Received: 24 September 2019 | Revised: 15 December 2019 | Accepted: 18 December 2019

DOI: 10.1002/jnr.24581

REVIEW

Interfacing behavioral and neural circuit models for habit formation

Talia N. Lerner

Department of Physiology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Correspondence
Talia Lerner, Department of Physiology, Northwestern University Feinberg School of Medicine, 303 E Chicago Ave, Ward 5-120, Chicago, IL 60611, USA.
Email: talia.lerner@northwestern.edu

Funding information
Brain and Behavior Research Foundation; National Institute of Mental Health, Grant/Award Number: DP2MH122401 and R00 MH109569

Abstract

Habits are an important mechanism by which organisms can automate the control of behavior to alleviate cognitive demand. However, transitions to habitual control are risky because they lead to inflexible responding in the face of change. The question of how the brain controls transitions into habit is thus an intriguing one. How do we regulate when our repeated actions become automated? When is it advantageous or disadvantageous to release actions from cognitive control? Decades of research have identified a variety of methods for eliciting habitual responding in animal models. Progress has also been made to understand which brain areas and neural circuits control transitions into habit. Here, I discuss existing research on behavioral and neural circuit models for habit formation (with an emphasis on striatal circuits), and discuss strategies for combining information from different paradigms and levels of analysis to prompt further progress in the field.

KEYWORDS

dopamine, goal-directed behavior, habit, striatum

1 | INTRODUCTION

How does the brain control behavior? Some actions are goal directed: we imagine the consequences of particular choices and take careful measures to ensure good, cost-efficient outcomes to our actions. Other actions are habitual: we respond to familiar situations by relying on established routines and practiced skills. Both of these goal-directed and habitual strategies may be useful for survival, depending on context. Automating a subset of routine behaviors by creating habits allows fast, efficient responding without significant cognitive demand, but leads to inflexible responding in the face of change (Dickinson, 1985; Packard & Knowlton, 2002; Yin & Knowlton, 2006). Thus, the brain must decide when it is appropriate to create habits from repeated actions and when it is more advantageous to stay goal directed. Importantly, not every brain will balance between goal-directed and habitual control in the same way: individual differences in habit learning rates may contribute to a variety of individual differences in reward-seeking strategies, and may also contribute to an individual's risk for disorders such as drug addiction (George & Koob, 2017). Habit formation is thus a key area for further study, to better understand how we use habits to navigate our daily lives and how we can manipulate habit formation circuits to mitigate disease risk and treat existing patients.

Maladaptive habit formation mechanisms have been hypothesized to contribute to a variety of neuropsychiatric problems, including obsessive compulsive disorder (OCD), autism, and drug addiction (Alvares, Balleine, Whittle, & Guastella, 2016; Everitt & Robbins, 2016; Gillan et al., 2014). While these disorders are distinct from each other when considered as a whole, they share the characteristic that problematic behavioral sequences are repeatedly executed and are difficult to inhibit. However, the exact contributions of habit per se to the particular symptoms of each disorder remain unclear. For example, in the context of drug addiction, it has been observed that habit-associated brain areas become engaged in drug seeking with

Edited by Joshua Plotkin. Reviewed by Jacqueline Barker, Marc Fuccillo, and Christina Gremel.

The peer review history for this article is available at https://publons.com/publon/10.1002/jnr.24581.

extended training (Everitt & Robbins, 2016) and that this engagement of habit areas is not necessary: constantly solving for new action–outcome contingencies to receive drug reward, which prevents habit formation, preserves many characteristics of drug addiction in a rodent model (including escalating use and punishment-resistant drug seeking; Singer, Fadanelli, Kawa, & Robinson, 2018). Thus, the role of habits in drug addiction has been questioned. Does it play a role in some aspects of drug addiction? Perhaps in some individuals but not others?

In fact, to determine how dysfunction of the habit system contributes to the development of a brain disorder such as addiction, we need two major things. First, we must better formalize how habits are defined *behaviorally*. As detailed below, many studies of habit use different methodologies, and while the tasks used may all be related to one another, there are also potentially important differences. These differences do not need to be erased, but understood and related to each other. In other words, we should take care not to artificially narrow our view of habit in pursuit of a clean definition; rather, the goal should be to understand how the primary features of habit contribute to many varied circumstances. Second, we must develop a circuit model for how habitual behavior is produced, such that the statement that "habit circuits" are engaged or disrupted is meaningful across analyses. One way to answer the question of whether habit is involved in controlling a behavior is with a behavioral probe such as outcome devaluation. Another way to connect across behavioral paradigms would be to ask whether similar neural circuits are engaged by related, putatively habit-inducing tasks. Below, I summarize knowledge and progress on these two issues, with a view as to how the field can proceed to develop a better interface between behavioral and circuit-level models of habit.

Significance

This article gives an overview of how habits have been conceptualized and studied behaviorally. It also reviews findings and hypotheses about the neural implementation of habits, with an emphasis on striatal circuits. The aim is to integrate discussions of behavioral and circuit-level approaches to the study of habit, and to motivate new research directions at the interface between these levels of investigation.

2 | TASKS TO PROBE HABITS IN ANIMAL MODELS

Colloquially, habits are simply actions that are performed regularly and are resistant to change, a definition which influences our intuitive understanding of habit and our communications with the public on the findings of our research about habits. Scientifically, however, habits have a narrower, more specific definition. A habit is developed when a stimulus–response association is formed. The stimulus is a familiar sensory cue or environmental context, which then triggers a responsive action without consideration of the expected outcome of that action and/or without consideration of the value of the action's outcome to the animal. Thus, habitual actions are performed automatically, even when they appear to be maladaptive.

When a habitual action does produce an adaptive outcome, it can be difficult to determine that the action was produced by force of habit rather than by goal-directed control. But under the stimulus–response definition of habit, one can test for habitual behavior by creating a situation in which habitual and goal-directed control systems will differ in the actions they produce. Generally, experimenters do this by manipulating the value of an outcome or by manipulating the action–outcome contingency within a task. Both approaches to probing for habit have been employed frequently in the literature, with variations in how the action–outcome contingency or outcome value is manipulated. As other reviews in this issue rightly point out, the probes chosen to evaluate habitual behavior can significantly influence study outcomes and interpretations (Schreiner et al., 2019; Woon et al., 2019). Table 1 summarizes the most common approaches that have been taken to probe habit. Some tests manipulate outcome values: they reduce the animals' motivation for the outcome (satiety-specific devaluation, in which an animal is prefed a reinforcer to reduce its drive to obtain the particular reinforcer) or they induce a negative valence to the outcome (LiCl pairing, in which an animal learns to associate a previously palatable reinforcer with malaise). Other tests manipulate the action–outcome contingency. Omission probes reverse the contingency of actions and outcomes, requiring animals to withhold their responding to earn rewards. Contingency degradation delivers rewards regardless of responding. Since both of these action–outcome contingency manipulations may

TABLE 1 Methods for probing habit formation

Probe type	Variable manipulated	Selected references
Satiety-specific devaluation	Outcome value	DeRusso et al. (2010), Gremel et al. (2016), Gremel and Costa (2013), Vandaele, Pribut, and Janak (2017), Yin, Knowlton, and Balleine (2005), Yin, Ostlund, Knowlton, and Balleine (2005)
LiCl taste aversion devaluation	Outcome value	Smith, Virkud, Deisseroth, and Graybiel (2012), Vandaele et al. (2017), Yin, Knowlton, and Balleine (2004)
Omission	Action–outcome contingency	DeRusso et al. (2010), Rossi and Yin (2012), Yu, Gupta, Chen, and Yin (2009)
Contingency degradation	Action–outcome contingency	Gourley, Olevska, Gordon, and Taylor (2013), Vandaele et al. (2017), Yin, Ostlund, et al. (2005)

read out slightly different aspects of behavioral flexibility, it is imperative to closely examine the methods used to measure habit in the existing literature on habit formation.

Another important issue is that under this methodology of identifying habits, habits are simply the *impairment* of goal-directed behavior. A goal-directed behavior should be responsive to both action–outcome contingency changes and changes in outcome value (Dickinson & Balleine, 1994), so the loss of either is used as evidence for habit. However, this definition of habit may be problematic. Using the probes described in Table 1, it is impossible to determine whether an apparent "habit" is the result of a strengthened stimulus–response association or a weakening of goal-directed mechanisms (Vandaele & Janak, 2018). Additionally, it has been argued that habitual and goal-directed control mechanisms operate in a hierarchical organization or in parallel rather than being mutually exclusive (Dezfouli & Balleine, 2013; Lee, Shimojo, & O'Doherty, 2014). In this case, traditional probe tasks would fail to capture important dynamics of the system. The use of so-called two-step decision tasks to assess model-free versus model-based learning is one attempt to simultaneously and nonexclusively measure contributions from habitual and goal-directed control systems to behavioral output (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Daw, Niv, & Dayan, 2005), but there is still substantial disagreement in the field about whether model-free and model-based learning map onto habitual versus goal-directed behavior as elicited by more traditional operant tasks and probe tests. Indeed, alternative computational frameworks to explain habit have recently been proposed (Miller, Shenhav, & Ludvig, 2019). Thus, the question of whether the two-step task can reliably measure habitual behavior is, for the moment, reserved. The traditional operant tasks used to elicit habit formation are discussed in the next section of this review.

2.1 | What kind of operant training induces habits?

The strategy that an animal uses to control its behavior is dependent (at least in part) on the external structure of the task it is asked to perform. A number of different tasks have been developed to elicit habit formation as measured by the probe approaches in Table 1. Random/variable ratio (RR) or random/variable interval (RI) schedules are the most commonly used. In an RR schedule, multiple responses (e.g., lever presses or nosepokes) are required for the subject to earn a reward. The exact number of responses, however, is variable. In an RI schedule, rewards are only available to be earned (by performing a lever press or nosepoke) after a certain period of time, which is variable. The subject must continue to respond to check if the response will be rewarded. In rodents, RI schedules are more effective than RR schedules at producing habits (Dickinson & Charnock, 1985; Dickinson, Nicholas, & Adams, 1983; Gremel & Costa, 2013; Yin & Knowlton, 2006). Both of these random schedules of reinforcement are in turn much more effective at eliciting habit than a fixed ratio schedule, where the relationship between the action and outcome is entirely predictable and stable (DeRusso et al., 2010). However, we have never fully understood why this should be true.

In fact, some recent work challenges the view that it is merely uncertainty which promotes habitual responding. Vandaele and colleagues demonstrated that a fixed ratio schedule (FR5) can in fact lead to rapid habit formation when it is bracketed by lever insertion and removal (Vandaele et al., 2017). This "discrete trials" version of the FR5 task, termed DT5, suggests that habit formation is accelerated by cued task bracketing, which seems in contrast to uncertainty. However, cues may help accelerate habit formation by creating clear stimuli for stimulus–response associations to form around, and by bracketing tasks into clearly defined action sequences. The fact that both an RI60 task and a DT5 task are effective at eliciting habitual responding raises the question of whether these apparently very differently structured tasks actually engage different circuit-level routes to habitual performance. *In vivo* recordings during these tasks may help to clarify, and will be discussed further in the second part of the review.

Some additional operant tasks have also been designed to elicit habits. In particular, Graybiel and colleagues have taken advantage of a T-maze task in multiple studies of habit (Kubota et al., 2009; Smith & Graybiel, 2016; Smith et al., 2012; Thorn, Atallah, Howe, & Graybiel, 2010). In the T-maze task, rats run down the long arm of a T-maze and are cued halfway down as to which direction they should turn at the end to receive reward. This task differs from a classic RI reinforcement schedule in important ways: animals are rewarded every time they make a correct decision (no uncertain waiting periods that induce high response rates), and they must perform a sensory discrimination to determine the correct decision for each trial. Nevertheless, after overtraining rats are unable to adjust their behavior after outcome devaluation, continuing to run down the T-maze and turn to the devalued side when instructed (Smith et al., 2012). Results from these studies are compelling and form a consistent body of literature, yet it remains unclear (as for the DT5 task) whether habitual performance in the T-maze is elicited via similar or different circuit-level mechanisms as habitual performance observed after RI60 training.

2.2 | Motor skill learning as habit

Habit formation and motor skill learning are related to each other. They are often discussed in parallel, and sometimes conflated. Motor skill learning involves the chunking of action sequences into fluidly executed motions requiring minimal cognitive engagement. The ability to learn new motor skills depends on similar brain areas as habit formation. For example, skill learning in mice on an accelerating rotarod test depends on dopamine-dependent shifts in encoding between the dorsomedial and dorsolateral striatum (DLS) (Yin et al., 2009), similar to the shifts in encoding observed as habitual performance emerges during operant training (Yin et al., 2004; Yin, Knowlton, et al., 2005; Yin, Ostlund, et al., 2005), which is also dependent on dopamine (Faure, Haberland, Condé, & Massioui, 2005). Learning on the accelerating rotarod has also been used to model acquired repetitive behaviors in mouse models of autism, which

have coincident changes in striatal circuitry (Rothwell et al., 2014). In human patients with Parkinson's disease, in whom nigrostriatal dopamine signaling is impaired, there are deficits in new motor skill acquisition (Kawashima, Ueki, Kato, Ito, & Matsukawa, 2018) as well as in habit (Bannard et al., 2019; Knowlton, Mangels, & Squire, 1996; Witt, Nuhsman, & Deuschl, 2002).

Motor skill acquisition has also been assessed in rodents using a variety of skilled reaching tasks and fast, timing-dependent sequences of lever pressing (Jin & Costa, 2010; Jin, Tecuapetla, & Costa, 2014; Kawai et al., 2015; Xu et al., 2009). Basal ganglia circuit function and striatal dopamine inputs are again at the heart of these learned behaviors. Dopamine cells projecting to the dorsal striatum signal the beginning and end of learned action sequences, and help to control learned sequence-related activity in the striatum (Jin & Costa, 2010; Jin et al., 2014). Motor skill acquisition also stabilizes dendritic spines in motor cortex (Xu et al., 2009), but provocatively, motor cortex was found to be dispensable for the execution of a previously learned motor task, suggesting that subcortical circuits can independently support motor execution after learning has occurred (Kawai et al., 2015). Indeed, lesions of the DLS, which prevent habit formation (Yin et al., 2004), also prevent learned motor skill execution (Dhawale, Wolff, Ko, & Ölveczky, 2019).

Another model system in which motor skill learning has been investigated is songbirds. Songbirds have a specialized song learning circuit called the anterior forebrain pathway, which includes areas analogous to cortex, basal ganglia, and thalamus in mammals (Doupe, Perkel, Reiner, & Stern, 2005). In songbirds such as the zebra finch, song is a highly stereotyped and easily quantified motor output. The combination of such an elegant motor output with a brain circuit that is dedicated to producing it (and separated from the circuits controlling other movements) makes birdsong a very appealing system for studying the relationship between brain activity and behavior. From studies of birdsong we know that the song-related basal ganglia, and its dopaminergic inputs, are required for song learning to take place (Brainard & Doupe, 2000; Gadagkar et al., 2016). The implication is that this type of skill learning too may bear relationships with habit learning in mammals. Thus, advances in our understanding of how song production is controlled in the avian brain stand to inform many of our studies in mammals, including those involving habit.

Whether singing in birds or mice is a “habit” is not obvious. Despite the similarities in brain structures required for motor skill learning and habit formation, the relationship between these behaviors, especially as defined by performance in the probe tests listed in Table 1, remains to be formalized. Skills such as singing are performed in the absence of external rewards like sucrose pellets, but changes in behavior can be driven by sensory feedback and internal template matching, which also drives dopaminergic reward prediction error signals (Gadagkar et al., 2016). Eventually, skilled singing is rewarded by mating opportunities in the wild, but the behavior is learned well before mating occurs (e.g., given an appropriate tutor, male zebrafinch learn and crystalize their song around the same time they reach sexual maturity, ~90 days post hatching). In laboratory animals, actual mating may never occur as a consequence for singing, yet the behavior is still learned and performed. Thus, song learning is an interesting but perhaps exceptional context is which a motor skill is acquired due to an innate drive. Still, the study of song learning has provided important principles for motor learning more generally, such as the critical role of variability in motor performance to learning (Dhawale, Smith, & Ölveczky, 2017). The variability that drives motor learning appears to be created by basal ganglia circuits, which also support habit formation (Dhawale et al., 2017). What role does behavioral variability play in habit formation? And could that be a key to understanding why difference reward schedules promote it?

A primary difference between tests of motor skills and habits is timing. Motor skills generally involve precision of action on the millisecond timescale, whereas habits encoding relationships between lever pressing and reward retrieval from a separate reward delivery port involve learning about events separated by seconds. Whether there are common striatal and dopaminergic mechanisms capable of mediating both millisecond- and second-scale feedback to alter behavior, particularly transitions to habit, is largely unknown. One study in songbirds showed that birds are capable of learning from millisecond-scale auditory feedback; however, the authors did not explore whether a dopaminergic mechanism mediates that effect (Charlesworth, Tumer, Warren, & Brainard, 2011). In rodents, millisecond timescale dorsal striatal dopamine signals can bias animals toward changes in action, a plausible mechanism for inducing fast behavioral adaptations in a motor sequence in response to salient feedback (da Silva, Tecuapetla, Paixão, & Costa, 2018; Howe & Dombeck, 2016; Jin & Costa, 2010).

2.3 | Grooming behavior

Grooming is a repetitive behavior that mice perform spontaneously without training. It follows a stereotyped sequence, starting from the nose and working back across the face and body. Grooming meets our intuitive or colloquial definition of habit as a regularly performed behavior, and thus is often discussed in relation to habit. Grooming is also similar to the skilled motor tasks described above, requiring fine coordinated sequences of movement to execute. Like birdsong, it is a stereotyped behavior acquired early in life. But is grooming a habit in the formal psychological sense? Self-injurious overgrooming is observed in several mouse models of OCD and autism, contributing the hypothesis that habit plays a role in these disorders (Peça et al., 2011; Shmelkov et al., 2010; Welch et al., 2007). The fact that mice with OCD/autism-related mutations will continue to groom even when the behavior is apparently harmful does suggest a connection to habit: these mice seem unable to discontinue a behavior even when the action is leading to a maladaptive outcome.

In addition to spontaneous grooming, mice can be induced to groom when a water drop is applied to the head (Burguière, Monteiro, Feng, & Graybiel, 2013). In a mouse model of OCD (Sapap3 model; Welch et al., 2007), water-induced grooming behavior transitions

into additional spontaneous grooming bouts, providing support for the idea that repetitive behaviors in OCD are the result of exuberant habit formation that quickly disconnects actions from desired outcomes. However, to what degree grooming behavior is or is not related to other types of habit, such as learned operant behaviors and motor skills, is not well established.

2.4 | Compulsive behavior

Habits have been widely hypothesized to contribute to addiction. In particular, habits may contribute to compulsive drug seeking, usually defined as drug seeking in the face of negative consequences. While in humans the negative consequences of drug seeking typically involve the loss of money, jobs, and important social relationships, as well as negative long-term health effects, in animals these negative consequences are often modeled simply as electrical shocks. Other simple methods for modeling the negative consequences of drug taking in animal models include inducing malaise associated with the drug via LiCl or histamine treatment, or adding bitterants to the drug (primarily quinine added to alcohol) to cause an aversive taste response (Vanderschuren, Minnaard, Smeets, & Lesscher, 2017). Extended drug taking in rodents leads to the perseverance of drug-seeking behavior even when shocks are also delivered as a consequence for seeking and it is hypothesized that habits play a role in this perseverance (Belin, Mar, Dalley, Robbins, & Everitt, 2008; Chen et al., 2013; Pelloux, Everitt, & Dickinson, 2007; Vanderschuren & Everitt, 2004).

While generally studied in the context of drug abuse, compulsive responding is not limited to responding for drugs. Rodents will also tolerate electrical shocks to receive sucrose in some circumstances (Datta, Martini, Fan, & Sun, 2018; Nieh et al., 2015). As mentioned above, mice with OCD-linked gene mutations will continue to groom even when it causes pain and injury, an obvious negative consequence. Thus, it is important to understand how negative feedback plays a role in shaping the emergence of habitual behavior, and whether habit formation circuits drive punishment-resistant reward seeking both generally and in particular circumstances or disorders.

Many tests for compulsive responding ask rodents to learn a new action–outcome association between their previously learned action and a new aversive outcome such as a shock. The tests also potentially change the perceived cost of obtaining a rewarding outcome or indirectly reduce the value of the outcome since it is paired with aversion, depending on the timing of the aversive feedback. Thus, tests for "compulsivity" are similar to the probes designed to test for habit formation. Shock paradigms to test for compulsion vary in their methods. Some punish lever pressing with certainty, while others deliver shock probabilistically (e.g., Chen et al., 2013; Deroche-Gamonet, Belin, & Piazza, 2004). Some studies in monkeys delay the aversive outcome, although most rodent studies using shock deliver it immediately (Epstein & Kowalczyk, 2018; Vanderschuren et al., 2017; Woolverton, Freeman, Myerson, & Green, 2012). One important difference between shock delivery and aversive pairing (e.g., LiCl pairing) is that aversive pairing directly degrades the value of the reward, whereas shocks that occur immediately as a consequence for lever pressing punish the action but leave the reward value intact (for further commentary on this issue, see: Epstein & Kowalczyk, 2018; Vanderschuren et al., 2017). Differentiating between compulsive and habitual responding is thus a challenge, although one which may be surmountable with the addition of circuit-level investigations demonstrating whether similar or different neural mechanisms are involved in each.

Notably, habitual and compulsive responding do not always track together. When Singer and colleagues trained rats to solve a new operant "puzzle" each day to get cocaine, they found that rats still escalated their cocaine intake and continued to seek cocaine when a footshock consequence was imposed (Singer et al., 2018). However, the behavior in theory could not be fully automated, since the actions required to get the outcome were changing each day. Additionally, blocking dopamine signaling in the DLS did not interrupt cocaine-seeking behavior in this paradigm, in contrast to other studies (Giuliano, Belin, & Everitt, 2019; Murray et al., 2014; Vanderschuren, Ciano, & Everitt, 2005). These results demonstrate that compulsive drug seeking does not absolutely require dopamine signaling in the DLS. However, the data do not preclude the involvement of "habit" defined in the behavioral sense, for example, by outcome devaluation procedures.

Another example of the dissociation between compulsive and habitual responding comes from Willuhn and colleagues. They found that dopamine signaling in the DLS in response to cocaine self-administration develops over weeks of training (Willuhn, Burgeno, Everitt, & Phillips, 2012). DLS dopamine signaling in this case was required for the selection of drug-seeking actions; however, the behavioral paradigm used—a short access (1 hr/day) self-administration paradigm—is generally not sufficient to achieve compulsive (shock-resistant) drug seeking. Therefore, a model emerges in which habit-related brain systems may become engaged in behavior independently of compulsive responding. The involvement of DLS may precede the development of compulsive behavior, but, intriguingly, a transition to reliance on the DLS system predicts vulnerability to compulsivity (Giuliano et al., 2019). Still, based on these few studies it is difficult to fully assess the relationship between two closely related concepts. Additionally, it is not known whether cocaine hijacks habitual and compulsive neural mechanisms in a non-naturalistic way or whether responding for natural rewards such as sucrose would produce similar dissociations. One study found that the development of compulsive sucrose-seeking behavior in rats *could not* predict the development of compulsive cocaine-seeking behavior, suggesting that the neurobiological basis for the engagement of habit may differ under conditions of cocaine use (Datta et al., 2018).

2.5 | Avoidance learning

The vast majority of behavioral studies of habit have focused on paradigms in which animals receive valued rewards for their actions

(positive reinforcement). However, animals also learn from aversive outcomes (positive punishment), from the relief of aversive outcomes (negative reinforcement), and from the removal of rewarding outcomes (negative punishment). It unfortunately remains unclear whether and how striatal habit mechanisms are engaged by feedback mechanisms other than positive reinforcement. Studies of compulsivity help address the role of positive punishment, but what about the roles of negative reinforcement and punishment? It has been suggested that active avoidance learning, in which animals perform an action to prevent a shock from occurring, may invoke habit (LeDoux, Moscarello, Sears, & Campese, 2017). Habit is an appealing explanation for why animals continue to perform actions that prevent negative consequences, since as the animal correctly performs preventative actions they essentially begin to perform the actions in extinction (i.e., if the animal's actions prevent the negative consequence from occurring 100% of the time, then no obvious outcomes occur as the performance of the behavior continues).

The importance of understanding how habits contribute to avoidance is a key question in the study of anxiety disorders and OCD. Human OCD patients show stronger learning of avoidance habits than control subjects (Gillan et al., 2014). However, this enhanced habit formation is associated with an increase in activity in the caudate (analogous to rodent DMS), an area for goal-directed control, but not changes in the putamen (analogous to rodent DLS), suggesting that avoidance habits in OCD may be the result of impaired goal-directed systems rather than strengthened habit learning (Gillan et al., 2015). Human studies of avoidance habits also show that a history of early-life stress, which is associated with vulnerability to a number of psychiatric disorders, promotes the development of avoidance habits (Patterson, Craske, & Knowlton, 2019). Despite these interesting human findings, animal studies on avoidance learning have largely not considered habit.

If we can develop better ways to model avoidance habits in rodents, there are many interesting circuit-level hypotheses to explore, including the roles for dopamine and dorsal striatal circuits. Dopamine, which is thought to be important for habit formation when learning from positive reinforcement, is also likely important for learning from aversive outcomes. Subsets of dopamine neurons increase their activity for aversive outcomes and for cues predicting aversive outcomes (Lammel et al., 2012; Lerner et al., 2015; Matsumoto & Hikosaka, 2009; Menegas, Akiti, Amo, Uchida, & Watabe-Uchida, 2018), which could allow for the invigoration of actions by punishment. Additionally, some dopamine neurons projecting to the nucleus accumbens (NAc) respond to safety cues (indicating that shocks will not occur) and encode a "safety prediction error" signal (Stelly et al., 2019). Dopamine neurons projecting to the caudal tail of the striatum are also potentially interesting in the context of avoidance learning, as ablation of these neurons has been shown to reduce avoidance (Menegas et al., 2018). However, whether the activity of any of these dopamine neurons can control habit formation in the context of avoidance learning is not yet determined. Future studies more thoroughly examining the role of habit-related neural circuitry in learning from different types of reinforcement and punishment may help the field to clarify its definitions of habitual control over behavior.

3 | TOWARD A CIRCUIT MODEL FOR HABIT FORMATION

As behavioral work on habit and related tasks has proceeded as described above, so too has work to create a convincing circuit model for habit formation. Such a model is essential for progress in the field. Without a circuit model for habit formation, we cannot be sure if the various tasks being used to study habits and other potentially related behaviors (see Table 2) converge on similar circuits. Furthermore,

TABLE 2 Behavioral paradigms for habit formation and related behaviors

Behavioral paradigm	Key features
Random interval (RI) training	Positive reinforcement, uncertainty in timing leads to high levels of responding
Fixed ratio discrete trials (DT5) training	Positive reinforcement, cueing of discrete trials is a key to habit formation
T-maze	Positive reinforcement, sensory discrimination task leading to habit with extensive overtraining
Motor skill learning (e.g., accelerating rotarod, skilled reaching tasks, vocal learning)	Mixed reinforcement/punishments depending on the task, or can be performed without explicit external feedback. Precise motor timing requirements may engage habit mechanisms to ensure fluid action sequences
Grooming	Robust innate repetitive behavior. Self-injurious overgrooming may invoke positive punishment and model OCD symptoms such as excessive hand washing and trichotillomania
Compulsive drug or sucrose seeking	Positive Punishment for seeking drug or sucrose. Tests animals' sensitivity to the addition of an aversive outcome for seeking positive reinforcement
Avoidance learning	Negative reinforcement. Animals learn to act to avoid aversive outcomes. Important model for determining how habits contribute to avoidance, for example, in anxiety disorders
Two-step task	Many different types of reinforcement or punishment may be used. The two-step task allows one to assess the parallel contributions of "model-free" versus "model-based" behavior to performance

without a strong working model of normal habit formation, we are limited in our ability to systematically test whether habit circuits are altered in animal models of neuropsychiatric disease.

What is the current state of circuit models for habit formation? Extensive work has identified striatal learning systems in habit formation and this work provides us with a set of brain regions on which to focus. Specifically, the DLS is imperative for supporting habit formation and motor skill acquisition (Yin & Knowlton, 2006; Yin et al., 2004, 2009). Lesions to the DLS prevent habit formation (Yin et al., 2004), as do lesions of the dopaminergic inputs to the DLS from the substantia nigra pars compacta (SNc; Faure et al., 2005). Pharmacological blockade of dopaminergic signaling in the DLS also impairs motor skill acquisition on the accelerating rotarod (Yin et al., 2009) and habitual cocaine and heroin seeking (Belin & Everitt, 2008; Hodebourg et al., 2019; Willuhn et al., 2012).

This DLS learning system works in parallel with other striatal learning systems centered around the dorsomedial striatum (DMS) and ventral striatum (or nucleus accumbens, NAc) to regulate reward processing, incentive motivation, and action selection. Lesions to the DMS generally bias rodents away from goal-directed instrumental behavior and toward habit (Gremel & Costa, 2013; Yin, Knowlton, et al., 2005; Yin, Ostlund, et al., 2005), but the effects of DMS lesions are different if the anterior versus posterior DMS (pDMS) is targeted. Anterior DMS (aDMS) lesions do not have major effects on habit formation, as measured by outcome devaluation or by contingency degradation. pDMS lesions reduce instrumental performance and increase habit formation (Yin, Ostlund, et al., 2005). Lesions of the pDMS, but not aDMS, also bias rodents toward egocentric rather than allocentric navigation strategies in a T-maze task, a finding that is consistent with increased striatal-driven habit learning (Yin & Knowlton, 2004). Lesions to the NAc do not have major effects on measures of habit such as outcome devaluation and contingency degradation (Corbit, Muir, & Balleine, 2001; de Borchgrave, Rawlins, Dickinson, & Balleine, 2002). However, NAc core lesions impair instrumental performance and NAc medial shell lesions impair Pavlovian-instrumental transfer (Balleine & Killcross, 1994; Corbit et al., 2001).

Together, these lesion studies have crudely mapped the striatal subregions participating in different aspects of instrumental learning and habit formation, but a critical outstanding question in the field is to what degree these systems interact. Are the NAc, DMS, and DLS systems all engaged simultaneously in learning, and to what degree and at what level in the circuitry do they coordinate or compete to control behavioral output?

There is behavioral evidence for an interaction between striatal subregions in gating the transition to habit. Using NAc lesions paired with contralateral infusions of dopamine receptor antagonists in the DLS (to disconnect NAc activity from the control of DLS dopamine activity), Belin and Everitt demonstrated that cross talk between the NAc and DLS is important for habitual cocaine seeking (Belin & Everitt, 2008). However, this study, while foundational, did not provide circuit-level insight into the nature of the interaction taking place.

3.1 | The ascending spiral hypothesis

One prominent and influential hypothesis in the field regarding the interaction between striatal subsystems is the "ascending spiral" hypothesis (Haber, Fudge, & McFarland, 2000; Yin & Knowlton, 2006). The ascending spiral hypothesis posits that the NAc *disinhibits* DMS dopamine signaling, causing dopamine-dependent plasticity of corticostriatal connections in the DMS. In turn, DMS *disinhibits* dopamine signaling and dopamine-dependent corticostriatal plasticity in the DLS. The ascending spiral hypothesis originally arose from anatomical data collected in monkeys. Haber et al. (2000) used combinations of anterograde and retrograde tracers injected into the striatum to demonstrate a plausible route of indirect information flow from more ventromedial to more dorsolateral regions of the striatum through the dopaminergic midbrain. Axons originating from the ventral striatum overlapped with cell bodies of dopamine neurons projecting to the central striatum, and axons originating from the central striatum overlapped with cell bodies of dopamine neurons projecting to the DLS. While intriguing, a major limitation of this study is that the authors could not determine whether synapses were actually made between the labeled axons and cell bodies in their preparations; the argument was made based purely on proximity of the labels rather than functional measurements. In fact, notably, the ascending spiral hypothesis *does not* propose that direct connections are made between striatal axons and midbrain dopamine neurons. Since the striatum contains only GABAergic projection neurons, direct connections between the central striatum axons and DLS-projecting cell bodies, for example, would be inhibitory. Thus, it was proposed that there are disinhibitory connections, in which GABAergic striatal projection neurons would contact GABAergic cells in the nearby substantia nigra pars reticulata (SNr), which would then be the cells to contact the dopamine neurons projecting back to the DLS. Despite the appeal of this hypothesis for learning theories, the original data do not speak to the possibility of disynaptic disinhibition. In fact, the ascending spiral hypothesis is potentially in conflict with the observation that DMS lesions (at least of the pDMS) *accelerate* the emergence of habitual control over behavior, the opposite effect of what might be expected in this framework (Gremel & Costa, 2013; Yin, Knowlton, et al., 2005; Yin, Ostlund, et al., 2005). Thus, it is imperative to test the ascending spiral hypothesis more rigorously to determine its appropriate role in a circuit model for habit formation.

Disinhibitory inputs, which are central to the ascending spiral hypothesis, are posited on the basis of separate knowledge of striatal inputs to midbrain SNr GABA neurons, and SNr GABA inputs to dopamine neurons. The direct pathway of the striatum sends GABAergic projections to SNr cells (Albin, Young, & Penney, 1989; DeLong, 1990), although the SNr is not uniformly inhibited by direct pathway stimulation *in vivo* (Freeze, Kravitz, Hammack, Berke, & Kreitzer, 2013). In turn, dopamine neurons in the SNc receive strong GABAergic inputs from the SNr (Tepper & Lee, 2007; Tepper, Martin, & Anderson, 1995). SNr GABA neurons have tonic,

linear current–frequency relationships (Richards, Shiroyama, & Kitai, 1997), meaning that a disinhibition circuit through these neurons would likely lead to corresponding graded changes in SNc dopamine neuron tonic firing rather than inducing bursts. Dopamine burst firing relevant to habit formation could be induced by concurrent excitatory inputs, whose efficacy might be strengthened by decreased inhibition from the SNr, but such a circuit then needs to be included explicitly in the ascending spiral hypothesis model.

A careful study of the morphology of SNr GABA neurons showed that SNr axons extend into the SNc in a longitudinal band across the ventral tier. This band encompasses the location of SNc dopamine neurons projecting to the striatal subregion from which the traced SNr cell would receive inputs, but may also extend beyond those boundaries (Mailly, Charpier, Menetrey, & Deniau, 2003). However, since this study was morphological and did not measure functional synaptic strengths, at present we still do not know if there is a specific connectivity from DMS to SNr GABA neurons that project to DLS-projecting dopamine neurons, or the strength of that connectivity if it does exist. Additionally, it does not appear that all SNr GABA neurons make synapses onto dopamine neurons (although further work is needed to characterize different streams of SNr output; Rizzi & Tan, 2019). Disinhibition could potentially work only in closed reciprocal loops (e.g., DMS disinhibiting DMS-projecting dopamine neurons) or in a "descending spiral" (e.g., DLS disinhibiting DMS-projecting dopamine neurons) as well as in an ascending spiral.

A second oft-cited reference related to the ascending spiral hypothesis is Ikemoto et al. (2007), which was conducted in rats. In this study, the retrograde tracer Fluoro-Gold was injected into various striatal sites and the locations of labeled dopaminergic cell bodies in the midbrain were reported. Indeed, there was a clear organization of dopamine cell bodies found, and dopaminergic projections to the dorsal striatum were found to arise primarily from the SNc. However, no distinction was made in this study between DMS and DLS within the dorsal striatum, and no anterograde tracing (parallel to what Haber et al. (2000) completed in monkeys) was done to examine the overlap of output-defined dopamine neuron cell bodies with inputs from distinct striatal subregions. Additionally, as was true in the Haber et al. (2000) study, no experiments (e.g., electrophysiological measurements) were carried out to verify functional synaptic connections within a striatonigrostriatal spiral, meaning there is still no direct evidence that such a circuit could mediate disinhibition during habit formation.

More direct evidence of disinhibitory control over dopamine neurons exists in the ventral tegmental area (VTA), which contains dopamine neurons projecting to the NAc. However, the patterns of disinhibitory control do not clearly follow predictions of the ascending spiral hypothesis. NAc neurons projecting to the VTA preferentially target VTA GABA neurons, leading to a disinhibition of VTA dopamine signaling following optogenetic stimulation of NAc terminals in the VTA (Bocklisch et al., 2013). This disinhibition appears to operate in a closed reciprocal loop, rather than an open ascending spiral. Supporting this finding, in a study looking at specific striatal subregions it was found that NAc lateral shell neurons disinhibit dopamine neurons that project back to the lateral shell in a reciprocal loop through VTA GABA neurons (Yang et al., 2018). Reciprocal loop dopamine disinhibition may also be important for songbird vocal learning (Gale & Perkel, 2010).

It has been suggested that the ventral pallidum (VP) is well suited to mediate a disinhibitory ascending spiral connecting the NAc and dorsal striatum (Root, Melendez, Zaborszky, & Napier, 2015). The VP is a source of inhibitory afferent control for SNc dopamine neurons that receive inputs from the NAc, making this suggestion plausible. Indeed, the VP is required for Pavlovian-instrumental transfer (Leung & Balleine, 2013) as would be predicted for such a circuit. However, the role of the VP in mediating transitions from goal-directed instrumental behavior to habit is not established. This transition may require a different mechanism, hypothesized by the ascending spiral hypothesis to be a disinhibitory connection between the DMS and DLS.

Disynaptic disinhibition of dopamine neurons is not the only route by which striatal activity might influence dopamine release. It is also important to consider the role of direct striatal inputs to dopamine neurons, which constitute a major source of their afferent control. Monosynaptic rabies-tracing experiments have provided a useful overview of the brain-wide inputs to midbrain dopamine neurons (Beier et al., 2015; Lerner et al., 2015; Menegas et al., 2015; Watabe-Uchida, Zhu, Ogawa, Vamanrao, & Uchida, 2012). These experiments confirmed that dopamine neurons receive direct inputs from striatum and demonstrated the relative numbers of inputs received in comparison with other brain areas. Notably, SNc dopamine neurons receive ~50% of their inputs from the dorsal striatum and an additional significant portion from the NAc (Lerner et al., 2015; Watabe-Uchida et al., 2012). Dopamine neurons in the VTA also receive inputs from both the dorsal and ventral striatum (Beier et al., 2015, 2019). Therefore, in both the SNc and VTA, there is potential for direct monosynaptic inhibition in addition to disynaptic disinhibition of dopamine neurons.

Direct inhibitory inputs are not a part of the ascending spiral hypothesis as it is currently set forth, and in fact these inputs appear to follow an opposite pattern: DLS inputs to DMS-projecting dopamine neurons are common and strong, as measured by both rabies tracing and electrophysiology (Lerner et al., 2015). Similarly, rabies-tracing experiments showed that the dorsal striatum sends large numbers of inputs to dopamine neurons that project to the NAc lateral shell (although this NAc lateral shell-projector population also sent inputs to DMS and DLS, complicating the interpretation; Beier et al., 2015).

Since both inhibition and disinhibition circuits may connect striatal activity to dopamine neuron activity, it is reasonable to ask which type of modulation dominates at behaviorally relevant time points. It is not clear if the inhibition and disinhibition circuits operate together (i.e., are active at the same times during behavior), especially as these circuits may arise from different striatal neuron populations. Striosomes (also known as patches) within the striatum project directly to dopamine neurons, whereas the matrix compartment of the striatum contains direct pathway projections to SNr GABA

neurons. Striosome and matrix neurons receive different cortical inputs, which may drive their engagement in behavior separately (Friedman et al., 2015; Smith et al., 2016). In general, striosomes receive input from more "limbic" areas, as opposed to the associational and sensorimotor cortical inputs to DMS and DLS matrix neurons, respectively. *In vivo* imaging from striosome neurons shows some differences in activity patterns between compartments, with striosome neurons responding more strongly to reward-predicting cues than matrix neurons (Bloem, Huda, Sur, & Graybiel, 2017). Thus, one can hypothesize that striatal inhibition of dopamine neurons dominates during cue presentation, especially after extensive training.

Striosomes are likely an important part of the habit formation circuit. Partial ablation of striosome neurons with a selective toxin called dermorphin–saporin causes deficits in learning on the rotarod (Lawhorn, Smith, & Brown, 2009) as well as in habit formation in a more traditional operant test (Jenrette, Logue, & Horner, 2019). One possible mechanism for these effects on learning could be a resulting imbalance in the regulation of striatal dopamine release (Shumilov, Real, Valderrama-Carvajal, & Rivera, 2018). Imbalances between activity in the striosome and matrix compartments have been proposed to contribute to the development of neurological and psychiatric disorders including Huntington's disease, L-DOPA-induced dyskinesias, dystonia, and drug addiction (Crittenden & Graybiel, 2011). Understanding these imbalances and the circuit mechanisms by which they might contribute to symptomatology will be a key to generate new clinical interventions.

In conclusion, while the ascending spiral hypothesis has been influential in the habit formation field, convincing circuit- and synaptic-level evidence of disinhibition has not been demonstrated, leaving the door open to other possibilities. Although striatonigrostriatal loops might mediate NAc to DMS to DLS information transfer, we should not focus on them to the exclusion of other possibilities. Other possible circuits that could promote communication between the DMS and DLS include corticostriatothalamic loops (e.g. Aoki et al., 2019), lateral connections made between striatal subregions (including through interneurons), and basal ganglia loops downstream of the striatum (e.g., through the globus pallidus externa, which sends projections back to the striatum). An ascending spiral circuit might also work in parallel with circuits that dampen rather than promote habit. Silencing of the DLS, particularly direct pathway striatal neurons in the DLS, promotes early goal-directed instrumental learning and PFC-DMS circuit engagement (Bergstrom et al., 2018). Thus, inputs from the DLS onto midbrain DMS dopamine circuits may serve to slow the acquisition of habits through a "descending spiral."

3.2 | Shifting patterns of DMS and DLS involvement in behavior with habit formation

In vivo electrophysiological recordings show that patterns of activity in the DMS and DLS change with habit formation and motor skill acquisition, but different tasks can produce different results, calling into question whether the same circuits and plasticity mechanisms are engaged by each (Gremel & Costa, 2013; Thorn et al., 2010; Vandaele et al., 2019; Yin et al., 2009). Thorn et al. (2010) used the T-maze task (described above) paired with tetrode recordings in the DMS and DLS. Similar percentages of task-responsive neurons were found in each striatal subregion, however, the patterns of activity differed across training. Responses in the DLS tended to occur at action boundaries of the task (locomotion onset, turn, goal). Goal responses in particular seemed to emerge and strengthen with overtraining, after rats reached a performance criterion, perhaps reflecting an emerging reward responsiveness. In contrast, DMS neurons responded most strongly in the middle of the task as the rats progressed down the long arm of the T-maze track. Strong responses to the cue onset (the signal telling rats which direction to turn) occurred mid-training, but faded with overtraining. These results seem in line with the idea that DMS is most actively engaged in action–outcome learning during an earlier phase of task experience, whereas DLS becomes engaged in creating habits later on.

Gremel and Costa (2013) recorded DMS and DLS neurons using a more traditional operant task. They trained mice to pursue rewards on an RI schedule (promoting habitual responding) in one context and an RR schedule (promoting goal-directed responding) in another context. This clever study design allowed them to assess within-subject differences depending on the training context. Similar to Thorn et al. (2010), this group found roughly equal percentages of task-responsive neurons in DMS and DLS. While some neurons responded specifically in one context, many were modulated in both the RI and RR contexts. The observation of task-responsive neurons in both DMS and DLS in both contexts questions the notion of a hard distinction between the two systems as habitual control emerges. When looking at the magnitude of the changes observed in DMS and DLS, however, some differences were observed in this study. After training, DMS neurons had a larger increase in their lever press–associated firing in the RR context when the reward had been devalued. DLS neurons had a smaller increase their lever press–associated firing in RR context when the reward was valued. In contrast to Thorn et al. (2010), Gremel and Costa did not find disengagement of DMS task-responsive neurons over training in the habitual context, nor did they find any changes in DLS task responsiveness with habit (RI context).

Another study by Vandaele et al. (2019) used the DT5 task (Table 2, described above). Like Gremel and Costa (2013), this group observed that both the DMS and DLS remained substantially task-responsive late into training, in this case many weeks after habits (as assessed by satiety-specific devaluation) had formed. The continued engagement of DMS in habitual behavior questions the notion that behavioral control completely shifts to DLS circuits with overtraining. DLS may still be required for the initial transition to habit, but the consolidation of habit memory may take place elsewhere. Indeed, in this study, pharmacological inactivation of DLS late in training had modest effects on behavior, slightly decreasing lever press rates, but overall did not prevent performance of the task.

Finally, Yin et al. (2009) used the accelerating rotarod for their study examining the participation of DMS and DLS neurons in motor skill acquisition. They found a pattern of early DMS engagement in the task and later DLS engagement as the task was mastered and performance plateaued. In this case, the findings appear more similar to Thorn et al. (2010), with DMS disengaging later.

These four studies clearly drive home the point that DMS and DLS engagement in behavior may be highly task dependent. They leave future researchers with the difficult job of parsing which responses are truly required for habit formation in general, and which are task specific. The reasons that certain tasks maintain DMS engagement while others cause it to diminish will be a particularly interesting avenue for future work. Such investigations will be important for clarifying whether there are multiple neural circuit implementations of habit available to an animal. Better connecting the emergence of habitual behavior with each of these recordings will also be key. Since behavioral probes for habit (Table 1) can only be done at discrete time points, it can be difficult to assess exactly when an individual animal is transitioning to habitual control, limiting the power of analyses. In the Vandaele et al. (2019) study, habit occurred early in training (after 10 sessions). Thus, habit per se could not be correlated with the late changes observed in DMS after many weeks. In contrast, the T-maze task using by Thorn et al. (2010) is more complicated to train. Training to criteria and then further overtraining until rats are insensitive to outcome devaluation generally takes much longer than when using the DT5 task (Smith et al., 2012). Whether these differences in training time or other aspects of the tasks are important for determining how DMS and DLS are engaged remain to be seen.

Notably, all of these studies which compared the *in vivo* activities of DMS and DLS neurons used relatively anterior recording coordinates. It remains unclear how the activities of posterior striatal regions are correlated with the emergence of habitual behavior, and this is a potentially important question. Lesions and inactivations of aDMS and pDMS differ in their effects, with pDMS lesions being more effective at promoting the early emergence of habitual control (Yin, Ostlund, et al., 2005). However, since most recordings are done in the aDMS it is difficult to know how to align the two literatures. Additionally, the posterior DLS (pDLS), including the far caudal tail of the striatum, is an understudied area, rich in cells projecting directly to substantia nigra dopamine neurons (Lerner et al., 2015; Menegas et al., 2015). Dopamine cells projecting to the caudal tail of the striatum also have unique input connectivity patterns (Menegas et al., 2015). Thus, it will be illuminating for future studies to determine the functions of circuits involving the pDLS and caudal tail of the striatum in the emergence of habitual behavior and to incorporate these striatal subregions into a refined circuit model of habit formation.

3.3 | Plasticity of habit circuits with learning

To assess the validity of any circuit model of habit formation that is developed, we must determine what types of plasticity take place during training to mechanistically cause the observed *in vivo* shifts in striatal function over time. A growing body of evidence points to the involvement of corticostriatal plasticity mechanisms in habit formation. Long-term synaptic plasticity at cortical inputs onto DMS and DLS neurons depends critically on dopamine, and blocking dopamine signaling during learning impairs habit or motor skill acquisition (Faure et al., 2005; Yin et al., 2009). Additionally, inhibition of adenosine A2A receptors or endocannabinoid CB1 receptors, both of which are known to be important actors in corticostriatal plasticity pathways (Lerner, Horne, Stella, & Kreitzer, 2010; Lerner & Kreitzer, 2011, 2012; Shen, Flajolet, Greengard, & Surmeier, 2008; Surmeier, Plotkin, & Shen, 2009), interferes with habit formation (Gremel et al., 2016; Hilário, Clouse, Yin, & Costa, 2007; Li et al., 2016; Yu et al., 2009).

Using an accelerating rotarod task, Yin et al. (2009) showed that AMPA:NMDA ratios at excitatory inputs onto DLS neurons are decreased specifically in the later stages of learning, after performance has plateaued. Additionally, LTD in the DLS was more readily observed in slices made from mice trained to the late stages of learning, suggesting that LTP had occurred *in vivo*. Another study using the rotarod task found that the engagement of cortical inputs to the DMS and DLS changes dynamically during learning (Kupferschmidt, Juczewski, Cui, Johnson, & Lovinger, 2017). PFC inputs to the DMS peak in activity early in learning and disengage later, while M1 motor cortical inputs to the DLS remain strong. However, a limitation in both of these studies is that differences in changes onto direct versus indirect pathway striatal neurons were not examined.

Corticostriatal plasticity may act disparately on the direct and indirect pathways within the striatum over the course of goal-directed and habitual learning. In the pDMS, AMPA:NMDA ratios increase onto direct pathway neurons but decrease onto indirect pathway neurons after training to a goal-directed stage of behavior (Shan, Ge, Christie, & Balleine, 2014). No changes in AMPA:NMDA ratios were observed in the DLS at this early stage in learning. After longer training on a RI60 schedule to induce habitual control, indirect pathway neurons in the DLS showed a reduced amplitude of spontaneous EPSCs (sEPSCs), suggesting that LTD onto these neurons had occurred (Shan, Christie, & Balleine, 2015). The average amplitude of recorded sEPSCs from each mouse negatively correlated with its press rate in the last RI60 training session, suggesting that this reduction in sEPSC amplitude is specifically involved in the escalation of responding behavior associated with habit.

In addition to plasticity in the strengths of corticostriatal coupling with the direct and indirect pathways, shifts in timing may play a role in habit formation. In a study using acute brain slices containing the DLS, O'Hare et al. (2016) found that changes in the relative timing of direct versus indirect pathway activity in response to cortical stimulation correlated with habitual behavior: direct pathway striatal neurons fired before indirect pathway striatal neurons in habitual mice, whereas the inverse was true in goal-directed mice.

Corticostriatal plasticity is further implicated in habit-related behaviors because mice prone to developing OCD-like repetitive

overgrooming behaviors all have corticostriatal synaptic deficits in common (Peça et al., 2011; Shmelkov et al., 2010; Welch et al., 2007), and overgrooming can be induced by repetitive corticostriatal stimulation (Ahmari et al., 2013) perhaps by engaging endocannabinoid-dependent long-term depression mechanisms important for the development of habitual responding (Gremel et al., 2016).

The fact that corticostriatal plasticity is gated not only by dopamine but also by a host of other neuromodulators suggests that dopaminergic circuits like those invoked by the ascending spiral hypothesis may not be the only mechanism by which transitions to habit are influenced. Alterations to circuits that gate the release of neuromodulators like adenosine, acetylcholine, and endogenous opioids in the striatum could also contribute to habit formation under different circumstances and in different disorders.

Sites of plasticity other than corticostriatal synapses may additionally play a role in shaping the function of the striatal circuitry regulating habit. Plasticity in cortical circuits upstream of the corticostriatal projections is one example. As another example, if dopamine inputs to the DMS and DLS are regulated by an ascending spiral from other striatal regions as proposed, then plasticity of inputs onto SNr GABA and/or SNc dopamine neurons might regulate habit formation through the spiral. Some inputs to SNc dopamine neurons are altered by exposure to drugs of abuse such as cocaine (Beaudoin et al., 2018), which could provide a basis for understanding how these drugs engage habit circuits. Overall, there are likely many distinct sites of plasticity occurring during habit formation. Plasticity events at these different sites might act together and be interdependent on one another. Understanding which synaptic changes occur at which points in training could help answer the question of why certain reinforcement schedules lead to the emergence of habitual control on different timescales.

4 | CONCLUSION

As a field, we have developed an array of behavioral tasks to study habit. What is now required is to better formalize our definitions of habit, thinking broadly across behavioral fields to integrate studies of instrumental responding, motor skill learning, repetitive behaviors, compulsive behaviors, and avoidance learning. Furthermore, a circuit model for habit—encompassing specific descriptions of circuits and synaptic changes that mediate changes in network activity occurring with habit formation—will provide a foundation to compare mechanisms across tasks. This review has focused on striatal mechanisms, but in fact many additional brain circuits may play a role as well and should be incorporated into our theories. Ultimately, a convincing circuit model for habit is indispensable for understanding the complex relationships between habit and habit-related behavioral tasks, and is required to make substantive progress on addressing the question of whether dysfunctions in habit circuits indeed contribute to the symptoms observed in various neuropsychiatric disorders such as OCD, autism, and addiction.

ACKNOWLEDGMENTS

This work was funded by the NIH (R00MH109569 and DP2MH122401) and by a NARSAD Young Investigator Award from the Brain & Behavior Research Foundation.

CONFLICT OF INTEREST

The author has no conflicts of interest to declare.

AUTHOR CONTRIBUTION

Conceptualization, T.L.; *Writing – Original Draft*, T.L.; *Writing – Review & Editing*, T.L.

ORCID

Talia N. Lerner https://orcid.org/0000-0002-3875-0654

REFERENCES

Ahmari, S. E., Spellman, T., Douglass, N. L., Kheirbek, M. A., Simpson, H. B., Deisseroth, K., … Hen, R. (2013). Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science*, *340*, 1234–1239. https://doi.org/10.1126/science.1234733

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences*, *12*, 366–375. https://doi.org/10.1016/0166-2236(89)90074-X

Alvares, G. A., Balleine, B. W., Whittle, L., & Guastella, A. J. (2016). Reduced goal-directed action control in autism spectrum disorder. *Autism Research*, *9*, 1285–1293. https://doi.org/10.1002/aur.1613

Aoki, S., Smith, J. B., Li, H., Yan, X., Igarashi, M., Coulon, P., … Jin, X. (2019). An open cortico-basal ganglia loop allows limbic control over motor output via the nigrothalamic pathway. *Elife*, *8*, e49995. https://doi.org/10.7554/eLife. 49995

Balleine, B., & Killcross, S. (1994). Effects of ibotenic acid lesions of the nucleus accumbens on instrumental action. *Behavioral Brain Research*, *65*, 181–193. https://doi.org/10.1016/0166-4328(94)90104-X

Bannard, C., Leriche, M., Bandmann, O., Brown, C. H., Ferracane, E., Sánchez-Ferro, Á., … Stafford, T. (2019). Reduced habit-driven errors in Parkinson's disease. *Scientific Reports*, *9*, 3423. https://doi.org/10.1038/s41598-019-39294-z

Beaudoin, G. M. J., Gomez, J. A., Perkins, J., Bland, J. L., Petko, A. K., & Paladini, C. A. (2018). Cocaine selectively reorganizes excitatory inputs to Substantia Nigra pars compacta dopamine neurons. *Journal of Neuroscience*, *38*, 1151–1159. https://doi.org/10.1523/JNEUROSCI.1975-17.2017

Beier, K. T., Gao, X. J., Xie, S., DeLoach, K. E., Malenka, R. C., & Luo, L. (2019). Topological organization of ventral tegmental area connectivity revealed by viral-genetic dissection of input-output relations. *Cell Reports*, *26*, 159–167.e6. https://doi.org/10.1016/j.celrep.2018.12.040

Beier, K. T., Steinberg, E. E., DeLoach, K. E., Xie, S., Miyamichi, K., Schwarz, L., … Luo, L. (2015). Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. *Cell*, *162*, 622–634. https://doi.org/10.1016/j.cell.2015.07.015

Belin, D., & Everitt, B. J. (2008). Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron*, *57*, 432–441. https://doi.org/10.1016/j.neuron.2007.12.019

Belin, D., Mar, A. C., Dalley, J. W., Robbins, T. W., & Everitt, B. J. (2008). High impulsivity predicts the switch to compulsive cocaine-taking. *Science*, *320*, 1352–1355. https://doi.org/10.1126/science.1158136

Bergstrom, H. C., Lipkin, A. M., Lieberman, A. G., Pinard, C. R., Gunduz-Cinar, O., Brockway, E. T., … Holmes, A. (2018). Dorsolateral striatum

engagement interferes with early discrimination learning. *Cell Reports*, *23*, 2264–2272. https://doi.org/10.1016/j.celrep.2018.04.081

Bloem, B., Huda, R., Sur, M., & Graybiel, A. M. (2017). Two-photon imaging in mice shows striosomes and matrix have overlapping but differential reinforcement-related responses. *ELife*, *6*, e32353.

Bocklisch, C., Pascoli, V., Wong, J. C. Y., House, D. R. C., Yvon, C., de Roo, M., … Lüscher, C. (2013). Cocaine disinhibits dopamine neurons by potentiation of GABA transmission in the ventral tegmental area. *Science*, *341*, 1521–1525. https://doi.org/10.1126/science.1237059

Brainard, M. S., & Doupe, A. J. (2000). Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature*, *404*, 762–766. https://doi.org/10.1038/35008083

Burguière, E., Monteiro, P., Feng, G., & Graybiel, A. M. (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science*, *340*, 1243–1246. https://doi.org/10.1126/science.1232380

Charlesworth, J. D., Tumer, E. C., Warren, T. L., & Brainard, M. S. (2011). Learning the microstructure of successful behavior. *Nature Neuroscience*, *14*, 373–380. https://doi.org/10.1038/nn.2748

Chen, B. T., Yau, H.-J., Hatch, C., Kusumoto-Yoshida, I., Cho, S. L., Hopf, F. W., & Bonci, A. (2013). Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature*, *496*, 359–362. https://doi.org/10.1038/nature12024

Corbit, L. H., Muir, J. L., & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *Journal of Neuroscience*, *21*, 3251–3260. https://doi.org/10.1523/JNEUROSCI.21-09-03251.2001

Crittenden, J. R., & Graybiel, A. M. (2011). Basal ganglia disorders associated with imbalances in the striatal striosome and matrix compartments. *Frontiers in Neuroanatomy*, *5*, 59. https://doi.org/10.3389/fnana.2011.00059

da Silva, J. A., Tecuapetla, F., Paixão, V., & Costa, R. M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature*, *554*, 244–248. https://doi.org/10.1038/nature25457

Datta, U., Martini, M., Fan, M., & Sun, W. (2018). Compulsive sucrose- and cocaine-seeking behaviors in male and female Wistar rats. *Psychopharmacology*, *235*, 2395–2405. https://doi.org/10.1007/s00213-018-4937-1

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*, 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*, 1704–1711. https://doi.org/10.1038/nn1560

de Borchgrave, R., Rawlins, J. N. P., Dickinson, A., & Balleine, B. W. (2002). Effects of cytotoxic nucleus accumbens lesions on instrumental conditioning in rats. *Experimental Brain Research*, *144*, 50–68. https://doi.org/10.1007/s00221-002-1031-y

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neurosciences*, *13*, 281–285. https://doi.org/10.1016/0166-2236(90)90110-V

Deroche-Gamonet, V., Belin, D., & Piazza, P. V. (2004). Evidence for addiction-like behavior in the rat. *Science*, *305*, 1014–1017. https://doi.org/10.1126/science.1099020

DeRusso, A., Fan, D., Gupta, J., Shelest, O., Costa, R. M., & Yin, H. H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Frontiers in Integrative Neuroscience*, *4*, 17.

Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Computational Biology*, *9*, e1003364. https://doi.org/10.1371/journal.pcbi.1003364

Dhawale, A. K., Smith, M. A., & Ölveczky, B. P. (2017). The role of variability in motor learning. *Annual Review of Neuroscience*, *40*, 479–498. https://doi.org/10.1146/annurev-neuro-072116-031548

Dhawale, A. K., Wolff, S. B. E., Ko, R., & Ölveczky, B. P. (2019). The basal ganglia can control learned motor sequences independently of motor cortex. *BioRxiv*, 827261. https://doi.org/10.1101/827261

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *308*, 67–78. https://doi.org/10.1098/rstb.1985.0010

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, *22*, 1–18. https://doi.org/10.3758/BF03199951

Dickinson, A., & Charnock, D. J. (1985). Contingency effects with maintained instrumental reinforcement. *The Quarterly Journal of Experimental Psychology Section B*, *37*, 397–416. https://doi.org/10.1080/14640748508401177

Dickinson, A., Nicholas, D. J., & Adams, C. D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *35*, 35–51. https://doi.org/10.1080/14640748308400912

Doupe, A. J., Perkel, D. J., Reiner, A., & Stern, E. A. (2005). Birdbrains could teach basal ganglia research a new song. *Trends in Neurosciences*, *28*, 353–363. https://doi.org/10.1016/j.tins.2005.05.005

Epstein, D. H., & Kowalczyk, W. J. (2018). Compulsive seekers: Our take. Two clinicians' perspective on a new animal model of addiction. *Neuropsychopharmacology*, *43*, 677–679. https://doi.org/10.1038/npp.2017.132

Everitt, B. J., & Robbins, T. W. (2016). Drug addiction: Updating actions to habits to compulsions ten years on. *Annual Review of Psychology*, *67*, 23–50. https://doi.org/10.1146/annurev-psych-122414-033457

Faure, A., Haberland, U., Condé, F., & El Massioui, N. (2005). Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *Journal of Neuroscience*, *25*, 2771–2780. https://doi.org/10.1523/JNEUROSCI.3894-04.2005

Freeze, B. S., Kravitz, A. V., Hammack, N., Berke, J. D., & Kreitzer, A. C. (2013). Control of basal ganglia output by direct and indirect pathway projection neurons. *Journal of Neuroscience*, *33*, 18531–18539. https://doi.org/10.1523/JNEUROSCI.1278-13.2013

Friedman, A., Homma, D., Gibb, L. G., Amemori, K.-I., Rubin, S. J., Hood, A. S., … Graybiel, A. M. (2015). A corticostriatal path targeting striosomes controls decision-making under conflict. *Cell*, *161*, 1320–1333. https://doi.org/10.1016/j.cell.2015.04.049

Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., & Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science*, *354*, 1278–1282. https://doi.org/10.1126/science.aah6837

Gale, S. D., & Perkel, D. J. (2010). A basal ganglia pathway drives selective auditory responses in songbird dopaminergic neurons via disinhibition. *Journal of Neuroscience*, *30*, 1027–1037. https://doi.org/10.1523/JNEUROSCI.3585-09.2010

George, O., & Koob, G. F. (2017). Individual differences in the neuropsychopathology of addiction. *Dialogues in Clinical Neuroscience*, *19*, 217–229.

Gillan, C. M., Apergis-Schoute, A. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Fineberg, N. A., … Robbins, T. W. (2015). Functional neuroimaging of avoidance habits in obsessive-compulsive disorder. *American Journal of Psychiatry*, *172*, 284–293. https://doi.org/10.1176/appi.ajp.2014.14040525

Gillan, C. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Voon, V., Apergis-Schoute, A. M., … Robbins, T. W. (2014). Enhanced avoidance habits in obsessive-compulsive disorder. *Biological Psychiatry*, *75*, 631–638. https://doi.org/10.1016/j.biopsych.2013.02.002

Giuliano, C., Belin, D., & Everitt, B. J. (2019). Compulsive alcohol seeking results from a failure to disengage dorsolateral striatal control over behavior. *Journal of Neuroscience*, *39*, 1744–1754. https://doi.org/10.1523/JNEUROSCI.2615-18.2018

Gourley, S. L., Olevska, A., Gordon, J., & Taylor, J. R. (2013). Cytoskeletal determinants of stimulus-response habits. *Journal of Neuroscience*, *33*, 11811–11816. https://doi.org/10.1523/JNEUROSCI.1034-13.2013

Gremel, C. M., Chancey, J. H., Atwood, B. K., Luo, G., Neve, R., Ramakrishnan, C., … Costa, R. M. (2016). Endocannabinoid modulation of orbitostriatal circuits gates habit formation. *Neuron*, *90*, 1312–1324. https://doi.org/10.1016/j.neuron.2016.04.043

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, *4*, 2264. https://doi.org/10.1038/ncomms3264

Haber, S. N., Fudge, J. L., & McFarland, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *Journal of Neuroscience*, *20*, 2369–2382. https://doi.org/10.1523/JNEUROSCI.20-06-02369.2000

Hilário, M. R. F., Clouse, E., Yin, H. H., & Costa, R. M. (2007). Endocannabinoid signaling is critical for habit formation. *Frontiers in Integrative Neuroscience*, *1*, 6. https://doi.org/10.3389/neuro.07.006.2007

Hodebourg, R., Murray, J. E., Fouyssac, M., Puaud, M., Everitt, B. J., & Belin, D. (2019). Heroin seeking becomes dependent on dorsal striatal dopaminergic mechanisms and can be decreased by N-acetylcysteine. *European Journal of Neuroscience*, *50*, 2036–2044. https://doi.org/10.1111/ejn.13894

Howe, M. W., & Dombeck, D. A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*, *535*, 505–510. https://doi.org/10.1038/nature18942

Ikemoto, S. (2007). Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Research Reviews*, *56*(1), 27–78.

Jenrette, T. A., Logue, J. B., & Horner, K. A. (2019). Lesions of the patch compartment of dorsolateral striatum disrupt stimulus-response learning. *Neuroscience*, *415*, 161–172. https://doi.org/10.1016/j.neuroscience.2019.07.033

Jin, X., & Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, *466*, 457–462. https://doi.org/10.1038/nature09263

Jin, X., Tecuapetla, F., & Costa, R. M. (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nature Neuroscience*, *17*, 423–430. https://doi.org/10.1038/nn.3632

Kawai, R., Markman, T., Poddar, R., Ko, R., Fantana, A. L., Dhawale, A. K., … Ölveczky, B. P. (2015). Motor cortex is required for learning but not for executing a motor skill. *Neuron*, *86*, 800–812. https://doi.org/10.1016/j.neuron.2015.03.024

Kawashima, S., Ueki, Y., Kato, T., Ito, K., & Matsukawa, N. (2018). Reduced striatal dopamine release during motor skill acquisition in Parkinson's disease. *PLoS ONE*, *13*, e0196661. https://doi.org/10.1371/journal.pone.0196661

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*, 1399–1402. https://doi.org/10.1126/science.273.5280.1399

Kubota, Y., Liu, J., Hu, D., DeCoteau, W. E., Eden, U. T., Smith, A. C., & Graybiel, A. M. (2009). Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *Journal of Neurophysiology*, *102*, 2142–2160. https://doi.org/10.1152/jn.00522.2009

Kupferschmidt, D. A., Juczewski, K., Cui, G., Johnson, K. A., & Lovinger, D. M. (2017). Parallel, but dissociable, processing in discrete corticostriatal inputs encodes skill learning. *Neuron*, *96*, 476–489.e5. https://doi.org/10.1016/j.neuron.2017.09.040

Lammel, S., Lim, B. K., Ran, C., Huang, K. W., Betley, M. J., Tye, K. M., … Malenka, R. C. (2012). Input-specific control of reward and aversion in the ventral tegmental area. *Nature*, *491*, 212–217. https://doi.org/10.1038/nature11527

Lawhorn, C., Smith, D. M., & Brown, L. L. (2009). Partial ablation of mu-opioid receptor rich striosomes produces deficits on a motor-skill learning task. *Neuroscience*, *163*, 109–119. https://doi.org/10.1016/j.neuroscience.2009.05.021

LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2017). The birth, death and resurrection of avoidance: A reconceptualization of a troubled paradigm. *Molecular Psychiatry*, *22*, 24–36. https://doi.org/10.1038/mp.2016.166

Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, *81*, 687–699. https://doi.org/10.1016/j.neuron.2013.11.028

Lerner, T. N., Horne, E. A., Stella, N., & Kreitzer, A. C. (2010). Endocannabinoid signaling mediates psychomotor activation by adenosine A2A antagonists. *Journal of Neuroscience*, *30*, 2160–2164. https://doi.org/10.1523/JNEUROSCI.5844-09.2010

Lerner, T. N., & Kreitzer, A. C. (2011). Neuromodulatory control of striatal plasticity and behavior. *Current Opinion in Neurobiology*, *21*, 322–327. https://doi.org/10.1016/j.conb.2011.01.005

Lerner, T. N., & Kreitzer, A. C. (2012). RGS4 is required for dopaminergic control of striatal LTD and susceptibility to parkinsonian motor deficits. *Neuron*, *73*, 347–359. https://doi.org/10.1016/j.neuron.2011.11.015

Lerner, T. N., Shilyansky, C., Davidson, T. J., Evans, K. E., Beier, K. T., Zalocusky, K. A., … Deisseroth, K. (2015). Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell*, *162*, 635–647. https://doi.org/10.1016/j.cell.2015.07.014

Leung, B. K., & Balleine, B. W. (2013). The ventral striato-pallidal pathway mediates the effect of predictive learning on choice between goal-directed actions. *Journal of Neuroscience*, *33*, 13848–13860. https://doi.org/10.1523/JNEUROSCI.1697-13.2013

Li, Y., He, Y., Chen, M., Pu, Z., Chen, L. I., Li, P., … Chen, J.-F. (2016). Optogenetic activation of adenosine A2A receptor signaling in the dorsomedial striatopallidal neurons suppresses goal-directed behavior. *Neuropsychopharmacology*, *41*, 1003–1013. https://doi.org/10.1038/npp.2015.227

Mailly, P., Charpier, S., Menetrey, A., & Deniau, J.-M. (2003). Three-dimensional organization of the recurrent axon collateral network of the Substantia Nigra pars reticulata neurons in the rat. *Journal of Neuroscience*, *23*, 5247–5257. https://doi.org/10.1523/JNEUROSCI.23-12-05247.2003

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, *459*, 837–841. https://doi.org/10.1038/nature08028

Menegas, W., Akiti, K., Amo, R., Uchida, N., & Watabe-Uchida, M. (2018). Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nature Neuroscience*, *21*, 1421–1430. https://doi.org/10.1038/s41593-018-0222-1

Menegas, W., Bergan, J. F., Ogawa, S. K., Isogai, Y., Umadevi Venkataraju, K., Osten, P., … Watabe-Uchida, M. (2015). Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. *Elife*, *4*, e10032.

Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, *126*, 292–311. https://doi.org/10.1037/rev0000120

Murray, J. E., Dilleen, R., Pelloux, Y., Economidou, D., Dalley, J. W., Belin, D., & Everitt, B. J. (2014). Increased impulsivity retards the transition to dorsolateral striatal dopamine control of cocaine seeking. *Biological Psychiatry*, *76*, 15–22. https://doi.org/10.1016/j.biopsych.2013.09.011

Nieh, E. H., Matthews, G. A., Allsop, S. A., Presbrey, K. N., Leppla, C. A., Wichmann, R., … Tye, K. M. (2015). Decoding neural circuits that

control compulsive sucrose seeking. *Cell*, *160*, 528–541. https://doi.org/10.1016/j.cell.2015.01.003

O'Hare, J. K., Ade, K. K., Sukharnikova, T., Van Hooser, S. D., Palmeri, M. L., Yin, H. H., & Calakos, N. (2016). Pathway-specific striatal substrates for habitual behavior. *Neuron*, *89*, 472–479. https://doi.org/10.1016/j.neuron.2015.12.032

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, *25*, 563–593. https://doi.org/10.1146/annurev.neuro.25.112701.142937

Patterson, T. K., Craske, M. G., & Knowlton, B. J. (2019). Enhanced avoidance habits in relation to history of early-life stress. *Frontiers in Psychology*, *10*, 1876.

Peça, J., Feliciano, C., Ting, J. T., Wang, W., Wells, M. F., Venkatraman, T. N., … Feng, G. (2011). Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature*, *472*, 437–442. https://doi.org/10.1038/nature09965

Pelloux, Y., Everitt, B. J., & Dickinson, A. (2007). Compulsive drug seeking by rats under punishment: Effects of drug taking history. *Psychopharmacology*, *194*, 127–137. https://doi.org/10.1007/s00213-007-0805-0

Richards, C. D., Shiroyama, T., & Kitai, S. T. (1997). Electrophysiological and immunocytochemical characterization of GABA and dopamine neurons in the substantia nigra of the rat. *Neuroscience*, *80*, 545–557. https://doi.org/10.1016/S0306-4522(97)00093-6

Rizzi, G., & Tan, K. R. (2019). Synergistic nigral output pathways shape movement. *Cell Reports*, *27*, 2184–2198.e4. https://doi.org/10.1016/j.celrep.2019.04.068

Root, D. H., Melendez, R. I., Zaborszky, L., & Napier, T. C. (2015). The ventral pallidum: Subregion-specific functional anatomy and roles in motivated behaviors. *Progress in Neurobiology*, *130*, 29–70. https://doi.org/10.1016/j.pneurobio.2015.03.005

Rossi, M. A., & Yin, H. H. (2012). Methods for studying habitual behavior in mice. *Current Protocols in Neuroscience*, *60*(1), 8–29. https://doi.org/10.1002/0471142301.ns0829s60

Rothwell, P. E., Fuccillo, M. V., Maxeiner, S., Hayton, S. J., Gokce, O., Lim, B. K., … Südhof, T. C. (2014). Autism-associated neuroligin-3 mutations commonly impair striatal circuits to boost repetitive behaviors. *Cell*, *158*, 198–212. https://doi.org/10.1016/j.cell.2014.04.045

Schreiner, D. C., Renteria, R., & Gremel, C. M. (2019). Fractionating the all-or-nothing definition of goal-directed and habitual decision-making. *J Neurosci Res*, https://doi.org/10.1002/jnr.24545.

Shan, Q., Christie, M. J., & Balleine, B. W. (2015). Plasticity in striatopallidal projection neurons mediates the acquisition of habitual actions. *European Journal of Neuroscience*, *42*, 2097–2104. https://doi.org/10.1111/ejn.12971

Shan, Q., Ge, M., Christie, M. J., & Balleine, B. W. (2014). The Acquisition of goal-directed actions generates opposing plasticity in direct and indirect pathways in dorsomedial striatum. *Journal of Neuroscience*, *34*, 9196–9201. https://doi.org/10.1523/JNEUROSCI.0313-14.2014

Shen, W., Flajolet, M., Greengard, P., & Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*, *321*, 848–851. https://doi.org/10.1126/science.1160575

Shmelkov, S. V., Hormigo, A., Jing, D., Proenca, C. C., Bath, K. G., Milde, T., … Rafii, S. (2010). Slitrk5 deficiency impairs corticostriatal circuitry and leads to obsessive-compulsive–like behaviors in mice. *Nature Medicine*, *16*(5), 598–602.

Shumilov, K., Real, M. Á., Valderrama-Carvajal, A., & Rivera, A. (2018). Selective ablation of striatal striosomes produces the deregulation of dopamine nigrostriatal pathway. *PLoS ONE*, *13*, e0203135. https://doi.org/10.1371/journal.pone.0203135

Singer, B. F., Fadanelli, M., Kawa, A. B., & Robinson, T. E. (2018). Are cocaine-seeking "habits" necessary for the development of addiction-like behavior in rats? *Journal of Neuroscience*, *38*, 60–73.

Smith, J. B., Klug, J. R., Ross, D. L., Howard, C. D., Hollon, N. G., Ko, V. I., … Jin, X. (2016). Genetic-based dissection unveils the inputs and outputs of striatal patch and matrix compartments. *Neuron*, *91*, 1069–1084. https://doi.org/10.1016/j.neuron.2016.07.046

Smith, K. S., & Graybiel, A. M. (2016). Habit formation coincides with shifts in reinforcement representations in the sensorimotor striatum. *Journal of Neurophysiology*, *115*, 1487–1498. https://doi.org/10.1152/jn.00925.2015

Smith, K. S., Virkud, A., Deisseroth, K., & Graybiel, A. M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 18932–18937. https://doi.org/10.1073/pnas.1216264109

Stelly, C. E., Haug, G. C., Fonzi, K. M., Garcia, M. A., Tritley, S. C., Magnon, A. P., … Wanat, M. J. (2019). Pattern of dopamine signaling during aversive events predicts active avoidance learning. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 13641–13650. https://doi.org/10.1073/pnas.1904249116

Surmeier, D. J., Plotkin, J., & Shen, W. (2009). Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. *Current Opinion in Neurobiology*, *19*, 621–628. https://doi.org/10.1016/j.conb.2009.10.003

Tepper, J. M., & Lee, C. R. (2007). GABAergic control of substantia nigra dopaminergic neurons. *Progress in Brain Research*, *160*, 189–208.

Tepper, J. M., Martin, L. P., & Anderson, D. R. (1995). GABAA receptor-mediated inhibition of rat substantia nigra dopaminergic neurons by pars reticulata projection neurons. *Journal of Neuroscience*, *15*, 3092–3103. https://doi.org/10.1523/JNEUROSCI.15-04-03092.1995

Thorn, C. A., Atallah, H., Howe, M., & Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*, *66*, 781–795. https://doi.org/10.1016/j.neuron.2010.04.036

Vandaele, Y., & Janak, P. H. (2018). Defining the place of habit in substance use disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *87*, 22–32.

Vandaele, Y., Mahajan, N. R., Ottenheimer, D. J., Richard, J. M., Mysore, S. P., & Janak, P. H. (2019). Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training, *ELife*, *8*, e49536. https://doi.org/10.7554/eLife.49536

Vandaele, Y., Pribut, H. J., & Janak, P. H. (2017). Lever insertion as a salient stimulus promoting insensitivity to outcome devaluation. *Frontiers in Integrative Neuroscience*, *11*, 23.

Vanderschuren, L. J. M. J., Ciano, P. D., & Everitt, B. J. (2005). Involvement of the dorsal striatum in cue-controlled cocaine seeking. *Journal of Neuroscience*, *25*, 8665–8670. https://doi.org/10.1523/JNEUROSCI.0925-05.2005

Vanderschuren, L. J. M. J., & Everitt, B. J. (2004). Drug seeking becomes compulsive after prolonged cocaine self-administration. *Science*, *305*, 1017–1019. https://doi.org/10.1126/science.1098975

Vanderschuren, L. J., Minnaard, A. M., Smeets, J. A., & Lesscher, H. M. (2017). Punishment models of addictive behavior. *Current Opinion in Behavioral Sciences*, *13*, 77–84. https://doi.org/10.1016/j.cobeha.2016.10.007

Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A., & Uchida, N. (2012). Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron*, *74*, 858–873. https://doi.org/10.1016/j.neuron.2012.03.017

Welch, J. M., Lu, J., Rodriguiz, R. M., Trotta, N. C., Peca, J., Ding, J.-D., … Feng, G. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. *Nature*, *448*, 894–900. https://doi.org/10.1038/nature06104

Willuhn, I., Burgeno, L. M., Everitt, B. J., & Phillips, P. E. M. (2012). Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proceedings of the National Academy of Sciences*, *109*, 20703–20708. https://doi.org/10.1073/pnas.1213460109

Witt, K., Nuhsman, A., & Deuschl, G. (2002). Dissociation of habit-learning in Parkinson's and cerebellar disease. *Journal of Cognitive*

Neuroscience, *14*, 493–499. https://doi.org/10.1162/089892902317362001

Woolverton, W. L., Freeman, K. B., Myerson, J., & Green, L. (2012). Suppression of cocaine self-administration in monkeys: Effects of delayed punishment. *Psychopharmacology*, *220*, 509–517. https://doi.org/10.1007/s00213-011-2501-3

Woon, E. P., Sequeira, M. K., Barbee, B. R., & Gourley, S. L. (2019). Involvement of the rodent prelimbic and medial orbitofrontal cortices in goal-directed action: A brief review. *J Neurosci Res*, https://doi.org/10.1002/jnr.24567.

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., … Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*, *462*, 915–919. https://doi.org/10.1038/nature08389

Yang, H., de Jong, J. W., Tak, Y., Peck, J., Bateup, H. S., & Lammel, S. (2018). Nucleus accumbens subnuclei regulate motivated behavior via direct inhibition and disinhibition of VTA dopamine subpopulations. *Neuron*, *97*, 434–449.e4. https://doi.org/10.1016/j.neuron.2017.12.022

Yin, H. H., & Knowlton, B. J. (2004). Contributions of striatal subregions to place and response learning. *Learning & Memory*, *11*, 459–463. https://doi.org/10.1101/lm.81004

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*, 464–476. https://doi.org/10.1038/nrn1919

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*, *19*, 181–189. https://doi.org/10.1111/j.1460-9568.2004.03095.x

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action—Outcome learning in instrumental conditioning. *European Journal of Neuroscience*, *22*, 505–512. https://doi.org/10.1111/j.1460-9568.2005.04219.x

Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., … Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*, *12*, 333–341. https://doi.org/10.1038/nn.2261

Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*, *22*, 513–523. https://doi.org/10.1111/j.1460-9568.2005.04218.x

Yu, C., Gupta, J., Chen, J.-F., & Yin, H. H. (2009). Genetic deletion of A2A adenosine receptors in the striatum selectively impairs habit formation. *Journal of Neuroscience*, *29*, 15100–15103. https://doi.org/10.1523/JNEUROSCI.4215-09.2009

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Transparent Peer Review Report

How to cite this article: Lerner TN. Interfacing behavioral and neural circuit models for habit formation. *J Neuro Res*. 2020;98:1031–1045. https://doi.org/10.1002/jnr.24581