that subjects without OFC function have to resort to unlearning and relearning the associations whenever the task changes. This results in slower adjustment to subsequent reversals and is consistent with findings that OFC damage results in slower switching in tasks requiring reversals, rather than an inability to shift between the contingencies per se (Boulougouris et al. 2007; Butter 1969; Rudebeck & Murray 2008; Schoenbaum et al. 2000, 2003). This theory can also explain a wide range of other deficits seen with OFC damage or inactivation, such as those seen in delayed alternation tasks (Mishkin et al. 1969), reward devaluation (Gallagher et al. 1999, Ostlund & Balleine 2007), extinction learning (Butter et al. 1963), and intradimensional set shifting (Chase et al. 2012), by assuming that the same deficit in flexibly making use of different contingencies in different cognitive spaces (Wilson et al. 2014). Generally, the theory predicts that OFC deficits will be more pronounced whenever external cues need to be treated differently in different unobservable contexts. In simpler situations when one set of contingencies always applies (e.g., in basic Pavlovian or instrumental conditioning), the OFC is not necessary for action selection. That is, according to the theory, if task states are fully observable (e.g., there are different cues for each contingency), then other regions can select actions without OFC input. However, if inference about the current location in the state space is necessary due to some partial observability of the task, then the OFC becomes critical.

Wilson et al.'s (2014) state-representation theory of OFC function is also supported by the finding that OFC activity appears to reflect encoding of distinct contingencies in different states (Schoenbaum et al. 2000, Schuck et al. 2016, Sharpe & Schoenbaum 2016). For instance, Schoenbaum et al. (2000) recorded activity in the OFC of rats performing an odor discrimination task. Initially, rats were trained to sample an odor at the start of a trial. Presentation of one odor indicated that the rat would be rewarded at a well delivering sucrose solution, while presentation of another odor indicated that the rat would not be rewarded in that trial. Across this initial phase of learning, neurons in the OFC gradually showed discrimination between the rewarded and unrewarded odors, where one ensemble of neurons exhibited high firing rates for the rewarded odor, and another ensemble exhibited equally high firing rates for the unrewarded odor (indicating a coding of stimulus–reward associations rather than of value). The contingencies were then reversed so that the previously rewarded odor was now unrewarded, and vice versa. Under these circumstances, a large proportion of OFC neurons encoding the odors in the initial phase stopped responding. Instead, new neurons were recruited, which encoded the new odor–outcome relationships. This pattern of activity supports the notion advocated by Wilson et al. (2014) that OFC activity distinguishes between the contingencies that are relevant to different states within the task using distinct neuronal ensembles.

The state theory of OFC function suggests that OFC neurons reflect state encoding at a higher resolution than just a broad encoding of context that may be stable across long periods of time (Wilson et al. 2014). That is, consistent with a conception of state as defined in the computational theory of reinforcement learning (Sutton & Barto 1998), neural encoding in the OFC seems to track the current state of the task in a manner that changes on a moment-to-moment basis (Schuck et al. 2016, Wilson et al. 2014). For example, even in simple tasks where rats are presented with a series of odors that dictate either a go or a no-go response, OFC neurons exhibit changes in activity at each phase of a trial (Schoenbaum & Eichenbaum 1995). More specifically, OFC neurons are synchronized to all of the different events that take place within a single trial (e.g., odor delivery, response, reward delivery) so that their activity changes as the animal travels through the different



states associated with a particular trial. At the single-unit level, this is implemented through distinct neurons encoding different events within a trial.

Activity in the OFC might therefore be considered analogous to driving from one place to another within a given city. In this analogy, the broad context is the particular city that you are in. However, the state (location) changes with every block that you drive. Neural activity in the OFC therefore resembles the blue dot that signals where you are in a Google map rather than encoding a static representation of the city.

### 5. THE PRELIMBIC CORTEX ALLOWS HIGHER-ORDER INFORMATION TO GUIDE CHOICE

The prelimbic (PL) cortex in rats is also important for maintaining the flexibility of choice performance in circumstances where the environmental contingencies have changed; however, this region appears to do so in a qualitatively different way from the OFC. [The PL cortex is broadly analogous to the dorsolateral and ventromedial prefrontal cortices in primates, depending on the demands of the task (see **Figure 1**), although this mapping is not uncontested (Neubert et al. 2014, Uylings et al. 2003). In this review, we refer to the PL cortex but note that our framework is relevant for human, nonhuman primate, and rodent cognitive systems.] Damage to or inactivation

	Human brain	Rodent brain
<b>Frontal cortices</b>	<p><b>Dorsolateral prefrontal cortex</b></p> <ul style="list-style-type: none"> <li>• Attentional set shifting</li> <li>• Contextual modulation</li> <li>• Strategy shifting</li> <li>• Rule learning</li> <li>• Task switching</li> </ul> <p><b>Medial prefrontal cortex</b></p> <ul style="list-style-type: none"> <li>• Goal-directed behavior</li> </ul> <p><b>Orbitofrontal cortex</b></p> <ul style="list-style-type: none"> <li>• Reversal learning</li> <li>• Delayed alternation</li> <li>• Pavlovian devaluation</li> </ul>	<p><b>Prelimbic cortex</b></p> <ul style="list-style-type: none"> <li>• Attentional set shifting</li> <li>• Contextual modulation</li> <li>• Strategy shifting</li> <li>• Rule learning</li> <li>• Task switching</li> <li>• Contextual modulation</li> </ul> <p><b>Orbitofrontal cortex</b></p> <ul style="list-style-type: none"> <li>• Reversal learning</li> <li>• Delayed alternation</li> <li>• Pavlovian devaluation</li> </ul>
<b>Striatal regions</b>	<p><b>Caudate</b></p> <ul style="list-style-type: none"> <li>• Goal-directed behaviour</li> <li>• Attentional set shifting</li> <li>• Strategy shifting</li> <li>• Task switching</li> <li>• Contextual modulation</li> </ul> <p><b>Putamen</b></p> <ul style="list-style-type: none"> <li>• Habits</li> <li>• Task bracketing</li> </ul>	<p><b>Dorsomedial striatum</b></p> <ul style="list-style-type: none"> <li>• Goal-directed behaviour</li> <li>• Attentional set shifting</li> <li>• Strategy shifting</li> <li>• Task switching</li> <li>• Contextual modulation</li> </ul> <p><b>Dorsolateral striatum</b></p> <ul style="list-style-type: none"> <li>• Habits</li> <li>• Task bracketing</li> </ul>
<b>Hippocampal regions</b>	<p><b>Hippocampus</b></p> <ul style="list-style-type: none"> <li>• Spatial navigation</li> <li>• Associative learning</li> <li>• Transitive inference</li> </ul>	<p><b>Hippocampus</b></p> <ul style="list-style-type: none"> <li>• Spatial navigation</li> <li>• Associative learning</li> <li>• Transitive inference</li> </ul>

**Figure 1**

Human and rodent homologies. Despite the remarkable functional and physiological homology between the human and rodent brain, there is some physiological differentiation between frontal cortices. Specifically, the prelimbic region of the rat regulates the flexibility of behavior by higher-order rules as well as goal-directed behavior, while in the human these functions seem to be distributed across both the dorsolateral prefrontal cortex and the medial prefrontal cortex.



of the PL cortex does not impair performance on tasks that may benefit from tracking a particular state within a particular trial (e.g., an intradimensional set shift, reversal learning, or delayed alternation; Chase et al. 2012, Chudasama et al. 2001, Gisquet-Verrier & Delatour 2006), as seen with OFC damage. Rather, manipulation of activity in the PL cortex results in impairments in task performance when there is a change in the current task demands that requires the use of a different choice strategy altogether. For example, lesions of the PL cortex disrupt the ability of rats to switch from using an odor strategy to using a place strategy to retrieve rewards in a T-maze task (Ragozzino et al. 2003), or from use of a visual strategy to use of a response strategy in choosing the appropriate lever-press response that leads to delivery of reward (Floresco et al. 2008). This effect extends to shifts in attentional strategies. Lesions of the PL cortex disrupt extradimensional set shifting, where subjects are required to shift attention from one stimulus dimension (e.g., texture) to another dimension (e.g., odor) to forage effectively for reward (Birrell & Brown 2000).

Conversely, damage to the PL cortex does not disrupt behavior when there is a reversal in contingencies within the same attentional or strategic dimension, nor does it disrupt performance following an intradimensional set shift where subjects need to learn a new contingency within the same dimension (Birrell & Brown 2000, Boulougouris et al. 2007, Floresco et al. 2008, Ragozzino et al. 2003). Indeed, the deficit resulting from inactivity of PL neurons appears to be due to a perseveration of the response associated with the previous strategy or attentional set but not the previous response per se (Chudasama & Muir 2001, Sharpe & Killcross 2012). This is in contrast to OFC lesions, which disrupt changes in performance within a particular attentional set (i.e., an intradimensional set shift) (Chase et al. 2012), suggesting that subjects without an OFC cannot use an attentional set to constrain performance in the first place.

The PL cortex has also been found to modulate behavior in the presence of task-setting cues that are physically available in the environment. For example, the PL cortex is critical to the ability of subjects to both acquire and express context-dependent behaviors under conditions where different contexts signal different contingencies and warrant corresponding changes in response (Marquis et al. 2007; Sharpe & Killcross 2015a,b; Willcocks & McNally 2013; Zelikowsky et al. 2013). In one study, rats were trained with a contextual biconditional discrimination (Sharpe & Killcross 2015b). This design involved presenting rats with two stimuli in two visually different contexts. In one context, one stimulus (e.g., white noise) predicted delivery of shock, while presentation of another stimulus (e.g., clicker) did not result in delivery of shock. However, in the other context, these contingencies were reversed: The clicker predicted the shock, while the white noise was presented without shock. Rats learned these contingencies across time, and the stimuli elicited a fearful response in a context-dependent manner. Sharpe & Killcross (2015b) then showed that inactivation of the PL cortex during either acquisition or expression disrupted context-sensitive responding. That is, rats without PL function failed to use the contextual cues to disambiguate the stimuli and express fear to each stimulus in the context in which it predicted delivery of shock.

Note that this procedure, where the contingency shift is signaled by the present contextual cues, is critically different from a reversal task where the change in contingencies is not signaled. That is, in a contextual biconditional discrimination, subjects can use the contextual cues to dictate which contingencies are relevant on a particular trial, whereas a reversal task relies on the knowledge of recent successful actions to perform the correct response on the current trial. While the PL cortex is necessary for contextual cueing of a contingency shift in the former, the OFC has been implicated in the latter, where contextual cues are not available to disambiguate the correct response (consistent with the notion that the OFC encodes the “you are here” signal within the current state map, which is not needed when external stimuli cue the correct action).

The role of the PL cortex in exerting control over choice behavior when circumstances change also extends to scenarios without explicit task-setting cues or higher-order rules governing





contingencies. Indeed, activity in the PL cortex during learning is also necessary for choice behavior to reflect changes in the subjective value of the reward toward which subjects are working (Balleine & Dickinson 1998, Gläscher et al. 2008, Killcross & Coutureau 2003, O'Doherty et al. 2004, Ostlund & Balleine 2005, Tanaka et al. 2008, Tran-Tu-Yen et al. 2009). Lesions or inactivation of the PL cortex of rats trained to associate two instrumental responses with two different rewards render the rats' instrumental responses insensitive to changes in the value of the reward. That is, if one of the food rewards is devalued—by either giving rats free access to the food or pairing its consumption with illness—then rats without PL activity during the initial acquisition of the response will fail to adapt responses after devaluation and stop responding for the now devalued reward (Balleine & Dickinson 1998, Killcross & Coutureau 2003, Ostlund & Balleine 2005). Note that inactivation of the PL cortex after the rats have learned the associations does not disrupt the sensitivity of choice behavior to the current goal value (Tran-Tu-Yen et al. 2009), suggesting that the role for the PL cortex in acquisition of goal-directed behaviors is transient (but see Whitaker et al. 2017). However, in humans, activity in the ventromedial prefrontal cortex (which is analogous to the rodent PL cortex in terms of goal-directed behaviors; see **Figure 1**) (Balleine & O'Doherty 2010, Tanaka et al. 2008), encodes relative goal value across both the acquisition and testing phases of tasks that involve goal-directed behaviors (Gläscher et al. 2008, O'Doherty 2011, Tanaka et al. 2008). This suggests that activity in the PL cortex may still be relevant for the expression of goal-directed behaviors but is not necessary under certain conditions.

One influential theory of prefrontal function in cognitive neuroscience that can explain the contribution of the PL cortex in these different types of tasks is that advocated by Cohen and colleagues (Cohen et al. 1990, MacDonald et al. 2000, Miller & Cohen 2001). This theory originally emerged to explain the Stroop effect, which occurs when subjects are presented with color words written in different colored inks and asked to name the color of the ink. The Stroop effect refers to the increased reaction time to name the ink color of incongruent word-color pairs (e.g., the word GREEN written in red ink) relative to congruent word-color pairs (e.g., the word RED written in red ink). Cohen and colleagues account for this effect by arguing that subjects need to exert cognitive control in the incongruent case to suppress the prepotent tendency to read the word itself (the so-called word superiority effect; MacLeod & Dunbar 1988).

Cohen and colleagues (1990) used a neural network model to explain the Stroop effect. In the model, stimulus-response associations develop across a lifetime and lead individuals to associate words with utterances or responses that match the words (i.e., see the word GREEN, say “green”). Similarly, prior training has associated colors with the response of naming the color (i.e., see the color red, say “red”). The strength of each stimulus-response association, manifest in the weights between the relevant units in the network, is roughly proportional to the number of times that it has been acted out. In this example, the number of times that you have seen the word green and read the word “green,” relative to the number of times you have seen the color red and said or thought “red.” The strength of the stimulus-response association determines the ability of the stimulus to elicit the response. Hence, the stimulus of the word GREEN written in red ink elicits more strongly the tendency to read the word as “green” relative to the tendency to say “red.”

However, importantly, in the model, these stimulus-response associations involve an intermediary node (in an internal, hidden layer of the network). This hidden layer provides a substrate for other factors, such as cognitive control, to influence the level of activation of the association. Specifically, Cohen et al. (1990) argue that the prefrontal cortex can use the task demands to influence activation of specific stimulus-response pathways to facilitate production of a response that is appropriate for the current circumstances. For example, when the task demand is to name the color of the ink, the prefrontal cortex can directly influence the strength of activation of



the color-naming stimulus–response association by preactivating intermediate nodes that connect colors to their names, even before the stimulus appears. This extra activity in the hidden layer is sufficient to overcome the stronger associations between color words and the word-reading response.

The PL cortex may be the region within the prefrontal cortex that allows task demands to influence activation of the stimulus–response pathway and resolve response competition, as in the Cohen et al. (1990) model. To test this hypothesis, Haddon et al. (2008) developed a rodent variant of the Stroop task. In the task, rats were trained with two different contingencies in two different contexts. In one context (called the auditory context), rats received two auditory stimuli. One auditory stimulus (A1) predicted that a left lever-press response (R1) would produce a reward (O1), whereas a second auditory stimulus (A2) predicted that a right lever-press response (R2) would produce the same reward (O1). In another context—the visual context—rats received two visual stimuli (V1 and V2) that dictated that the left or right lever press, respectively, would be reinforced with another outcome (O2). Rats then received either congruent or incongruent compounds in an unrewarded session. In the congruent case, rats were placed in one context (e.g., the auditory context) and given an audio-visual compound where both elements of the compound predicted that the same lever-press response would be reinforced (e.g., A1V1). Thus, rats could just make the appropriate lever press associated with both elements (i.e., R1). However, in the incongruent case, rats were given an audio-visual compound that suggested that opposing responses would be rewarded (e.g., A1V2). Just as humans performing the Stroop task can use the current task demands to perform the accurate response, rats could use the current contextual cues to resolve the response conflict and choose the context-appropriate response (e.g., R1 in the auditory context). Furthermore, the ability of rats to accurately select the correct response was related to the amount of training that they had had with the contingencies in either context. For instance, if rats received much more training in the visual context, then they showed greater errors in lever-press responding to an incongruent compound where the correct response was dictated by the auditory stimulus (Haddon & Killcross 2011). This is analogous to findings in humans, where people are slower to respond when required to say the name of the color rather than to read the written word in the incongruent color–word pairs owing to greater experience with reading words (Cohen et al. 1990, MacLeod & Dunbar 1988).

The similarity of this rodent task to the Stroop task in humans allows the use of causal techniques to test whether the PL region is the site within the prefrontal cortex that is necessary for the exertion of cognitive control in the rodent. Initially, Haddon & Killcross (2006) tested whether lesions of the entire medial prefrontal cortex of the rat—including the PL cortex, as well as the infralimbic and anterior cingulate cortices—impaired performance in this task. Indeed, they found that, while these rats could learn each of the contingencies in each context well, they failed to use the contextual cues to resolve response conflict in the incongruent trials in the test session. This was despite performing normally on the congruent compounds. Furthermore, this was not due to an inability to distinguish the contexts: Following testing on the Stroop task, rats received free access to the outcome associated with one of the contexts (e.g., O1 associated with the auditory context). Their food magazine entries in both contexts was then assessed. Like animals in the control group, rats with lesions of the medial prefrontal cortex were capable of showing a selective reduction of magazine entries in the context associated with the devalued reward relative to the context associated with the still-valued reward, demonstrating that they had learned about the contextual cues and associated the different outcomes with the two different contexts. Thus, the role of the medial prefrontal cortex in this rodent Stroop task was specific to using contextual cues in a top-down manner to influence the appropriate choice of response. Marquis et al. (2007) later showed that the deficit seen with large lesions of the medial prefrontal cortex was in fact due



to damage to the PL cortex. Specifically, inactivating the PL cortex just prior to the compound test sessions produced the same deficit seen with large lesions; this was not the case with selective manipulation of the infralimbic or anterior cingulate cortices (Marquis et al. 2007).

These data provide strong evidence that the PL cortex provides top-down control over behavior, as theorized in models of prefrontal function (Cohen et al. 1990, MacDonald et al. 2000, Miller & Cohen 2001). That is, the PL cortex appears to be the region that allows rats to use higher-order cues to make a choice that reflects the current demands of the task. Conceptualizing PL function as synonymous with the role of the prefrontal cortex in the Cohen et al. (1990) framework allows for the reconciliation of work implicating the PL cortex in attentional set shifting, strategy shifting, and contextual modulation of choice behavior in both appetitive and aversive learning procedures (Sharpe & Killcross 2018). Accordingly, a role for the PL cortex in these behaviors can be commonly viewed as using higher-order information related to the demands of the task—such as attentional or strategy set or current contextual cues—to select a response that is appropriate to the current circumstance. Going back to our map analogy, while activity in the OFC may reflect the current position within the relevant cityscape, we might view the PL cortex as providing the cityscape context. That is, in this framework, the PL cortex may be considered critical for uploading the relevant cognitive map of the particular city that you are in, which details the potential routes though which you may travel, while the OFC tracks the movement on these routes.

We note that the role of the PL cortex in execution of goal-directed behaviors is probably not mediated by top-down modulation. More specifically, goal-directed behavior is commonly conceptualized as the result of a forward search through an associative structure comprising the response–outcome contingencies rather than top-down modulation of response by value of a goal (Balleine & Dickinson 1998, Daw et al. 2005, Dickinson & Balleine 2002). Thus, while the PL cortex plays an integral role in the development of the associative structures that support goal-directed behaviors, it may do so in a manner that is functionally distinct from its role in top-down modulation by context, strategy, or attentional set shifting. This distinction is supported by findings in humans that goal-directed behaviors and those regulated by contexts are mediated by distinct neural circuitry (see **Figure 1**). We discuss this further in Section 6.

## 6. HOW MIGHT THESE DISTINCT REGIONS WORK IN CONCERT TO INFLUENCE FLEXIBLE BEHAVIOR?

In the above sections, we provided a brief overview of the current literature on seemingly disparate neural systems. We began by discussing the role of the striatum in influencing choice behavior on an ongoing basis, with the DMS contributing to flexible behavior that is sensitive to the current environment, while the DLS facilitates the execution of automatic, habitual responses when a task is well learned. In contrast to the role of striatal regions in directly influencing the motor aspects of choice behavior, the hippocampus appears to be involved in forming associative relationships between events to facilitate a representation of the structure of the environment—at a conceptual as well as spatial level—regardless of whether these representations are critical to the current demands of the task. The OFC, in contrast, does not appear to represent all contingencies, but rather it selectively represents the contingencies that are relevant at the current moment, namely, the fine-grained state of the task within the current cognitive map at the current moment in time. Finally, the PL cortex is argued to be necessary for subjects to use higher-order cues in the environment—such as attentional set, strategy, or contextual cues—to exert top-down control over responses and resolve response competition when several courses of action could be appropriate for different task demands. Essentially, the PL cortex is important for providing or uploading the relevant cognitive map, which contains the contingencies appropriate to that circumstance.



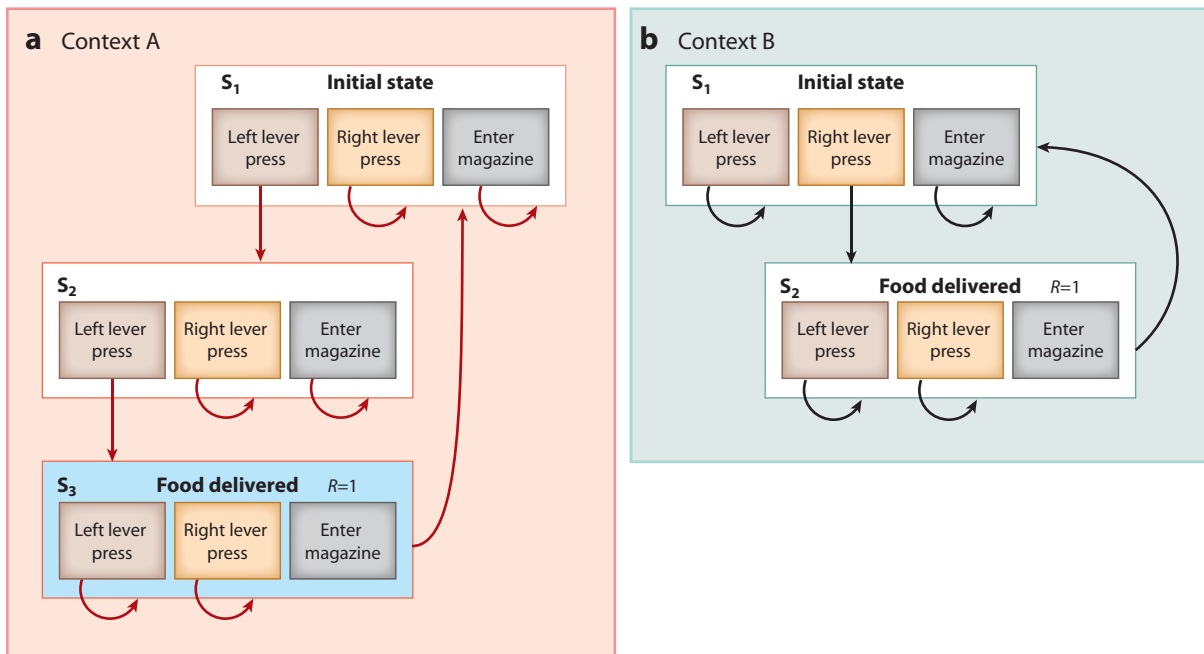
Do these systems act as separate nodes contributing to choice behavior through separable neural systems? Or can they be viewed as an integrative system that works together to influence flexible behavior? The existing literature appears to support the latter idea. For instance, both the OFC and the PL cortex receive information from the hippocampus that may be critical for accurately representing the current state (i.e., the current position within the cognitive map) (Wikenheiser et al. 2017) and higher-order task-set information that dictates the correct map for the current circumstance (Orsini et al. 2011, Zelikowsky et al. 2013), respectively. In turn, lesions of the OFC disrupt representation of the current state in the DMS (Stalnaker et al. 2016), and disconnection of PL cortex projections to the DMS impairs goal-directed behaviors (Hart et al. 2018), strategy shifting (Baker & Ragozzino 2014), and attentional modulation (Christakou et al. 2001). Taken together, this evidence indicates that one way to view these dissociable regions as an integrative system is to suppose that the hippocampus relays information about the general structure of the environment to both the OFC and the PL cortex, subsequently allowing these regions to influence striatal activity and directly impact choice behavior.

Consider a simple example where rats are trained to press two levers in two different contexts. In one context, context A, two left lever presses will produce reward [i.e., a fixed ratio 2 (FR2) schedule on the left lever], while in another context, context B, one right lever press will produce reward (i.e., an FR1 schedule on the right lever). A way to conceptualize this contextual task would be to integrate Cohen and colleagues' (1990) influential model of the Stroop task with model-based reinforcement learning (Daw et al. 2005). According to Daw et al. (2005), subjects trained on this set of contingencies would segregate the task into a series of consecutive states and learn, through experience, the action-dependent transitions between states in the task and the immediate reward values of each of the transitions (see **Figure 2**). In our example, in the first state ( $S_1$ ) in context A, pressing the left lever will transition the task to  $S_2$  (the state of the left lever having been pressed once), at which point pressing the left lever again will result in  $S_3$ , in which food is delivered to the food magazine. If the subject now chooses to enter the magazine they will retrieve food and transition back to  $S_1$  (see **Figure 2a**). However, if they press the right lever in either  $S_1$  or  $S_2$ , then food will not be delivered, and they will stay in their current state. When subjects are placed in context B and need to press the right lever only once, they learn a separate state transition diagram that describes the transition probabilities associated with that specific context (**Figure 2b**). Through experience and learning, the transition probabilities of each model of the environment come to reflect an accurate representation of the task, allowing the subject to flexibly switch between the models according to context and perform contextually appropriate responses.

In this framework, we would view the DMS as harboring the state transition diagram (often called the transition matrix) of the task. Critically, we suggest that the particular state transition diagram that is in effect in the DMS and the representation of the current location of the animal within it at each point in time are determined by frontal modulation. Thus, upon transitioning to a new state (e.g., due to the context changing or transitioning to another state within the current context), frontal regions can upload a particular state transition matrix—or influence representation of a particular state within the matrix—into the DMS to promote choice of a response that is suitable to the current moment in the particular environment.

Critically, we argue that the OFC and PL cortex influence different aspects of the representations in the DMS. We view the PL cortex as determining which state transition matrix is in effect in DMS through top-down modulation by higher-order information that is currently available. In our example above, the PL cortex would be necessary to increase activity related to the context-relevant state transition matrix, so that the state transitions trained in that context dominate behavior (e.g., context A; see **Figure 2a**). This might be envisioned at the neuronal level as priming a representation of all the states comprising the contextually relevant state transition





**Figure 2**

Proposed framework of the influence of the prelimbic (PL) cortex and the orbitofrontal cortex (OFC) on choice behavior. In this example, rats are first trained on (a) a contextual discrimination where placement in one context, context A, predicts that two left lever presses will yield reward [i.e., fixed ratio 2 (FR2)], while (b) another context, context B, will predict that one right lever presses will yield reward (i.e., FR1). As a consequence of this training, subjects develop a separate state transition diagram for each context (Daw et al. 2005). We propose that the PL cortex is necessary to activate the correct state transition matrix relevant to the current context in the dorsomedial striatum (DMS) (e.g., red, context A) through top-down modulation of activity in DMS neurons, in a manner similar to the modulation described by Miller & Cohen (2001). The OFC, in contrast, modulates DMS activity by identifying and highlighting the current state that the animal is in [e.g., S<sub>3</sub> (blue), a state that is accompanied by a reward of one food pellet ( $R = 1$ )], which can also be thought of as highlighting the relevant contingencies at a particular point within a trial. Importantly, the state represented in OFC is local—it is not the general context of the task (which we associate with PL cortex modulation); instead, the OFC representation changes within and across trials as the subject progresses through the task, transitioning from state to state. Together, these influences of the OFC and PL cortex therefore facilitate a state- and context-specific response that is appropriate for the current circumstance.

matrix (see **Figure 2a**). Such modulation would make it more likely that the response that is enacted is the action that is associated with reward delivery in that particular context. This result is consistent with the PL cortex playing a role in both the acquisition and the expression of context-specific behaviors (Orsini et al. 2011; Sharpe & Killcross 2015a,b; Zelikowsky et al. 2013) and would extend to the role of the PL cortex in other behaviors (e.g., attentional modulation and task switching) (Birrell & Brown 2000, Floresco et al. 2008, Ragozzino et al. 2003, Sharpe & Killcross 2012).

In addition to playing a role in switching between state transition diagrams, the PL cortex also appears to be necessary for the development of the state transition matrices in the DMS. More specifically, the role of the PL cortex in the development of goal-directed behavior implicates the PL cortex in the development of the action-dependent state transitions themselves (i.e., the probability of transitioning, upon performing some specific action, from one state to another, possibly accompanied by earning a reward) (Balleine & Dickinson 1998, Daw et al. 2005, Dickinson & Balleine 1994). Interestingly, data show that the PL cortex is necessary for basic acquisition of goal-directed behaviors but not their expression after learning has been completed (Hart et al.



2018, Ostlund & Balleine 2005, Smith et al. 2012, Tran-Tu-Yen et al. 2009). This suggests that the PL cortex entrains neuronal ensembles in the DMS to represent the state transition matrix early in training. However, once this representation is learned (and providing there is no change in other aspects of the environment that would warrant a top-down modulation to change the state diagram being represented, also dependent on the PL cortex), the DMS appears to be endowed with the capacity to perform a forward search through the state transition matrix and evaluate actions in relation to their future reward value to produce a response that is goal directed (that is, a response that is immediately sensitive to changes in the goal). Interestingly, while the PL cortex is not necessary for changing responses when goal values are altered after learning (e.g., through devaluation), it is necessary for responding to changes in action-outcome contingency (Balleine & Dickinson 1998). For instance, without PL function, subjects fail to stop responding when contingencies are changed such that the reward is delivered even in the absence of the response (Balleine & Dickinson 1998). This suggests a role for the PL cortex in calculation (or updating) of the transition probabilities within a particular state transition matrix even after initial learning, despite the state transition matrix being stored in the DMS after learning has taken place.

However, Schuck et al. (2016) and Wilson et al. (2014) have argued that the OFC influences choice behavior by indicating which state the subject is currently in within a particular model of the environment, which becomes critical when external stimuli are not sufficient for determining this. For instance, the OFC is necessary for determining whether the subject is currently in  $S_1$  or  $S_2$  in **Figure 2a**—two states that are externally identical (in both, there is no reward in the food magazine) but are at different places in the task diagram. Animals can differentiate such externally identical states (as indicated by, for instance, reaction times) and can even take different actions in each state if required (Mishkin et al. 1969). This influence of the OFC on state representation may not have significant consequences for behavior in simple tasks that do not require tracking of which state you are currently in within a particular cognitive map or in which the current state is explicitly signaled by environmental stimuli. However, when tasks become more complex (e.g., due to an unsignaled reversal of contingencies), the OFC is critical for performing the task optimally.

Support for the idea that the OFC and PL cortex play distinct roles in influencing task-relevant representations in the DMS is also provided by the different ways in which these regions influence DMS activity. For example, OFC modulation of state representation (the “you are here” within the cognitive map) in the DMS appears to occur through cholinergic interneurons, rather than directly on medium spiny neurons (MSNs), in line with recent work implicating cholinergic interneurons in representation of state (Apicella 2007, Bradfield et al. 2013a, Kimura et al. 2004). In one particularly elegant study, Bradfield et al. (2013a) demonstrated that cholinergic interneurons in the DMS are not necessary for flexible behavior during initial learning when a single set of contingencies determine the correct course of action. However, when these contingencies are shifted and subjects have to represent the two contingencies as separate cognitive states, cholinergic interneurons are necessary for exhibiting flexible behavior (Bradfield et al. 2013a).

Recently, Stalnaker et al. (2016) recorded activity in cholinergic interneurons during an instrumental task and indeed found evidence for the encoding of state. In this case, rats responded to two different odors that signaled varying levels of reward in one of two wells on any particular trial. Critically, the reward contingencies were changed periodically such that an odor would predict a high-value reward of a particular identity in one well during one trial block but a low-value reward of a different identity in another block. Stalnaker et al. (2016) found that activity in cholinergic interneurons encoded information that was relevant to the specific trial block in effect. Furthermore, state-relevant information was encoded across all events occurring within a trial, not only at the moment of choice. For example, state-relevant information in cholinergic





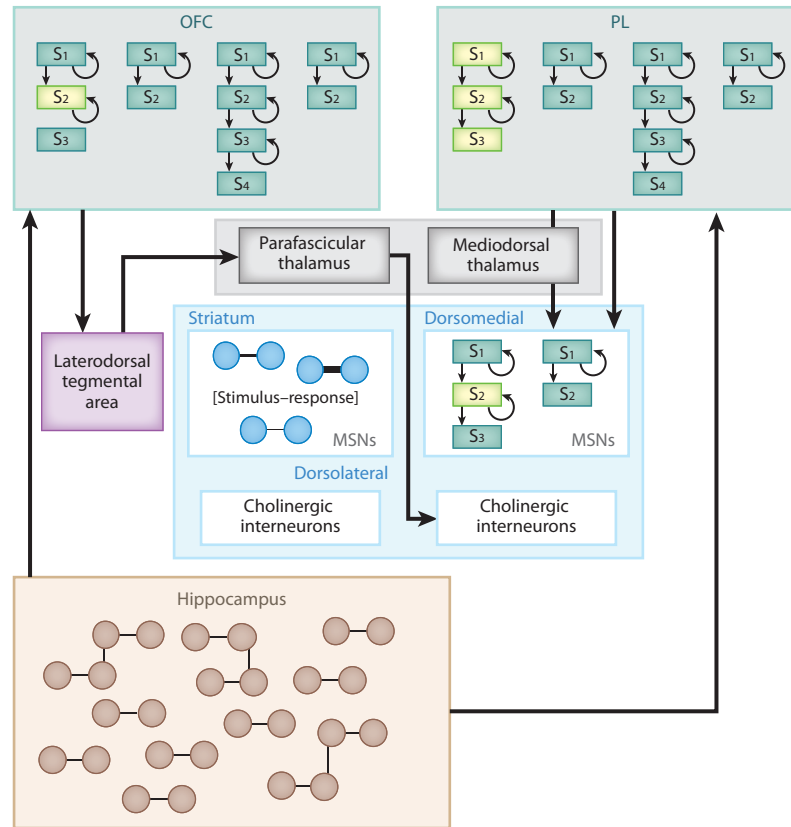
interneurons could be decoded during trial initiation, odor delivery, choice, reward delivery, or the signaled end of a trial and was maintained between these events. This is consistent with the encoding of a particular state that reflects the current contingencies available at a particular moment in time, rather than signaling a particular contingency associated, for instance, with the highest-valued reward at the choice point. Remarkably, miscoding of states in DMS cholinergic interneurons predicted inaccuracy in the task on a trial-by-trial basis, as seen in the OFC in humans (Schuck et al. 2016). That is, the coding of state in cholinergic interneurons predicted the behavior of the animals when they made an error on the behavioral task.

Input from the OFC was necessary for cholinergic interneurons to acquire state-specific activity. Lesions of the OFC significantly reduced the ability to decode state information from cholinergic interneurons during the task. Furthermore, the information that appeared to be lost without the OFC was that related to the unobservable aspects of a trial. That is, without OFC input, cholinergic interneurons did not contain information related to the current state outside of reward delivery. In fact, without OFC input, decoding of external stimuli (identity of odors and rewards) from cholinergic interneurons was better than it was in rats with intact OFC function. This indicates that the OFC is necessary for cholinergic interneurons to track the states through a trial in a manner that is not dependent on external stimuli, as suggested by the state-representation theory of OFC function (Wilson et al. 2014).

In contrast, the PL cortex modulates striatal activity through direct effects on MSNs in the striatum (Hart et al. 2018; Kasanetz et al. 2006, 2008; Surmeier et al. 2007). Specifically, PL neurons and DMS MSNs display correlated subthreshold depolarizations (i.e., up states), where stimulation or inhibition of PL afferents can produce or end, respectively, an up state in DMS MSNs. Interestingly, while recording state-specific activity in cholinergic interneurons, Stalnaker et al. (2016) found that the same pattern of activity relevant to state was not seen in MSNs. Instead, MSN activity appeared to be specific to the moment at which a choice was made (Stalnaker et al. 2016). This supports the direct relevance of MSN activity to choice behavior, consistent with the long literature characterizing the effects of these neurons on direct motor output through pallidal pathways (Alexander & Crutcher 1990, Cui et al. 2013, Grillner et al. 2005). In line with this characterization, the PL cortex may influence activity associated with a particular state diagram through direct modulation of MSNs to influence the current choice at the moment that it is made.

The proposed framework also fits with the physiological characteristics of this circuit. That is, the OFC and PL cortex send projections to the DMS through distinct routes (**Figure 3**). The PL region sends dense direct projections to the DMS (Hart et al. 2018, Voorn et al. 2004). In addition, the PL cortex sends indirect projections to the DMS through the mediodorsal thalamus (Vertes 2004), a region also critical for behavioral flexibility (Bradfield et al. 2013b, Corbit et al. 2003). In contrast, there are relatively sparse projections directly from the OFC to the DMS or indirectly through the mediodorsal thalamus (but see Renteria et al. 2018 for projections to the ventral DMS and potential indirect projections through the nucleus accumbens). Instead, an important candidate for the pathway through which the OFC can influence DMS activity is the parafascicular thalamus, a nucleus providing extensive input to the cholinergic interneurons in the DMS (Groenewegen & Berendse 1994, Lapper & Bolam 1992). In particular, the OFC sends dense projections to the laterodorsal tegmental area, which provides dense cholinergic input to the parafascicular thalamus (Cornwall et al. 1990, Groenewegen & Berendse 1994). This is a particularly interesting pathway given the current framework, as input from the parafascicular nucleus to DMS cholinergic interneurons has been shown to be necessary for rats to use states to compartmentalize old and new learning in the striatum, in line with the state representation theory (Bradfield & Balleine 2017, Bradfield et al. 2013a). The physiological characteristics of this circuit





**Figure 3**

A proposed circuit underlying the differential effects of the OFC and the PL cortex on action selection. We propose that the hippocampus relays information about the general structure of the environment to the OFC and PL cortex. In turn, the PL cortex is proposed to exert top-down control to arbitrate between different possible state diagrams (models; the currently relevant models are in lighter green) of the task through direct projections to the dorsomedial striatum, as well as indirect projections through the mediodorsal thalamus. The OFC, in contrast, projects to the dorsomedial striatum through the laterodorsal tegmental area and parafascicular thalamus to influence signaling of cholinergic interneurons. Accordingly, the OFC can modulate the activity of a particular state within a state diagram (*light green*). Together, the OFC and PL cortex allow the effect of the dorsomedial striatum on action selection to be specific to the current state of the environment. Abbreviations: MSN, medium spiny neuron; OFC, orbitofrontal cortex; PL, prelimbic; S<sub>1</sub>–S<sub>4</sub>, state 1–state 4.

therefore support the notion that the PL cortex and OFC exert effects on choice performance through distinct mechanisms.

Finally, the DLS does not seem to contain information that relates to models of the task as seen in the DMS, but rather contains so-called model-free action propensities (Daw et al. 2005). This is consistent with the absence of projections from prefrontal regions to the DLS (Voorn et al. 2004), given that we argue that these projections are critical for training and modulating the state model. As others have argued, the DLS likely contains information that reflects the absolute strength of a response in terms of the number of times it has been acted out (Balleine et al. 2007, Jog et al. 1999, Miller et al. 2017b). Such Thorndikean associations are typically referred to as habits and are relatively automatic and not sensitive to changes in the current

environment (Balleine & Dickinson 1998, Miller et al. 2017b, Tricomi et al. 2009, Yin et al. 2006). According to these theories, with extensive training, the strongest response in the DLS acquires greater influence over behavior regardless of the current model in the DMS, in line with the development of rigid habitual behaviors with increasing experience on a task (Dickinson & Balleine 1994, Jog et al. 1999, Killcross & Coutureau 2003). Essentially, these two competitive systems would facilitate the trade-off between voluntary goal-directed behavior, supported by the development of state transition matrices within the DMS, and behavior that reflects automatic performance of the response that has reliably lead to reward in the past, consistent with activity in the DLS.

## 7. CONCLUSIONS

In this review, we propose an integrated framework of how the OFC and PL cortex exert control over choice behavior. Specifically, we argue that both the OFC and the PL cortex receive input from the hippocampus about the current scenario. In turn, the OFC and PL cortex use this information to modulate activity in the DMS to select the appropriate course of action in that circumstance. Critically, the OFC and PL cortex influence choice through different circuits and to different ends. We propose that the PL cortex is required for using higher-order information to exert cognitive control over which state diagram is currently represented in the DMS. This is likely to occur through direct PL modulation of MSN activity. In addition, the PL cortex entrains the direct development of the state transition matrices in the DMS, which are relevant to goal-directed behaviors. The OFC, in turn, represents the currently occupied state within the relevant state diagram, becoming critical when states are aliased and can only be distinguished based on internal information and inference processes. This “you are here” signal facilitates a representation of the choices and contingencies relevant at the current moment in time, conveyed through DMS cholinergic interneurons. Specifically, OFC input through the parafascicular thalamus allows cholinergic interneurons to reflect state information across a task to influence a representation of choice in the DMS. Thus, parallel processing in the PL cortex and OFC ensures a response that is appropriate to both the particular state space currently relevant to the task and the particular location within that state space. Future research would benefit from investigating whether the PL cortex influences state encoding in the OFC, how the OFC and PL cortex influence state encoding in the hippocampus, and whether a role for these regions in state encoding can be generalized to other regions receiving input from the OFC and PL cortex.

### FUTURE ISSUES

1. Does the PL cortex influence state encoding in the OFC? The current framework posits that the PL cortex is necessary for uploading the current state space on the basis of higher-order information present in the environment (e.g., contextual cues, attentional set). In contrast, the OFC is theorized to track the movement through the current state space to allow responses that are appropriate to the current position within the state space, even if this position is externally ambiguous. Information about the current relevant state space in the PL cortex may be relayed to the OFC to allow the OFC to track the position within the current state space. Effectively, the PL cortex may modulate activity in the OFC to allow the OFC to activate ensembles that are specific to the current state space. This hypothesis has yet to be tested directly. For example, would lesioning of the PL cortex disrupt the representation of the current state seen in the OFC cortex?



2. Does the OFC contribute to the development of the state space? The literature suggests that the PL cortex is also involved in the development of the state space. Specifically, disconnection of the PL from the DMS during acquisition of instrumental behavior prohibits subjects from exhibiting goal-directed behavior (theorized to result from a forward search through the transitions within the state space) (Daw et al. 2005). However, it may also be the case that the OFC is involved in the development of the state space. While lesions or inactivation of the OFC do not disrupt goal-directed behavior in simple instrumental settings (Ostlund & Balleine 2007), they disrupt goal-directed behavior in a Pavlovian setting (Gallagher et al. 1999). Thus, the OFC may contribute to the development of the state space under some circumstances. That is, the OFC may be involved in learning the state diagram, not only locating oneself within it after it has been learned. One way to test whether this is the case is to ask whether inactivation of the OFC during learning disrupts subsequent use of a state space in other tasks, and if so, how this influence differs from that of the PL cortex.
3. What is the significance of projections from the OFC and PL cortex to the hippocampus? We have advocated a role for the hippocampus in facilitating state encoding in the OFC and PL cortex by providing information about the general structure of the environment. Others have also suggested that the hippocampus contains candidate information that can be used to create a state space (Schuck et al. 2016). However, this is quite an impoverished view of the interactions among these regions. In particular, both the OFC and the PL cortex project densely to hippocampal regions. How might the OFC and PL cortex inform information coding in the hippocampus? Could the OFC and PL cortex influence representation of state in the hippocampus?
4. What other systems may benefit from state encoding in the OFC and PL cortex? Of course, many other systems receive input from the OFC and PL cortex. For example, both the OFC and the PL cortex project densely to the nucleus accumbens, the basolateral amygdala, and the lateral hypothalamus (Sharpe et al. 2017, Vertes 2004). This raises the question of whether these regions specifically receive information related to the encoding of state information to facilitate learning and behavior. Perhaps information processing in the OFC and PL cortex contributes to other aspects of learning that can be dissociated from the state encoding proposed in this framework.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

This work was supported by a National Health and Medical Research Council CJ Martin fellowship awarded to M.J.S., grant 1R01DA042065 from the National Institute on Drug Abuse at the National Institutes of Health (NIH) awarded to Y.N., and grant W911NF-14-1-0101 from the Army Research Office (ARO) awarded to Y.N. The opinions expressed in this article are the authors' own and do not reflect the view of the NIH, the US Department of Health and Human Services, or the ARO.



## LITERATURE CITED

- Alexander GE, Crutcher MD. 1990. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13:266–71
- Apicella P. 2007. Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends Neurosci.* 30:299–306
- Baker PM, Ragozzino ME. 2014. Contralateral disconnection of the rat prelimbic cortex and dorsomedial striatum impairs cue-guided behavioral switching. *Learn. Mem.* 21:368–79
- Balleine BW, Delgado MR, Hikosaka O. 2007. The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27:8161–65
- Balleine BW, Dickinson A. 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–19
- Balleine BW, O'Doherty JP. 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35(1):48–69
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. 2005. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437:1158–61
- Baunez C, Robbins T. 1999. Effects of dopamine depletion of the dorsal striatum and further interaction with subthalamic nucleus lesions in an attentional task in the rat. *Neuroscience* 92:1343–56
- Berger TW, Alger B, Thompson RF. 1976. Neuronal substrate of classical conditioning in the hippocampus. *Science* 192:483–85
- Birrell JM, Brown VJ. 2000. Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J. Neurosci.* 20:4320–24
- Boulougouris V, Dalley JW, Robbins TW. 2007. Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behav. Brain Res.* 179:219–28
- Bradfield LA, Balleine BW. 2017. Thalamic control of dorsomedial striatum regulates internal state to guide goal-directed action selection. *J. Neurosci.* 37:3721–33
- Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW. 2013a. The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. *Neuron* 79:153–66
- Bradfield LA, Hart G, Balleine BW. 2013b. The role of the anterior, mediodorsal, and parafascicular thalamus in instrumental conditioning. *Front. Syst. Neurosci.* 7:51
- Brasted P, Bussey T, Murray E, Wise S. 2003. Role of the hippocampal system in associative learning beyond the spatial domain. *Brain* 126:1202–23
- Brasted PJ, Wise SP. 2004. Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur. J. Neurosci.* 19:721–40
- Butter CM. 1969. Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in *Macaca mulatta*. *Physiol. Behav.* 4:163–71
- Butter CM, Mishkin M, Rosvold HE. 1963. Conditioning and extinction of a food-rewarded response after selective ablations of frontal cortex in rhesus monkeys. *Exp. Neurol.* 7:65–75
- Castañé A, Theobald DE, Robbins TW. 2010. Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behav. Brain Res.* 210:74–83
- Chase EA, Tait DS, Brown VJ. 2012. Lesions of the orbital prefrontal cortex impair the formation of attentional set in rats. *Eur. J. Neurosci.* 36:2368–75
- Christakou A, Robbins TW, Everitt BJ. 2001. Functional disconnection of a prefrontal cortical–dorsal striatal system disrupts choice reaction time performance: implications for attentional function. *Behav. Neurosci.* 115:812–25
- Chudasama Y, Bussey TJ, Muir JL. 2001. Effects of selective thalamic and prelimbic cortex lesions on two types of visual discrimination and reversal learning. *Eur. J. Neurosci.* 14:1009–20
- Chudasama Y, Muir JL. 2001. Visual attention in the rat: a role for the prelimbic cortex and thalamic nuclei? *Behav. Neurosci.* 115:417–28
- Clarke HF, Robbins TW, Roberts AC. 2008. Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J. Neurosci.* 28:10972–82



- Cohen JD, Dunbar K, McClelland JL. 1990. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97:332–61
- Corbit LH, Muir JL, Balleine BW. 2003. Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur. J. Neurosci.* 18:1286–94
- Corcoran KA, Desmond TJ, Frey KA, Maren S. 2005. Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *J. Neurosci.* 25:8978–87
- Cornwall J, Cooper J, Phillipson O. 1990. Afferent and efferent connections of the laterodorsal tegmental nucleus in the rat. *Brain Res. Bull.* 25:271–84
- Coyle JT, Snyder SH. 1969. Antiparkinsonian drugs: inhibition of dopamine uptake in the corpus striatum as a possible mechanism of action. *Science* 166:899–901
- Cui G, Jun SB, Jin X, Pham MD, Vogel SS, et al. 2013. Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* 494:238–42
- Davachi L, Wagner AD. 2002. Hippocampal contributions to episodic encoding: insights from relational and item-based learning. *J. Neurophysiol.* 88:982–90
- Daw ND, Niv Y, Dayan P. 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8:1704–11
- Devan BD, Hong NS, McDonald RJ. 2011. Parallel associative processing in the dorsal striatum: segregation of stimulus–response and cognitive control subregions. *Neurobiol. Learn. Mem.* 96:95–120
- Dickinson A, Balleine B. 1994. Motivational control of goal-directed action. *Anim. Learn. Behav.* 22:1–18
- Dickinson A, Balleine B. 2002. The role of learning in the operation of motivational systems. In *Stevens' Handbook of Experimental Psychology*, ed. CR Gallistel, pp. 497–534. Hoboken, NJ: Wiley
- Dusek JA, Eichenbaum H. 1997. The hippocampus and memory for orderly stimulus relations. *PNAS* 94:7109–14
- Eagle D, Robbins T. 2003. Inhibitory control in rats performing a stop-signal reaction-time task: effects of lesions of the medial striatum and d-amphetamine. *Behav. Neurosci.* 117:1302–17
- Eagle DM, Wong JC, Allan ME, Mar AC, Theobald DE, Robbins TW. 2011. Contrasting roles for dopamine D1 and D2 receptor subtypes in the dorsomedial striatum but not the nucleus accumbens core during behavioral inhibition in the stop-signal task in rats. *J. Neurosci.* 31:7349–56
- Eichenbaum H. 2000. Hippocampus: mapping or memory? *Curr. Biol.* 10: R785–87
- Emmons EB, De Corte BJ, Kim Y, Parker KL, Matell MS, Narayanan NS. 2017. Rodent medial frontal control of temporal processing in the dorsomedial striatum. *J. Neurosci.* 37:8718–33
- Floresco SB, Block AE, Maric T. 2008. Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behav. Brain Res.* 190:85–96
- Fuchs RA, Branham RK, See RE. 2006. Different neural substrates mediate cocaine seeking after abstinence versus extinction training: a critical role for the dorsolateral caudate–putamen. *J. Neurosci.* 26:3584–88
- Gallagher M, McMahan RW, Schoenbaum G. 1999. Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* 19:6610–14
- Gardner MP, Conroy JS, Shaham MH, Styer CV, Schoenbaum G. 2017. Lateral orbitofrontal inactivation dissociates devaluation-sensitive behavior and economic choice. *Neuron* 96:1192–203.e4
- Gershman SJ, Blei DM, Niv Y. 2010. Context, learning, and extinction. *Psychol. Rev.* 117:197–209
- Gisquet-Verrier P, Delatour B. 2006. The role of the rat prelimbic/infralimbic cortex in working memory: not involved in the short-term maintenance but in monitoring and processing functions. *Neuroscience* 141:585–96
- Gläscher J, Hampton AN, O'Doherty JP. 2008. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb. Cortex* 19:483–95
- Graybiel AM. 2005. The basal ganglia: learning new tricks and loving it. *Curr. Opin. Neurobiol.* 15:638–44
- Grillner S, Hellgren J, Menard A, Saitoh K, Wikström MA. 2005. Mechanisms for selection of basic motor programs: roles for the striatum and pallidum. *Trends Neurosci.* 28:364–70
- Groenewegen HJ, Berendse HW. 1994. The specificity of the “nonspecific” midline and intralaminar thalamic nuclei. *Trends Neurosci.* 17:52–57
- Haddon JE, George D, Killcross S. 2008. Contextual control of biconditional task performance: evidence for cue and response competition in rats. *Q. J. Exp. Psychol.* 61:1307–20





- Haddon JE, Killcross S. 2006. Prefrontal cortex lesions disrupt the contextual control of response conflict. *J. Neurosci.* 26:2933–40
- Haddon JE, Killcross S. 2011. Inactivation of the infralimbic prefrontal cortex in rats reduces the influence of inappropriate habitual responding in a response-conflict task. *Neuroscience* 199:205–12
- Hart G, Bradfield LA, Balleine BW. 2018. Prefrontal corticostriatal disconnection blocks the acquisition of goal-directed action. *J. Neurosci.* 38:1311–22
- Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, Averbeck BB. 2015. The role of frontal cortical and medial-temporal lobe brain areas in learning a Bayesian prior belief on reversals. *J. Neurosci.* 35:11751–60
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. 1999. Building neural representations of habits. *Science* 286:1745–49
- Kasanetz F, Riquelme LA, Della-Maggiore V, O'Donnell P, Murer MG. 2008. Functional integration across a gradient of corticostriatal channels controls UP state transitions in the dorsal striatum. *PNAS* 105:8124–29
- Kasanetz F, Riquelme LA, O'Donnell P, Murer MG. 2006. Turning off cortical ensembles stops striatal UP states and elicits phase perturbations in cortical and striatal slow oscillations in rat in vivo. *J. Physiol.* 577:97–113
- Killcross S, Coutureau E. 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13:400–8
- Kimura M, Minamimoto T, Matsumoto N, Hori Y. 2004. Monitoring and switching of cortico-basal ganglia loop functions by the thalamo-striatal system. *Neurosci. Res.* 48:355–60
- Kish SJ, Shannak K, Hornykiewicz O. 1988. Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. *N. Engl. J. Med.* 318:876–80
- Langston RF, Ainge JA, Couey JJ, Canto CB, Bjerknes TL, et al. 2010. Development of the spatial representation system in the rat. *Science* 328:1576–80
- Lapper S, Bolam J. 1992. Input from the frontal cortex and the parafascicular nucleus to cholinergic interneurons in the dorsal striatum of the rat. *Neuroscience* 51:533–45
- Liljeholm M, O'Doherty JP. 2012. Contributions of the striatum to learning, motivation, and performance: an associative account. *Trends Cogn. Sci.* 16:467–75
- MacDonald AW, Cohen JD, Stenger VA, Carter CS. 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288:1835–38
- MacLeod CM, Dunbar K. 1988. Training and Stroop-like interference: evidence for a continuum of automaticity. *J. Exp. Psychol. Learn. Mem. Cogn.* 14:126–35
- Mahut H, Zola-Morgan S, Moss M. 1982. Hippocampal resections impair associative learning and recognition memory in the monkey. *J. Neurosci.* 2:1214–20
- Marquis JP, Killcross S, Haddon JE. 2007. Inactivation of the prelimbic, but not infralimbic, prefrontal cortex impairs the contextual control of response conflict in rats. *Eur. J. Neurosci.* 25:559–66
- Matus-Amat P, Higgins EA, Barrientos RM, Rudy JW. 2004. The role of the dorsal hippocampus in the acquisition and retrieval of context memory representations. *J. Neurosci.* 24:2431–39
- McDonald RJ, White NM. 1993. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav. Neurosci.* 107:3–22
- McDonald RJ, White NM. 1995. Information acquired by the hippocampus interferes with acquisition of the amygdala-based conditioned-cue preference in the rat. *Hippocampus* 5:189–97
- McIntyre CK, Pal SN, Marriott LK, Gold PE. 2002. Competition between memory systems: acetylcholine release in the hippocampus correlates negatively with good performance on an amygdala-dependent task. *J. Neurosci.* 22:1171–76
- Miller EK, Cohen JD. 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24:167–202
- Miller KJ, Botvinick MM, Brody CD. 2017a. Dorsal hippocampus contributes to model-based planning. *Nat. Neurosci.* 20:1269–76
- Miller KJ, Shenhav A, Ludvig E. 2017b. Habits without values. bioRxiv 067603
- Milner B, Corkin S, Teuber H-L. 1968. Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of HM. *Neuropsychologia* 6:215–34



- Mishkin M, Vest B, Waxler M, Rosvold HE. 1969. A re-examination of the effects of frontal lesions on object alternation. *Neuropsychologia* 7:357–63
- Miyachi S, Hikosaka O, Miyashita K, Kárádi Z, Rand MK. 1997. Differential roles of monkey striatum in learning of sequential hand movement. *Exp. Brain Res.* 115:1–5
- Morris R, Garrud P, Rawlins JA, O'Keefe J. 1982. Place navigation impaired in rats with hippocampal lesions. *Nature* 297:681–83
- Murray EA, O'Doherty JP, Schoenbaum G. 2007. What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J. Neurosci.* 27:8166–69
- Naumann M, Pirker W, Reiners K, Lange KW, Becker G, Brücke T. 1998. Imaging the pre- and postsynaptic side of striatal dopaminergic synapses in idiopathic cervical dystonia: a SPECT study using [123I] epidepride and [123I]  $\beta$ -CIT. *Mov. Disord.* 13:319–23
- Neubert F-X, Mars RB, Thomas AG, Sallet J, Rushworth MF. 2014. Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81:700–13
- Nieuwenhuis S, Heslenfeld DJ, von Geusau NJA, Mars RB, Holroyd CB, Yeung N. 2005. Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage* 25:1302–9
- O'Doherty JP. 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Ann. N. Y. Acad. Sci.* 1239:118–29
- O'Doherty JP, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–54
- O'Keefe J, Nadel L. 1978. *The Hippocampus as a Cognitive Map*. Oxford, UK: Clarendon Press
- O'Keefe J, Speakman A. 1987. Single unit activity in the rat hippocampus during a spatial memory task. *Exp. Brain Res.* 68:1–27
- Orsini CA, Kim JH, Knapska E, Maren S. 2011. Hippocampal and prefrontal projections to the basal amygdala mediate contextual regulation of fear after extinction. *J. Neurosci.* 31:17269–77
- Ostlund SB, Balleine BW. 2005. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.* 25:7763–70
- Ostlund SB, Balleine BW. 2007. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J. Neurosci.* 27:4819–25
- Padoa-Schioppa C, Assad JA. 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–26
- Palminter RD. 2008. Dopamine signaling in the dorsal striatum is essential for motivated behaviors. *Ann. N. Y. Acad. Sci.* 1129:35–46
- Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, et al. 2016. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nat. Neurosci.* 19:845–54
- Parkes SL, Ravassard PM, Cerpa J-C, Wolff M, Ferreira G, Coutureau E. 2017. Insular and ventrolateral orbitofrontal cortices differentially contribute to goal-directed behavior in rodents. *Cereb. Cortex* 28:2313–25
- Quinlan MG, Hussain D, Brake WG. 2008. Use of cognitive strategies in rats: the role of estradiol and its interaction with dopamine. *Horm. Behav.* 53:185–91
- Ragozzino ME, Kim J, Hassert D, Minniti N, Kiang C. 2003. The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. *Behav. Neurosci.* 117:1054–65
- Ragozzino ME, Mohler EG, Prior M, Palencia CA, Rozman S. 2009. Acetylcholine activity in selective striatal regions supports behavioral flexibility. *Neurobiol. Learn. Mem.* 91:13–22
- Ragozzino ME, Ragozzino KE, Mizumori SJ, Kesner RP. 2002. Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav. Neurosci.* 116:105–15
- Redish AD. 1999. *Beyond the Cognitive Map: From Place Cells to Episodic Memory*. Cambridge, MA: MIT Press
- Renteria R, Baltz ET, Gremel CM. 2018. Chronic alcohol exposure disrupts top-down control over basal ganglia action selection to produce habits. *Nat. Commun.* 9:211
- Rudebeck PH, Murray EA. 2008. Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *J. Neurosci.* 28:8338–43
- Rudebeck PH, Saunders RC, Prescott AT, Chau LS, Murray EA. 2013. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat. Neurosci.* 16:1140–45



- Rudy JW, Sutherland RJ. 1989. The hippocampal formation is necessary for rats to learn and remember configural discriminations. *Behav. Brain Res.* 34:97–109
- Samejima K, Ueda Y, Doya K, Kimura M. 2005. Representation of action-specific reward values in the striatum. *Science* 310:1337–40
- Schendan HE, Searl MM, Melrose RJ, Stern CE. 2003. An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37:1013–25
- Schoenbaum G, Chiba AA, Gallagher M. 2000. Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. *J. Neurosci.* 20:5179–89
- Schoenbaum G, Eichenbaum H. 1995. Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. *J. Neurophysiol.* 74:733–50
- Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M. 2003. Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learn. Mem.* 10:129–40
- Schuck NW, Cai MB, Wilson RC, Niv Y. 2016. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron* 91:1402–12
- Schultz W, Stauffer WR, Lak A. 2017. The phasic dopamine signal maturing: from reward via behavioural activation to formal economic utility. *Curr. Opin. Neurobiol.* 43:139–48
- Sharpe MJ, Killcross S. 2012. The prelimbic cortex contributes to the down-regulation of attention toward redundant cues. *Cereb. Cortex* 24:1066–74
- Sharpe MJ, Killcross S. 2015a. The prelimbic cortex uses contextual cues to modulate responding towards predictive stimuli during fear renewal. *Neurobiol. Learn. Mem.* 118:20–29
- Sharpe MJ, Killcross S. 2015b. The prelimbic cortex uses higher-order cues to modulate both the acquisition and expression of conditioned fear. *Front. Syst. Neurosci.* 8:235
- Sharpe MJ, Killcross S. 2018. The modulation of attention and action in the medial prefrontal cortex of rats. *Psychol. Rev.* In press
- Sharpe MJ, Marchant NJ, Whitaker LR, Richie CT, Zhang YJ, et al. 2017. Lateral hypothalamic GABAergic neurons encode reward predictions that are relayed to the ventral tegmental area to regulate learning. *Curr. Biol.* 27:2089–100.e5
- Sharpe MJ, Schoenbaum G. 2016. Back to basics: making predictions in the orbitofrontal–amygdala circuit. *Neurobiol. Learn. Mem.* 131:201–6
- Smith KS, Virkud A, Deisseroth K, Graybiel AM. 2012. Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *PNAS* 109:18932–37
- Stalnaker TA, Berg B, Aujla N, Schoenbaum G. 2016. Cholinergic interneurons use orbitofrontal input to track beliefs about current state. *J. Neurosci.* 36:6242–57
- Stalnaker TA, Cooch NK, Schoenbaum G. 2015. What the orbitofrontal cortex does not do. *Nat. Neurosci.* 18:620–27
- Surmeier DJ, Ding J, Day M, Wang Z, Shen W. 2007. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci.* 30:228–35
- Sutton RS, Barto AG. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press
- Tanaka SC, Balleine BW, O’Doherty JP. 2008. Calculating consequences: brain systems that encode the causal effects of actions. *J. Neurosci.* 28:6750–55
- Thorn CA, Atallah H, Howe M, Graybiel AM. 2010. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66:781–95
- Tran-Tu-Yen DA, Marchand AR, Pape JR, Di Scala G, Coutureau E. 2009. Transient role of the rat prelimbic cortex in goal-directed behaviour. *Eur. J. Neurosci.* 30:464–71
- Tricomi E, Balleine BW, O’Doherty JP. 2009. A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29:2225–32
- Uylings HB, Groenewegen HJ, Kolb B. 2003. Do rats have a prefrontal cortex? *Behav. Brain Res.* 146:3–17
- Vertes RP. 2004. Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse* 51:32–58
- Volkow ND, Wang GJ, Fowler JS, Logan J, Jayne M, et al. 2002. “Nonhedonic” food motivation in humans involves dopamine in the dorsal striatum and methylphenidate amplifies this effect. *Synapse* 44:175–80
- Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM. 2004. Putting a spin on the dorsal–ventral divide of the striatum. *Trends Neurosci.* 27:468–74



- Wallenstein GV, Hasselmo ME, Eichenbaum H. 1998. The hippocampus as an associator of discontiguous events. *Trends Neurosci.* 21:317–23
- Wallis JD. 2012. Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.* 15:13–19
- Whitaker LR, Warren BL, Venniro M, Harte TC, McPherson KB, et al. 2017. Bidirectional modulation of intrinsic excitability in rat prelimbic cortex neuronal ensembles and non-ensembles after operant learning. *J. Neurosci.* 37:8845–56
- Wikenheiser AM, Marrero-Garcia Y, Schoenbaum G. 2017. Suppression of ventral hippocampal output impairs integrated orbitofrontal encoding of task structure. *Neuron* 95:1197–207.e3
- Wikenheiser AM, Redish AD. 2015. Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* 18:289–94
- Wikenheiser AM, Schoenbaum G. 2016. Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat. Rev. Neurosci.* 17:513–23
- Willcocks AL, McNally GP. 2013. The role of medial prefrontal cortex in extinction and reinstatement of alcohol-seeking in rats. *Eur. J. Neurosci.* 37:259–68
- Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014. Orbitofrontal cortex as a cognitive map of task space. *Neuron* 81:267–79
- Xu L, Anwyl R, Rowan MJ. 1998. Spatial exploration induces a persistent reversal of long-term potentiation in rat hippocampus. *Nature* 394:891–94
- Yin HH, Knowlton BJ, Balleine BW. 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19:181–89
- Yin HH, Knowlton BJ, Balleine BW. 2005. Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22:505–12
- Yin HH, Knowlton BJ, Balleine BW. 2006. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behav. Brain Res.* 166:189–96
- Yin HH, Mulcare SP, Hilário MR, Clouse E, Holloway T, et al. 2009. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat. Neurosci.* 12:333–41
- Zelikowsky M, Bissiere S, Hast TA, Bennett RZ, Abdipranoto A, et al. 2013. Prefrontal microcircuit underlies contextual learning after hippocampal loss. *PNAS* 110:9938–43
- Znamenskiy P, Zador AM. 2013. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497:482–85

