Alternative Approaches to Understanding Habit Learning in the Dorsolateral Striatum

Kenneth A. Amaya and Kyle S. Smith

Abstract Our daily lives are full of habits. Yet, even though we know them intimately, habits are surprisingly hard to pin down in laboratory settings in order to study them and their basis in the brain. For over a century, scholarly work on habits has been dominated by the idea that they are, fundamentally, a routine of behavior that gets set in motion as a sort of reflex in a given circumstance. This is captured by the theory that habits reflect an association between a stimulus or context and a response. Habitual behaviors of this sort are very different from behaviors that are driven cognitively, with a goal in mind, which is argued to reflect an association between the behavior and the outcome that occurs. In this framework, we can identify habits as behaviors that remain unchanged when the outcome changes, as well as when other environmental changes occur, because they look reflexive and not cognitively guided. In the brain, this insensitivity of behavior to change is known to rely on networks that include the dorsal-lateral aspect of the striatum. Ergo, the dorsolateral striatum (DLS), is regarded as a home base for stimulus-response learning for habits. However, when one examines habits further, there are explanatory gaps. First, there are alternative options for defining habits based on behaviors that appear stuck and unchangeable and that deserve closer attention. Second, when one digs further into the function of the dorsolateral striatum—the habit zone of the brain—it is not stimulus-response learning that jumps out, but rather other types of activity that may cause behaviors to look like habits. We view the research field as in rather muddy waters for trying to operationally define what habits are and for trying to understand how areas like the striatum make behaviors appear habitual. Yet, having several different working models to lean on and advances being made in both brain analyses and behavioral analyses, progress in this domain seems promising.

K. A. Amaya
Tufts University School of Medicine, Boston, MA, USA
e-mail: kenneth.amaya@tufts.edu

K. S. Smith (✉)
Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA
e-mail: kyle.s.smith@dartmouth.edu

Y. Vandaele (ed.), *Habits*, https://doi.org/10.1007/978-3-031-55889-4_7

Keywords Striatum · Motivation · Optogenetics · Vigor

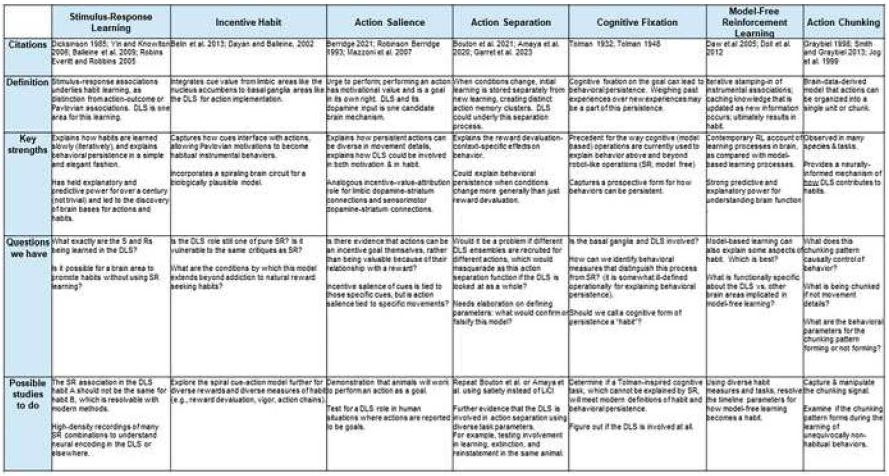
1 Introduction

Reward-based learning can be categorized as being reliant on either Pavlovian or instrumental (operant) conditioning. Within instrumental conditioning, two main types of behavioral strategies come into play: goal-directed behaviors and habits. Goal-directed behaviors are generally thought of as cognitively guided actions with a specific outlook on the goal, including the goal's value and its relationship with the actions. Using this behavioral strategy, the consequences of one's actions are weighed and used to determine whether to act. In contrast, habits are considered to be automatic stimulus-driven responses that are decoupled from the goal itself. Colloquially, habits are often casually categorized as "good" or "bad" and are spoken about as things that should be continued, like brushing one's teeth or going to bed early, or things that should be stopped or broken, like biting one's fingernails. The truth about habits is that their utility can quickly change based on external factors. For example, going to bed early might be good normally, but going to bed early on New Year's Eve when you plan to attend a party is a scenario where a good habit turns bad. One might need to transition back to being goal-directed to accommodate those New Year's plans. Or one might build different types of habits, some appropriate for workday bedtime routines and others for weekend party routines. In this vein, we consider two interplays critical to understanding instrumental behavior and its neural substrates. One interplay is the balance between goal-directed and habitual control over behavior, which has received a lot of attention in behavioral and neuroscience research. Another interplay is transitions between different types of habits depending on contextual factors and how one selects those habits to fit the different scenarios. This topic, namely, how habitual actions are contextualized and carried out appropriately, is a murkier area for research. In this chapter, we develop a few conceptual frameworks for understanding how habits can be controlled and selected as an active process in the brain, one resulting in part from activity in the dorsolateral striatum (DLS). We consider how traditional concepts of DLS function as promoting habits through stimulus-response (SR) learning can be unsatisfactory. Instead, we highlight some other concepts that could have explanatory power for understanding habits and how the DLS contributes to them. These notions include roles for the DLS in prospectively guiding habitual actions, contextualizing action plans to achieve behavioral inflexibility, and attributing motivational value to action plans.

2 Revisiting What SR Learning Tells Us About Habits and the DLS

To study the brain mechanisms underlying habit formation, scientists must be careful in how they determine if a behavior is a habit or not. We can consider a rat pressing a lever for food in an operant chamber. Just by looking at the animal, it is impossible to tell which of the behavioral strategies (goal-directed vs habit) it is using to solve the task. Cleverly, scientists have tested this by changing some aspect of the instrumental task and observing whether the behavior of the animal changes. In some cases, this has been done by changing the value of the outcome. Logically, if animals are relying on the cognitive, prospective goal-directed behavior, changing the value of the outcome should readily produce observable behavioral changes (less value, less behavior). Other times, the relationship between the action and outcome delivery is changed such that an action that once reliably produced outcome no longer produces outcomes in the same way, or even, outcomes are delivered on a schedule independent of the performance of the action. These measures highlight ways to differentiate between behaviors that might be goal-directed and behaviors that might be habits. Using these measures, there is now a wealth of evidence from multiple species and tasks showing that a brain area called the dorsolateral striatum (DLS; putamen homologue) and its connections are vital for behaviors to be expressed as habits. But with this wealth of information comes some curiosities.

Habits have been modeled as associations between a stimulus and the response it elicits (SR) (Fig. 1). Forming a habit is a process of repeating those responses when in the presence of those situations/contexts, thereby stamping in SR associations. SR associations lack a "goal" or "outcome" component, which is contrasted to goal-directed behaviors that rely on associations formed between an action and an outcome (AO). Therefore, as an SR association does not include an outcome (O) component, a habitual behavior based on SR associations should not be sensitive to changes related to the outcome. By extension, if a brain area underlies SR-based behavior, then this brain area should cause outcome-insensitivity in behavior. This is clearly seen in procedures such as reward devaluation, where the value of the outcome of an animal's behavior (i.e., the reward) is reduced. Animals can, particularly after a period of extensive training, show that they are habitual by continuing to engage in behaviors despite the outcome of those behaviors being devalued. Reducing DLS activity renders animals much more sensitive to the change in the outcome value (Yin & Knowlton, 2006; Balleine et al., 2009), providing strong evidence that the DLS underlies SR learning for habits; without a DLS, animals are always attending to the outcome. Similar findings have been obtained by changing the goal to include a new aversive stimulus (Jonkman et al., 2012) or by changing how the action and outcome are related to one another (Yin & Knowlton, 2006; Balleine et al., 2009). This has all been a very clear and convincing story in research. Yet, researchers have also pointed out that there are limitations in inferring SR habits from the behavioral measurements that we have at our disposal (De Houwer, 2019; Watson et al., 2022), that there are limitations to calling compulsive reward-seeking behaviors like addiction SR habits (Vandaele & Ahmed, 2021; Robinson & Berridge, 1993), and that there are functional aspects of habit-related brain function like in the DLS that are difficult to understand in SR terms (Smith & Graybiel, 2014). Below we raise and discuss several important questions.



	Stimulus-Response Learning	Incentive Habit	Action Salience	Action Separation	Cognitive Fixation	Model-Free Reinforcement Learning	Action Chunking
Citations	Dickinson 1985; Yin and Knowlton 2006; Balleine et al. 2009; Robins Everitt and Robbins 2005	Belin et al. 2013; Dayan and Balleine, 2002	Berridge 2021; Robinson Berridge 1993; Mazzoni et al. 2007	Bouton et al. 2021; Amaya et al. 2020; Garret et al. 2023	Tolman 1932; Tolman 1948	Daw et al 2005; Doll et al. 2012	Graybiel 1998; Smith and Graybiel 2013; Jog et al. 1999
Definition	Stimulus-response associations underlies habit learning, as distinction from action-outcome or Pavlovian associations. DLS is one area for this learning.	Integrates cue value from limbic areas like the nucleus accumbens to basal ganglia areas like the DLS for action implementation.	Urge to perform; performing an action has motivational value and is a goal in its own right. DLS and its dopamine input is one candidate brain mechanism.	When conditions change, initial learning is stored separately from new learning, creating distinct action memory clusters. DLS could underly this separation process.	Cognitive fixation on the goal can lead to behavioral persistence. Weighing past experiences over new experiences may be a part of this persistence.	Iterative stamping-in of instrumental associations; caching knowledge that is updated as new information occurs; ultimately results in habit.	Brain-data-derived model that actions can be organized into a single unit or chunk.
Key strengths	Explains how habits are learned slowly ('iteratively'), and explains behavioral persistence in a simple and elegant fashion. Has held explanatory and predictive power for over a century (not trivial) and led to the discovery of brain bases for actions and habits.	Captures how cues interface with actions, allowing Pavlovian motivations to become habitual instrumental behaviors. Incorporates a spiraling brain circuit for a biologically plausible model.	Explains how persistent actions can be diverse in movement details, explains how DLS could be involved in both motivation & in habit. Analogous incentive-value-attribution role for limbic dopamine-striatum connections and sensorimotor dopamine-striatum connections.	Explains the reward devaluation-context-specific effects on behavior. Could explain behavioral persistence when conditions change more generally than just reward devaluation.	Precedent for the way cognitive (model based) operations are currently used to explain behavior above and beyond robot-like operations (SR, model free). Captures a prospective form for how behaviors can be persistent.	Contemporary RL account of learning processes in brain, as compared with model-based learning processes. Strong predictive and explanatory power for understanding brain function	Observed in many species & tasks. Provides a neurally-informed mechanism of how DLS contributes to habits.
Questions we have	What exactly are the S and Rs being learned in the DLS? Is it possible for a brain area to promote habits without using SR learning?	Is the DLS role still one of pure SR? Is it vulnerable to the same critiques as SR? What are the conditions by which this model extends beyond addiction to natural reward seeking habits?	Is there evidence that actions can be an incentive goal themselves, rather than being valuable because of their relationship with a reward? Incentive salience of cues is tied to those specific cues, but is action salience tied to specific movements?	Would it be a problem if different DLS ensembles are recruited for different actions, which would masquerade as this action separation function if the DLS is looked at as a whole? Needs elaboration on defining parameters: what would confirm or falsify this model?	Is the basal ganglia and DLS involved? How can we identify behavioral measures that distinguish this process from SR? (it is somewhat ill-defined operationally for explaining behavioral persistence). Should we call a cognitive form of persistence a "habit"?	Model-based learning can also explain some aspects of habit. Which is best? What is functionally specific about the DLS vs. other brain areas implicated in model-free learning?	What does this chunking pattern causally control of behavior? What is being chunked if not movement details? What are the behavioral parameters for the chunking pattern forming or not forming?
Possible studies to do	The SR association in the DLS habit A should not be the same for habit B, which is resolvable with modern methods. High-density recordings of many SR combinations to understand neural encoding in the DLS or elsewhere.	Explore the spiral cue-action model further for diverse rewards and diverse measures of habit (e.g., reward devaluation, vigor, action chains).	Demonstration that animals will work to perform an action as a goal. Test for a DLS role in human situations where actions are reported to be goals.	Repeat Bouton et al. or Amaya et al. using satiety instead of LiCl Further evidence that the DLS is involved in action separation using diverse task parameters. For example, testing involvement in learning, extinction, and reinstatement in the same animal.	Determine if a Tolman-inspired cognitive task, which cannot be explained by SR, will meet modern definitions of habit and behavioral persistence. Figure out if the DLS is involved at all.	Using diverse habit measures and tasks, resolve the timeline parameters for how model-free learning becomes a habit.	Capture & manipulate the chunking signal. Examine if the chunking pattern forms during the learning of unequivocally non-habitual behaviors.

Fig. 1 Illustration of 7 approaches to understanding behavioral persistence and habit, and the potential function of the DLS. Highlighted from top to bottom are key citations, a definition, key strengths, questions we have, and possible experiments that could be done to help better understand each approach

2.1 *What Are the Ss and Rs?*

A possible starting point for understanding habits as consequences of SR learning, and the DLS as a neural instantiation of this process, is to consider how Ss and Rs links are formed. By the logic of SR learning, there must be a specific R (or set of Rs) that becomes reinforced and associated with a specific S. In some cases, this S-R association process seems straightforward. Two major flavors of habit research in neuroscience are operant lever-pressing behavior and maze-running behavior. In the operant version, animals might press a lever a certain number of times for reward. This is a simple action (lever press, R) triggered by a specific situation (task context, S). Animals learn this lever-pressing behavior rapidly, and after extensive training (or after using certain types of reward schedules; (Dickinson & Weiskrantz, 1985)), animals show us that they behave habitually by failing to change their response when the outcome changes. The DLS shows us that it is involved because when it is inactivated, animals are sensitive to such outcome changes when they otherwise would not be (Yin & Knowlton, 2006; Balleine et al., 2009; Malvaez & Wassum, 2018; Seger, 2018). In the maze version, a classic example is having animals traverse a plus-shaped environment. Animals start in one arm (e.g., north), and then they turn a direction (e.g., right) to get reward. If the animals are then placed in the opposite maze arm from the start arm (e.g., south), they show us that they are running habitually by making the same learned turn direction (e.g., turn right) rather than following spatial cues and going to the left side, where the reward had been physically located (Packard & McGaugh, 1992). In this case, the S is likely again the task context. As an aside, it is worth considering that the sensory milieu is presumably quite different when they move from north to south arms; yet, by still emitting the learned response of turning right, the S in such cases is the broader task environment itself. Regardless, this tendency of animals to perform the right turn when moved to the south arm increases with training time and is DLS-dependent (Packard & McGaugh, 1992; Packard, 2009). In other maze tasks, habits are also shown by animals undergoing extensive training and then being insensitive to outcome devaluation, with a key DLS role in this process as well (Jog et al., 1999; Smith & Graybiel, 2013; Crego et al., 2020). One can get even more simple with the S-R associations in understanding DLS function, such as simple motor movements like head bobbing (Tang et al., 2007) that can become outcome-insensitive. Still, some behaviors linked closely with reward consumption and lateral striatal function such as licking movements may not enter into habitual control, at least as judged by those movements remaining devaluation sensitive despite a lot of training (Tang et al., 2009).

There are other behaviors that require the DLS but challenge our ability to identify what either the Ss or the Rs are. The DLS plays a critical role in guiding behavior towards cues that are associated with rewards. In some tasks, animals must locate intra-task stimuli and approach them to achieve goals. This is called "win-stay" or "beacon" behavior, and it can become a behavior that is guided by the DLS (Sage & Knowlton, 2000; Berke et al., 2009; Kosaki et al., 2015; Crego et al., 2020). The S is obvious in this task (the cue). The R would seem to be the cue-initiated maze-turning behavior. However, to us, the R for the animal seems cue-directed, and to be more of a natural repertoire of approach behaviors. This type of cue-triggered approach bears some similarity to Pavlovian conditioned responses, particularly those that are measured as cue approach (e.g., sign-tracking as below), perhaps more so than similarity to classical instrumental actions. Thus, behaviors directed to cues can be DLS-dependent, but may not be clearly SR in nature.

In reverse, an example of a behavior that requires the DLS but might lack a clear S is grooming. Rodents exhibit a pattern of grooming that follows a chain of movement events as a sort of syntax. Performing this grooming routine in a patterned form requires the DLS (Cromwell & Berridge, 1996). In this case, the R is obvious: the movement pattern. The S is less obvious. The S cannot just be tactile feedback from the grooming movements, as the grooming patterns persist despite tactile feedback being removed via limb amputation (Fentress, 1973). Grooming can be evoked by stimuli like moisture on the body, but in most circumstances, the grooming appears to happen spontaneously. What it the S to trigger grooming patterns? Is it an internal impetus of some nature to carry out the full sequence of grooming movements? In many ways, grooming bears resemblance to another facet of DLS function, which is learning and performing skills. A few examples of skills that involve the DLS include rotarod learning (i.e., staying on a rotating rod to avoid falling) (Yin et al., 2009), performing an action a set number of times at a set rate (Jin & Costa, 2010), and running mazes a certain way to get reward (Packard & McGaugh, 1992; Jog et al., 1999) or to escape (Asem & Holland, 2015). Although it is rare to find studies that address both skill learning and habit learning in the same tests, which are needed to understand how they relate to one another, most habits do involve a skill aspect, and so it is not hard to understand why the DLS would be involved in both. However, there are cases in which measures of skill and measures of habit do not correlate perfectly (e.g., Smith & Graybiel, 2013; Vandaele & Janak, 2023; Garr & Delamater, 2019), suggesting a need for further research in this domain.

Finally, there are some behaviors that involve the DLS that arguably do not have SR underpinnings. One example is Pavlovian-to-instrumental transfer (PIT). Animals learn that a Pavlovian cue predicts reward, and separately that an action leads to reward. Then, when the cue is presented along with the action option for the first time, the action behavior is potentiated when the cue occurs (the "PIT effect"). Animals exhibit this PIT effect without ever having received explicit pairings of the cue and action together for associative learning. Yet, the DLS plays a critical role in this type of cue-potentiated action (Corbit & Janak, 2007), despite it not being explicitly learned. Still, it is possible that the Pavlovian cue, represented in limbic regions, engages a habit-like behavioral function of the DLS through wider circuits, plausibly involving nucleus accumbens-DLS coupling (Belin et al., 2013). An additional example is a behavior called sign-tracking, which is also DLS-dependent (DiFeliceantonio & Berridge, 2016; Naeem & White, 2016) (but see Fraser & Janak, 2017). Here, the insertion of a lever predicts reward delivery in a Pavlovian manner. Animals do not need to do anything for reward to occur. Yet, what many animals do is to approach the lever cue, then bite it, sniff it, and grab it as though it were food. This interaction with the lever cue is an appetitive process consisting of innate approach and consumption-like behaviors. For sign-tracking, the S is obvious (lever insertion), but the R is hard if not impossible to assign to a predictable pattern of muscle movements. What is the R in sign-tracking then? At the moment, it seems to be an ill-defined repertoire of cue-directed approaches and consummatory strategies. Notably, the sign-tracking response is not strictly an instrumental behavior. Although sign-tracking might plausibly be learned superstitiously as a behavior that must be done for a reward to occur, there is evidence that it does not result from such superstitious instrumental learning (Locurto et al., 1976; Chang & Smith, 2016; Ahrens et al., 2016). Thus, it is unclear what aspect of the R in sign-tracking is being reinforced in an instrumental learning way for an SR habit. Instead, for both PIT and sign-tracking, a more compelling account of these behaviors is to describe them in terms of incentive motivation rather than in terms of learned instrumental habits, despite the neural underpinnings of these behaviors involving the DLS (see also below for concepts of action salience and incentive habits).

Collectively, with these examples, we arrive at a set of behavioral phenomena that the DLS participates in, with some that have straightforward SR associations underlying them, some that have fuzzier SR associations, and some that may not have SR associations. However, we point out that even simple SR situations can be unclear. Within the context of simple instrumental learning like pressing a lever for a reward, the S that triggers the lever-pressing R is usually thought of as being a task context or the sight of the lever. This makes the most sense in free-operant conditions, in which the lever is always present for animals to press at their leisure. However, it is speculation what the S actually is in such cases.

2.2 What Does DLS Neural Activity Tell Us?

Another approach to this question of SR learning in the DLS is to take the perspective of the DLS itself: what do DLS neurons respond to during behavior and across learning? To begin with, the lateral striatum contains a topographical organization of neurons tuned to sensorimotor information in specific parts of the body (Künzle, 1977; Crutcher & DeLong, 1984; Selemon & Goldman-Rakic, 1985; Carelli & West, 1991). This organization would seem to set the stage nicely for DLS to represent the R component of a habitual behavior (when those neurons are engaged by an S). However, there is considerable plasticity in how these body-tuned activity patterns respond with repeated movement, such as occurring in task learning. For example, in rodents, neurons that respond to a specific body movement can actually decrease in activity as that specific movement is repeated (e.g., Carelli et al., 1997; Jog et al., 1999; Tang et al., 2007). In studies that engage more complex bodily movements to perform a response in a task, DLS neurons that represent that response can become linked to the stimuli that engage those responses in a nicely associated manner. In other words, neurons may be active during a particular response that was emitted by a particular cue (Stalnaker et al., 2010). However, the same type of SR-appearing neuronal response patterns are seen simultaneously in the medial striatum, an area that favors AO learning and purposeful behavior in opposition to SR-based habits. Similarly occurring DMS and DLS activity patterns with extensive training are seen in some other studies too (Vandaele et al., 2019), and there is an argument to be made for considering DMS and DLS activity as copiloting some ongoing behaviors (Schreiner et al., 2020). However, DMS and DLS activity do not always align with one another during habits (e.g., Thorn et al., 2010; Gremel & Costa, 2013; Vandaele & Janak, 2023). There is also some intriguing new evidence that DMS dopamine activity may help promote habits (Seiler et al., 2022; van Elzelingen et al., 2022), contrary to the AO/SR divide thought to delineate DMS/DLS function.

There are also studies showing that a predominant DLS activity pattern that develops with habit formation is one that emphasizes action "chunking" (Fig. 1). In 1999, Jog, Graybiel, and colleagues (Jog et al., 1999) discovered this striking pattern of activity in the DLS when animals performed a maze task. With training, DLS neuronal activity became accentuated at the start and end of the maze-running behavior, which was interpreted as potentially chunking together the running actions into a unit (similar to the way in which we remember some information like phone numbers) (Graybiel, 2008). Since then, this chunking pattern of activity has been observed in multiple DLS cell types, in multiple different types of tasks, multiple species, and several different brain areas beyond the DLS (Smith & Graybiel, 2016). The majority of studies in which this chunking pattern emerges in the DLS during behavior involve animals learning to perform a behavior for reward, and it aligns fairly well with the development of skillful patterns of movement. Although it emerges in parallel with the development of habitual running as assessed through outcome devaluation testing (see also below), it is most closely correlated with the speed and fluidity of behavior (Barnes et al., 2005; Smith & Graybiel, 2013). One might interpret this to reflect its important role in skillful behaviors that can, in some conditions, favor habit formation. Within the cells that contribute to this chunking activity pattern, which is roughly 2/3 of recorded projection cells, there are subtypes that respond to run-start, turn, run-stop, or combinations thereof (Barnes et al., 2005; Smith & Graybiel, 2013). Some cells do show modulation of their activity to task details like turn direction. However, by and large, this DLS activity pattern seems curiously independent from important task details, such as accuracy of responding, changes in behavior after task changes, and whether responding is sensitive or insensitive to changes in the goal value (Smith & Graybiel, 2013). For example, in one study, animals learned to traverse a T maze and to turn a certain direction based on an auditory cue they encounter while running. As they learn, DLS activity in fast-spiking interneurons responds robustly to the start and end of the behavior (see "chunking" below) (Kubota et al., 2009). Then the cue is suddenly switched in modality from auditory to tactile. Animals must learn this cue-turn association anew, and they make a lot of errors followed by a period of improvement. During this learning time, however, DLS activity at the start and end of behaviors does not change much (though DLS neurons do also develop a new firing response to the cue-turn decision point). One might argue that the overall response schema (run and turn) is familiar and unchanged in the task and that the key events that maintain a stable DLS activity response—run start and run stop—remain the same. Yet, a similar study using T-maze behavior has also found a lack of relationship between such start- and end-related DLS activity patterns to major changes in task performance after reward devaluation (Smith & Graybiel, 2013). In that study, animals learned to run a maze and, depending on a mid-maze cue, turn right for one reward (e.g., chocolate) or left for a different reward (e.g., sucrose). Then, one of the rewards was devalued and animals were returned to the maze. They first behaved habitually by running to the devalued goal when so instructed. Then, with repeated experience with the now-devalued goal on the maze, they shifted their behavior to avoid the devalued goal. During this period of major behavioral change, new learning, and a clear switch to AO-guided behavior, DLS activity during the maze runs hardly changed at all. In fact, it did not even distinguish different trials in which animals showed different levels of sensitivity to the devaluation. On different trials, animals might run to the devalued reward when so-instructed, run to the valued reward when so-instructed, or run to the valued goal location (where there was nothing) when instructed to run to the devalued reward. DLS activity was similar on each of these different types of trials. It is difficult to imagine how this DLS activity functioned to promote SR behavior on the maze when it was present in clearly non-habit behavioral conditions, such as during maze runs in which animals were avoiding the devalued goal in a seemingly purposeful manner. It is possible, however, that the DLS was helping control behavior prior to reward devaluation, while after reward devaluation, when behavior changed, the stable DLS representations of the task behavior were not “used” by the brain, and instead, other brain circuits for purposeful behavior were used.

Even so, there are arguments to be made that the DLS carries an active, online influence over actions that can either be more goal-directed or habitual. This influence seems particularly related to the *vigor* of performance. In the above study with reward devaluation (Smith & Graybiel, 2013), the activity at run-start was correlated at a tight, trial-level with measures of running vigor, including early points in learning (i.e., prior to habit expression proper), later habit expression, and even later return to goal-directed behaviors after reward devaluation. Notably, the stronger this start-related DLS activity was the more routed animals’ runs were and the less likely they were to exhibit signs of deliberation in vicarious trial-and-error head turns at maze decision points. This finding aligns with several related studies that have shown a role for the DLS, and its nigrostriatal dopaminergic inputs, in performance vigor for different varieties of ongoing movement routines (Aldridge et al., 2004; Haith et al., 2012; Barter et al., 2015; Panigrahi et al., 2015; Dudman & Krakauer, 2016; Dodson et al., 2016; da Silva et al., 2018), and it is also reminiscent of the movement control problems that result from nigrostriatal deterioration in Parkinson’s Disease (e.g., Mazzoni et al., 2007).

To gain a sense of how this start-related DLS activity causally controls habitual behaviors, we (Crego et al., 2020) transiently inhibited or excited this signal in several maze tasks. First, animals were trained on a response learning task, in which they learned to turn a single direction (e.g., right) on a plus-shaped maze for reward. The stimulation of the DLS for 0.5 sec at the start of runs made the full runs speedier and less deliberative. Stimulation in the middle of the runs, where DLS seemed to not be very active in prior work, did not have this effect on behavior. Reward devaluation showed that animals were behaving habitually, but run-start stimulation of the DLS was still able to increase the vigor of their habitual performance. Conversely, inhibition at run-start (but less so mid-run) tended to slow behavior and increase deliberations. This was especially true after reward devaluation. Animals with DLS inhibition at run-start mostly refused to run, thus showing an increase in outcome-sensitivity and a return to goal-directedness in their behavior. When they did run, they were far slower and deliberated on nearly every trial. We reasoned that the DLS manipulations at run start were particularly effective because animals could select and initiate their full running routine at that point. In other words, they would know from the start of a trial to go forward, turn right, and go forward again to get reward. To test this idea further, we exposed animals to a beacon task. In this task, they exited the start arm to reach the center of the plus maze and then needed to locate an intra-maze cue that signaled the location of the reward. To do this accurately, animals needed two action decision points—one to start running and one to approach the cue once it was located. In this case, DLS manipulations at run-start as well as at the cue location period produced the same changes in performance vigor. This set of results supports the idea that DLS activity at the point of choosing and initiating an action controls the level to which that action is vigorous and outcome-insensitive.

In both of the above tasks, running accuracy was not affected by the transient DLS manipulations. This was curious as response learning as well as learning beacon-directed behaviors is reduced with DLS perturbations like GABAergic inactivation or lesions (Asem & Holland, 2015; Packard & McGaugh, 1992; Yin & Knowlton, 2006; Balleine et al., 2009; Malvaez & Wassum, 2018; Kosaki et al., 2015). The dissociation between run accuracy being unaffected and vigor being affected in this work suggests that there are other epochs in performance or learning time in which the DLS must be active to learn and recall the task rules that is separate from the action selection points that we manipulated. On this point, much work remains to be done on other aspects of DLS action chunking pattern and how it relates to habitual behavior. It is unclear what role run-stop activity has, nor is it clear which circuits the heterogeneity in cell responses that make up this chunking pattern come from. There are even DLS cells that develop responses to the outcome of behavior (e.g., to reward or to lack of reward), and these outcome-related responses change with learning and habit formation (Schmitzer-Torbert & Redish, 2004; Smith & Graybiel, 2016). Moreover, there is growing appreciation for DLS activity in post-task consolidation windows for action learning (Goodman et al., 2017; Smith et al., 2021). How each of these different DLS signals relates to habits is a point of continued interest.

Collectively, the evidence that the DLS controls SR learning is underwhelming. Instead, we favor emphasizing the importance of the DLS in dictating performance vigor, at least in its activity at the time of action selection and initiation. This control over vigor appears to be ongoing and on-line. DLS activity seems less correlated at the trial level to outcome-insensitivity that can characterize habits, but outcome-insensitivity is clearly still under DLS control in a manner that can be linked—at a broader level—to the performance vigor that arises from the DLS (see also below). Thus, if SR learning does reside in the DLS, it has some explaining to do. When one takes the perspective of the DLS for how measures of habitual behavior can arise, SR learning representations do not jump out. That said, there is little point in abandoning an SR account of DLS function, as it certainly may be happening for some behaviors or in some pattern of neural activity or in some complex input-output circuit architecture, and, even if finding that mechanism remains elusive, the SR account of DLS function remains parsimonious as an interpretation of many dozens of studies. In this context, there are suggestions to consider brain areas that are important for habits less as uniform "habit" zones, and more as zones with multiple functional roles that might be connected with different circuit dynamics (Smith & Graybiel, 2016; Turner & Parkes, 2020; Watson et al., 2022). We simply hope to point out that finding the Ss and Rs that become associated with the DLS is a point of great interest in understanding its contributions to habitual behavior. It seems the Rs can be a set of muscle movements or a more complex "plan" or "schema" of behavior, it seems like the Ss can be discrete stimuli or else more hard-to-pinpoint representations of task environments or internal states, and it seems like the DLS can control aspects of behaviors that are both apparently habitual and apparently not habitual. It remains elusive how Ss and Rs are represented, reinforced, and linked up within the DLS, and this elusiveness is a point worth making in its own regard about an SR learning account of the DLS.

3 Revisiting What Behavioral Inflexibility Tells Us About Habits and the DLS

3.1 Conceptual Considerations

The above research raises the likelihood that the DLS can carry an active influence over ongoing behavior. In this section, we examine this notion further. The question to us becomes whether the DLS promotes the use of past action plans as habits (a retrospective outlook) or whether it is more actively engaged in habitual performance (a more prospective outlook). At face value, this may sound like a silly question. Habits are, essentially by definition, behaviors that we do automatically, semiconsciously, and consistently from occurrence to occurrence. We form behaviors that can, when they become habits, simply be set in motion when appropriate (or occasionally when inappropriate). This is the view of habits that has a retrospective orientation—we repeat what was done before. Ergo, a brain area that controls habits in this way must do so by triggering previously learned behaviors in a reflex-like manner, with a retrospective "do what was learned" focus. This is certainly a true form of habit that must have a brain basis. However, there is room for considering that there can also be a brain basis of habits that actively controls them in an online, forward-looking manner to help them be executed appropriately. Part of the reasoning behind this argument comes from different concepts for how behavioral fixity of the habit sort can arise from processes unlike canonical habit learning, and part of the reasoning comes from studies showing that the DLS may—at least in some conditions—have a prospective role for planning and guiding habits (Crego et al., 2023).

In early debates regarding SR versus cognitive control over different types of behavior, Tolman conducted a number of studies to show how behavior could be purposeful and goal-directed (Tolman, 1932, 1948). There were examples that seemed to show habit-like fixity in behavior, in which animals would continue to do something in a task despite it becoming disadvantageous. For example, animals may perseverate when task rules are changed. Instead of explaining such behaviors as SR-based, Tolman suggested that they were the result of cognitive fixation (i.e., purposefully following signs to get a goal in accordance with an understanding of the environment, but being fixated on how to go about doing it because earlier motivations were strong) (Fig. 1). As he writes, *"If rats are too strongly motivated in their original learning, they find it very difficult to relearn when the original path is no longer correct"* (Tolman, 1948). He likened it to similar forms of fixated behaviors in animals introduced to mazes, with some simply preferring to go a certain direction in a way that interferes with learning. By this logic, with learning, we can consider that animals build up expectations and behave purposefully based on goals. When there is a sudden addition of something new to the environment like an aversive component to the goal, animals might actually understand this new component but weigh the prior experiences above it as a form of Tolman-esque goal fixation (see also: Dayan & Berridge, 2014). Indeed, we now know that, for the habit measure of resistance to outcome devaluation after extended training, more experience with the outcome in its initially valued state makes it more likely that the outcome will continue to be pursued later after devaluation (Adams, 1982; Dickinson et al., 1995). As Dickinson et al. (1995) have shown and written, "*. . . the critical determinant of the overtraining effect [devaluation insensitivity] is the number of exposures to the pellets when the animals are hungry during training rather than the amount of instrumental conditioning."* Therefore, when animals continue working for a newly devalued outcome by "habit," one could argue that it is because they have had a long history of that outcome being valued, and that the probabilistic likelihood of it still being valuable is high. Zooming forward to today, we face a similar conundrum in many measures of habit. When animals continue to perform as they had learned to perform despite a new devaluation of the outcome, or despite no longer needing to perform the action when reward starts to occur freely, it may be hard to know if they are stuck in a habitual mode of responding or instead cognitively engaging with the task, considering what was learned versus what is new, maybe actively exploring newly imposed task conditions, or maybe even behaving purposefully based on prior (now experimentally "wrong") AO associations. These factors have been pointed out as a possible explanation of habit-like behavioral persistence in some cases, as have other factors like task confusion or problematic task inferences (Balleine & Dezfouli, 2019; Watson et al., 2022). While it is compelling to suggest that experimentally "wrong" persistent behaviors occur because animals are stuck in their old learned SR routines, there are other explanatory options that, in principle, can lead to such continued behaviors, including the process of prospective thinking to carry out a set of actions, be those learned rituals or new plans (Dezfouli & Balleine, 2012). One amusing demonstration of how habit-like behaviors could potentially be acquired as a form of cognitive fixation was a recent study that challenged rodents seeking drugs to solve a new set of puzzles every day to get those drugs (Singer et al., 2018). Despite the fact that this puzzle-solving challenge precluded the possibility of forming clear SR learning rules to get drugs—indeed the animals had considerable difficulty learning the task—the animals still developed many markers of addiction habits, which notably included drug seeking despite the addition of aversive consequences. In this study, animals with a fixation on getting drugs developed habit-like inflexible behaviors presumably in the absence of SR learning (however, in this case, behavior was not clearly linked to the DLS). Scientists recognize that identifying behaviors as habits if they persist despite task changes is a negative indicator, which carries inherent limitations. Essentially, we infer that animals are behaving habitually when we change the task or outcome but see that the animals do not change their behavior. In other words, when animals are not performing an instrumental action in an AO manner, such as by continuing to work for a devalued outcome or failing to adjust to a task change, then they must be performing by SR habit as a default (assuming control conditions rule out other factors like confusion). We need research in several domains to better flesh this out. First, we need more attention to the details by which animals respond. When they continue responding after a task change, are they responding in the exact same way as before or are they responding in new ways? Is the structure of the response the same or different? Simple measures of lever press choice and pressing rates, which are commonly reported performance indices, do not give us this. Second, we need to know if a continued response is prospective (forward-looking, more cognitive-like) or retrospective (backward-looking, more habit-like). Are animals reflexively behaving as they had or are they instead attuned to the task in a more purposeful manner? By extension, when an area like the DLS encourages behaviors to be continued despite task changes, we need to better resolve whether that area is promoting the use of learned routines (SR habits) or instead promoting exploratory/purposeful plans of action (not SR habits). In the next sections, we cover some recent insights into this.

3.2 A Closer Look at Outcome Devaluation

Despite a lot of work having been done since the era of Tolman, it is still difficult to tell if a behavior that persists after a change in task conditions is truly habitual or not, in the strict sense of habits being automatic and rather mindless reflexes. This issue has come into sharp focus due to the recent work on the "gold standard" measure of habits, outcome devaluation insensitivity. As above, habit research has traditionally used a test of sensitivity to outcome devaluation to assess whether animals are reliant on a habit to solve a task in the lab. Generally, two types of outcome devaluation have been used: devaluation by satiety and devaluation by conditioned taste aversion (CTA). In the former, animals are usually given free access to the outcome prior to behavioral testing, allowing them to become sated on that outcome. Then during the post-devaluation probe, animals are tested, often in extinction, to see if responding rates changed compared to basal levels. Similarly, outcome devaluation by conditioned taste aversion involves reducing the value of the outcome and testing animals afterward to see if their behavior flexibly changes to meet that reduced value. The difference in this method, however, is that rather than allowing animals to sate themselves on the outcome, the outcome is made explicitly aversive by inducing nausea. These two devaluation methods are distinct in that the outcome becomes relatively valueless (consider a value of $+1$ becoming a 0 during satiety) during satiety-based devaluations while in CTA, the outcome becomes, at least somewhat, aversive (value of $+1$ becoming negative). However, the crux of what is happening behaviorally remains similar because what was once valued is now valued less so, and behavior is tested to ascertain whether animals reduce instrumental responding (are goal-directed) or maintain performance despite these value decrements (habit).

Over time, articles have noted limitations to the outcome-devaluation-insensitivity measure of habits. It is a null result, and as a null result, habits may be one of many interpretations for why animals are insensitive to outcome value changes. Devaluations do not always transfer perfectly across contexts and are rarely permanent, so there would be some residual positive value that could drive continued outcome pursuit. Test phases are also commonly limited to very brief assays to see if animals will pursue a devalued outcome under extinction conditions, which is a task context that is rather unlike the task context they initially learned in. Even in such test phases, devaluation-sensitive behavioral evidence can be small in effect size and rarely if ever appears as a full cessation of behavior (Holland, 2008). Some behaviors also never develop devaluation insensitivity, such as behaviors in more complex task scenarios (Colwill & Rescorla, 1985), despite the fact that complex task behaviors can seem very well-learned.

Even beyond these caveats to the use of reward devaluation per se to identify habits, two recent studies—one examining cue-approach behavior (Amaya et al., 2020) and another examining lever-pressing behavior (Bouton et al., 2021)—found that a common way in which the CTA method is used to study outcome sensitivity/insensitivity can possibly lead to entirely wrong conclusions about habits. Specifically, if LiCl-reward pairings to devalue the reward are done in a non-task context, as is typical in experiments, animals placed back in the task after CTA learning can look habitual and show outcome-devaluation-insensitivity. If the CTA is instead done in the task context, animals will show that they are clearly not habitual and exhibit outcome-sensitivity in their behavior. This effect is not due to LiCl-induced nausea being associated with the testing room itself because LiCl exposure in the testing room without any reward consumption does not affect behavior (Bouton et al., 2021). The effect is also not due to a failure to learn and remember the LiCl-food-nausea association, because animals exhibit this learned aversion on the task when rewards are ultimately encountered (Amaya et al., 2020; Bouton et al., 2021). Thus, whether animals appear habitual or not can depend on where reward devaluations occurred. This is problematic. However, these studies do point to a fascinating role for the generalization of outcome values across different environments as a potential factor in the control of actions and habits (see below). Also, we should recognize that other forms of devaluation (e.g., satiety) and other types of tasks (e.g., maze running) have not been investigated in this way to our knowledge, and those conditions may or may not prove to be different.

3.3 A Role for Prospective Action Planning and Contextualizing in the DLS

In the above section, it was discussed that animals exhibit outcome-insensitivity when the CTA to devalue the reward is conducted in a context different from the task context, but they exhibit outcome-sensitivity when the CTA is done in the task context (Amaya et al., 2020; Bouton et al., 2021). This means that the persistence that animals show in a task after CTA, when the CTA is conducted elsewhere, is related to poor generalization of that CTA learning to the task. We highlight a recent effort to use models of action and habit learning to account for this phenomenon (Garrett et al., 2023). In reinforcement learning models, one framework holds that actions can be either model-free or model-based. Model-free learning is thought to be an iterative process, by which behaviors are learned through positive or negative feedback. This type of learning can underlie habits, and so, habits can be one destination of model-free learning systems. In contrast, model-based learning involves a more cognitive process of constructing a representation of the world and its possibilities. The argument in this new modeling effort (Garrett et al., 2023) is that the apparent devaluation-insensitivity we see when CTA is done outside of the task environment could be a model-based process. The devaluations done elsewhere create a distinct context of learning that becomes separate from the context of initial learning when the reward had still been valuable. This process is analogous to how extinction learning does not erase initial learning, but rather initial learning is preserved and extinction learning becomes a different learned experience. Animals then use history and context as important factors in either expressing or suppressing that initially learned behavior, be it after extinction or devaluation. The decision to continue performing actions despite outcome devaluation can thus be regarded as a cognitive type of response plan.

One way of viewing this all is that the brain can create an action context, and when new experiences come along, the brain can learn those in a separable context than initial learning. This may be an explicit contextual dissociation, such as when the new information is literally in a new context as in the case of CTA learning above. Theoretically, the context dissociation could also be less explicit such as when animals learn a behavior and contextualize it separately from learning that occurs when new information is added to the task. For example, when task conditions change (e.g., a task rule shift) or when the reward changes (e.g., when an aversive component is added to it), animals may parse out the initial learning from the new learning that occurs with the new task information and keep those learned situations separate.

This idea of contextualizing actions may have some explanatory power for understanding the DLS (Fig. 1). Recall that inhibition of the DLS renders animals more sensitive than they otherwise would be to changes in a learned task. For example, animals without a DLS are more sensitive to changes in action-outcome contingency (Yin & Knowlton, 2006; Balleine et al., 2009) and to the addition of an aversive stimulus to the goal (Jonkman et al., 2012). Similarly, animals are devaluation-insensitive if CTA learning is done outside of the task chamber, but DLS inhibition creates outcome sensitivity in that case (Yin & Knowlton, 2006). In normal situations, perhaps the animals are separating their initial learning from new learning when information changes (e.g., reward is devalued, task rules change, or an aversive stimulus is added). In these situations, they are "habitual" because they are behaving based on the old conditions rather than on the new ones. Perhaps the DLS is enabling this process. Thus, without a DLS, animals no longer separate old from new information, which leads to the information collapsing together and the most current (i.e., recent) information dominating behavior. In the devaluation situation in particular, even over-trained animals will show devaluation sensitivity if the CTA learning is done in the task chamber, thus making the new CTA information more directly relevant and embedded into the task information that had been learned (Amaya et al., 2020; Bouton et al., 2021). Perhaps embedding the CTA learning into the task conditions is very similar to what happens when the DLS is inhibited—new information becomes more integrated with old information to be used in behavior. In other words, the DLS could separate out initial task learning from the new CTA learning that occurred elsewhere, and without a DLS that separation of learning does not happen—the old and new learning are collapsed together as though the CTA was part of the task context when it was not. We note that another common form of reward devaluation, satiety by pre-feeding, has not been systematically tested to see if animals show satiety-insensitivity if the pre-feeding was done in the task chamber compared to it being done in a different context. It could be that satiety and LiCl devaluations are exactly the same. Or, it could be that satiety is different and that animals will still show devaluation insensitivity even if the satiety was done in the task chamber, in which case, this logic does not apply as well to the satiety assay of habit.

One prediction of the concept that the DLS parses action learning contexts is that it should be related to habit extinction, and indeed it is. In one study (Goodman et al., 2017), animals learned a response-based maze-running behavior, which is known to require the DLS. Reward was then removed and extinction of responding occurred. However, with the DLS inactivated during a post-task consolidation window, animals did not extinguish as well. In addition, by the logic of DLS parsing action contexts, it ought to play a role in the renewal or reinstatement of habitual reward seeking after extinction. One study on drug-seeking behavior (though not obviously habitual drug seeking) suggests this might be true (Bossert et al., 2009). Animals were trained to seek heroin in context A and then extinguished in context B. Placement back in context A caused a reinstatement of heroin seeking. Rats with a blockade of D1 dopamine receptors in the DLS showed less of this context-induced reinstatement in a manner not easily explainable by a motor deficit. The D1 blockade in the DMS did not interfere with reinstatement. Likewise, in this view of the DLS, it is ought to be important not only for response learning but also for changing that learning to a new response, such as in reversal learning tasks, and evidence suggests that this can be the case (Jackson et al., 2019; Bergstrom et al., 2020). We should be cautious in drawing major conclusions from a handful of studies that have many methodological dissimilarities. Still, the data begin to support a view that the DLS could plausibly be important for habit learning, performance, extinction, and reinstatement, and could be similarly important for learning a behavior and then for learning a new behavior that overrides the first one. We offer the action contextualization idea as one way to explain this diversity of DLS function. Systemic examinations of the role of the DLS in multiple stages of behavioral learning and behavioral change will be important to do in the future (Fig. 1). We add that another habit-promoting structure, the infralimbic cortex in rodents, could also be a part of an action-contextualization network for habits. In one study (Smith et al., 2012), the inactivation of the infralimbic cortex after reward devaluation caused animals to avoid the devalued goal more, thus reducing the habit. However, once animals had been avoiding the devalued goal for a while, further infralimbic inhibition caused animals to go back to pursuing the devalued goal as though to reinstate the old habit. This might be explainable if the infralimbic cortex were helping toggle different action strategies to promote the newest ones.

This line of thinking should require the DLS to be, at least sometimes, a structure for prospectively selecting and guiding actions in a more active form that might be appreciated from SR learning accounts alone. There are ways in which the DLS could be prospective in function. First, the DLS is routinely implicated in action learning and in the honing of ongoing behavior (e.g., Packard & McGaugh, 1992; Bailey & Mair, 2006; Yin et al., 2009; Bergstrom et al., 2020; Watson et al., 2022; Lerner, 2020) including as mentioned the extinction of habits (Goodman et al., 2017). This set of work would suggest a possible role for the DLS in the prospective planning of actions. Second, because many studies examine DLS activity when task conditions are stable, it becomes impossible to know if manipulations of the DLS that reduce an aspect of performance do so because planning has been affected or prior learning has been affected. Third, most prior inactivation procedures (e.g., lesions, pharmacologic inhibitions) would have affected DLS activity before, during, and after an action, leaving unresolved what causal role DLS activity during the action itself has—such as the neural "chunking" representations noted above.

In a recent unpublished study (Crego et al., 2023), we examined the possibility that the DLS could hold a prospective engagement in planning habit-like actions to help parse out new versus old action learning contexts. To do this, animals were trained on a simple FR1 lever-press task for reward over 7 days where each press led to reward. During this time, the animals developed a predictable routine of 1-press bouts (i.e., pressing once and then checking for reward). Once this routine was stamped in, we then suddenly increased the task demands to an FR3, which now required three presses for reward, while inhibiting the DLS during the lever pressing. Control animals showed signs of learning by increasing 3-press bouts. DLS inhibition reduced this, causing animals to continue acting as though they were in an FR1 task and pressing in 1-press bouts. Animals with DLS offline were favoring their initial learned action plan in a retrospective sort of manner. After additional FR3 training, the DLS was inhibited again, but without any task change. This inhibition did nothing to action strategies. Later, animals were shifted back from FR3 to FR1. During this shift, animals with DLS inhibition again favored 1-press bouts and less 3-press bouts. This again showed that DLS inhibition caused animals to revert to their initially learned action routine. Collectively, it appeared that an active DLS allowed control animals to learn new task rules better and change their behaviors, but without a DLS behavior collapsed back to what was first learned. We suggest these findings show that the DLS helps animals parse out new action learning from old action learning, and prospectively learn newer more optimal task strategies when it was useful to do so. Of interest, this DLS inhibition effect looked "maladaptive" when animals gave 1-press bouts when shifted to the FR3, but looked "adaptive" when shifted back to FR1. These DLS perturbations were pan-neuronal, and we recognize that different cell types may function differently from one another. For example, we have also inhibited cholinergic interneurons and found that this inhibition causes animals to adapt more readily to task changes, for example, from an FR1 to FR3, which is opposite to the pan-neuronal inhibition effect. Moreover, a recording study recently identified DLS neuronal activity that appears to encode prior (retrospective) information for behavior rather than forthcoming (prospective) information (Cunningham et al., 2021). So, there is certainly an interesting cell-type story to investigate in this research area.

Our study also uncovered a conditional effect of DLS in controlling performance vigor. In many prior studies, including our own, DLS inhibition will reduce performance vigor. This is particularly true when there is a "reason" to reduce the vigor of behavior, such as after devaluation of the reward when it may no longer make sense to seek the goal (Crego et al., 2020). In the action strategy study, we found that DLS inhibition likewise reduced vigor when animals were shifted from FR1 to FR3. However, when animals were later moved from FR3 to FR1, the same DLS inhibition oppositely increased vigor. This was a surprising result because DLS inhibition reduced vigor rather reliably in our prior maze study (Crego et al., 2020). It might be that the DLS participates differently in lever-press and maze-running conditions. Nevertheless, we interpret this surprising result as showing that, normally, the DLS encourages vigor when task demands rise (e.g., FR1➔FR3), while encouraging it less when task demands decrease (FR3➔FR1). With DLS inhibition, vigor was improperly low at the increase in task demand and improperly high at the reduction in task demand. This bidirectional effect of DLS inhibition on vigor, despite animals doing FR1-like behaviors in both cases, might indicate that DLS inhibition is not simply making animals prefer to do a strategy that requires less effort. Likewise, when the reward is not delivered on an FR1 schedule but animals behave in an FR1 manner, it is not an energetically "easy" strategy to use. However, there may well be an interaction of effort and task strategy, underlying the effects of DLS inhibition in this particular case.

In conclusion, we offer several points of consideration: (1) habits could in principle arise out of animals favoring prior learning conditions over new changed ones, as an active online process; (2) insensitivity to outcome devaluation, a defining characteristic of habits, could arise from animals behaving prospectively by considering initial learning conditions as separate from new outcome devaluation information gained elsewhere; (3) DLS activity during behavior can encourage a prospective learning strategy for acquiring new actions routines while also preserving the old learned routines; and (4) DLS controls performance vigor but can do so conditionally to match task demands in a similar action-context-specific manner. These points provide a means for considering the DLS as an active, prospective brain system rather than a retrospective, reflexive brain system for habitual behavior. We would not go so far as to suggest that these prospective representations of actions in the DLS reach conscious control, nor that habits ought to be considered consciously regulated processes. It simply suggests a different way of thinking about DLS function, and a new way of thinking about how the defining measures that we have for habits—vigorous and seemingly inflexible behaviors—could have underpinnings in a surprisingly forward-looking biobehavioral operation. It will be worth some deep consideration about how to define habits going forward. Do we need to stop using the word "habit" to describe a behavior if it can occur through a cognitive-like brain process? Or do we expand our definition of "habit" to encompass those processes? For example, if devaluation insensitivity can be achieved in a cognitive-like process, perhaps it should not be used to define a habit. However, if that cognitive-like process occurs rapidly and automatically, perhaps it is fine to call it a habit.

3.4 A Role for Motivation in Habits and the DLS

A complementary notion, in alignment with a prospective role for the DLS, is that actions themselves can become endowed with motivational value and become an actual goal in their own right (Robinson & Berridge, 1993) (Fig. 1). Although habits are recognized as behaviors that lack a representation of the value of the specific outcome that they lead to, it is important to recognize that they are still very much influenced by motivational states and can be conditional. We also might have a ritual in which we prepare and drink coffee, but we do this at certain times when the desire for coffee hits us and not at other times. We also don't often eat, or at least eat a bit less, when sated versus when hungry. In experiments, this is seen for example in data showing that the strength of habits increases when hungry versus when full (Dickinson et al., 1995). The same motivational modulation of habits would likely occur across different internal states, including, for example, from thirst, sexual arousal, and socializing need. Cues can evoke motivational states that are thought to regulate habitual responding as well. For example, a view of habitual drug-taking behavior called "incentive habit" proposes that incentive motivational signals residing in limbic areas become abnormally coupled with DLS-based habit systems, resulting in habits being generated through cues and the motivational states they trigger (Belin et al., 2013) (see also: Dayan & Balleine, 2002).

A different sort of way in which motivation might interface with habits is if actions themselves carry incentive value. In this view, actions can be something to do for their own sake, related to, but dissociable from, how those actions produce a rewarding outcome. There is a motivational urge to do them, and doing them is a goal of the organism in its own right (Glickman & Schiff, 1967; Robinson & Berridge, 1993; Mazzoni et al., 2007). The possibility that actions can have an incentive value was recognized by Robinson and Berridge when they proposed that Pavlovian cues can carry incentive salience and so too might actions (Robinson & Berridge, 1993). This notion holds that Pavlovian cues, by being paired with rewards, can become incentive targets, triggering motivation and attracting attention and behavioral pursuit. Sign-tracking, as noted above, is a behavior that can be explained by cues having an incentive value to them. In the brain, the incentive salience of cues has been linked to networks including the nucleus accumbens, ventral pallidum, and their dopaminergic innervation. As the model goes, animals learn cue-reward relationships and dopamine release endows those cues with attractiveness and value. Perhaps actions can have an incentive value too. If actions can be incentives, it may have some explanatory power for understanding the roles of the DLS and its dopaminergic input. What happens when you lose the DLS? One thing we have covered is that you lose performance vigor. Subjects will be slower to initiate behaviors, be slower or more deliberate in performing them, and be more scattered in performing established routines. One might use the word "motivation" instead of "vigor" to describe such effects. Another aspect of behavior that is lost with DLS dysfunction is measures of motivations themselves, such as sign-tracking and PIT. If DLS is contributing a motivational value to behaviors, it might also help explain how in many cases the behaviors themselves seem to be complex natural repertoires of behavior, rooted in a sort of "schema" of what to do, rather than specific movements that could become associated with stimuli in an SR sense, such as in cue-guided behaviors ("beacon"-based performance), sign-tracking, and grooming. Dopamine within limbic circuitry is regarded as being important for the incentive valuation of cues, and there are hints that dopamine in the sensorimotor system could function similarly to endow actions with incentive value. These hints come from findings that nigrostriatal dopamine regulates measures of behavioral vigor (Mazzoni et al., 2007; Barter et al., 2015; Panigrahi et al., 2015; Dodson et al., 2016; da Silva et al., 2018), and from findings that real-time nigrostriatal dopamine activity can itself increase action repetition in the absence of any explicit reward (Markowitz et al., 2023). There is also evidence from Parkinson's Disease. A similar idea to the notion that actions have motivational value is found in Mazzoni, Hristova, and Krakauer (Mazzoni et al., 2007). In studying the consequences of striatonigral dopamine depletion in Parkinson's patients, they suggested that there is low intrinsic motivation to move. They state, "*We introduce the idea that the motor system has its own motivation circuit, which operates analogously to but separately from explicit motivation. We suggest that this "motor motivation" works implicitly (*i.e.*, outside of awareness) and governs automatic and spontaneous behavior . . .* " (Mazzoni et al., 2007, pg. 7115). Their framework of a "motor motivation" seems close to the idea that dopamine relates to cost/benefit analyses of moving, but reads as a more goal-focused account of valuing an action rather than the incentive salience notion above in which the action itself is intrinsically valued. Prior modeling efforts of actions and habits have proposed something similar, in which behaviors can become a goal in a cognitive-like sense, in which animals purposefully engage with them as a point of focus (Dezfouli & Balleine, 2012).

There may be explanatory power for motivational processes in understanding why behaviors can persist despite reward devaluation or task changes as well. If actions acquire motivational value and become goals themselves, then their utility becomes dissociated from the value of the ultimate reward goal just like the value of cues can become disconnected from the value of the rewards they are associated with. Thus, it is worth considering that one reason animals may continue to perform behaviors even after task changes like outcome devaluation is because the actions themselves still carry some value to do. In other words, the focus of the animal is on the action itself rather than on the reward goal, and as a result, changes to the reward goal do not immediately affect the value of the action. Ultimately, the reward goal will be the focus, but for now, the brain desires that next action step.

This reminds us of a key feature of the incentive salience concept of Pavlovian conditioned responding, which is that it can become decoupled from the outcome value (unconditioned stimulus) itself (Robinson & Berridge, 1993). Normally, our wants track our likes. For example, we often respond to a cue for food appropriately to the extent that we will ultimately like that food. But there are also examples in which our response to the cue—the wanting that it evokes—is irrationally higher than how much we will like the reward that the cues predict. This heightened cue value can play a role in peoples' urges to eat past satiety or to take drugs despite the enjoyment of the drug being low from tolerance. Robinson and Berridge write, *"The neural system that is rendered hypersensitive ('sensitized') to activating stimuli is hypothesized to mediate a specific psychological function involved in the process of incentive motivation: namely the attribution of incentive salience to the perception and mental representation of stimuli and actions. This makes stimuli and their representations highly salient, attractive, and "wanted""* (Robinson & Berridge, 1993, Pg 249). In a more recent publication, Berridge adds, "*my colleagues and I have often wondered whether incentive salience can be attributed onto an action by mesostriatal systems, as "action salience," which would give motivational attraction and urgency to performing it*" (Berridge, 2021). A notable feature of the incentive valuation of reward-related stimuli is that it can, in some studies, grow in strength to tag the earliest relevant information as particularly meaningful. For example, one can set up a task in which a cue follows a second cue which then leads to reward. Animals will sign-track to both cues, showing that the cues have acquired motivational value. With time, sign-tracking grows to be greater to the first cue compared to the second, suggesting its value is highest (Smedley & Smith, 2018a, b). In brain motivation systems, areas like the ventral pallidum can represent the value of serially occurring cues with neuronal responses to both cues but with greater activity to the initial cue (although changes in the motivational state seem to preferentially affect neuronal activity at the second cue) (Tindell et al., 2004; Smith et al., 2011). In the domain of action learning, a recent study (van Elzelingen et al., 2022) trained animals to do one action (a seeking action), and then a second (a taking action) to get reward. Animals received periodic assays of action preference in which they could choose which of the two actions to take if they were simultaneously available. As animals learned the task and then were overtrained, they developed a strong preference for performing the first action (the seeking action), which was argued to reflect the development of a habit in the task. Although this finding could be interpreted as animals learning with time to understand and favor the *initiation* of a seeking-taking routine, it might also be interpreted as showing an incentive valuation that is highest to the seeking action. This preference of the seeking action in this study corresponded in the brain to DLS dopamine signaling, but actually more so to DMS dopamine signaling, which is a finding worthy of further research. We have noted above that in the DLS, aspects of a habit such as vigor can be dictated by the strength of DLS activity at the onset of complex behaviors—the onset being similarly the first task event. In turn, there is evidence that when ingrained behaviors are halted, such as after enough experience with a reward being devalued with CTA, after satiety, or during extinction, the initially learned behaviors are the most stubborn to get rid of. In a maze study example, when all outcomes are devalued, animals quit running the task in a marching-back fashion, first ceasing to approach the goal, then ceasing to turn to it, and then finally after a lot of trials ceasing to start running (Smith & Graybiel, 2013). During this, DLS activity chunking patterns deteriorated, much in the way they deteriorate when rewards are removed (Barnes et al., 2005). Earlier studies have shown this sort of "breakdown of operant chains from consummatory response backwards" (Morgan, 1974) for various behaviors during extinction. Taken together, these data indicate the possibility that some task scenarios endow the earliest task-relevant information, be it a cue or action, with importance. This importance could plausibly reflect a motivational value to engage with that cue or behavior.

Of course, much more work needs to be done to measure whether animals are performing actions as a means to reward only, or if there may be a motivational value for the actions that results in them being an incentive goal themselves (Robinson & Berridge, 1993). One approach might be to devise situations in which performing an action can be unequivocally a goal to the animal, a goal they might even work for. For example, some work (Carder & Berkowitz, 1970; Taylor, 1972) has shown that, if given a choice, animals can actually prefer to work for food rather than to receive free food, as though the work itself was valued. However, to the extent one wants to consider this "incentive action" idea as a real potential phenomenon, one may find examples all around them. It is not uncommon for routine to be described as something that is desired, rituals that seem pointless to be described as comforting and meaningful, warm-up patterns to be described as critical elements of athletic preparation, behavioral exertion like weight lifting to be described as enjoyed, and working hard to be described in adulating terms. These anecdotes suggest an intuitive appeal for the idea that behaviors done repeatedly can acquire a positive motivational value. In research settings, there are many examples of animals performing actions despite them lacking a clear goal. For instance, a review by Morgan (Morgan, 1974) highlights ways in which instinctual or learned actions in animals can persist beyond satiety and thus beyond valuing the outcome that those actions lead to (e.g., food). Once sated, animals' food-seeking actions often do not cease immediately but gradually decrease over time/trials. Diverse behaviors such as lever pressing, digging, and feline predation have all been seen to continue despite satiation and well past the point of animals consuming much of the reward when becomes available. The insensitivity of behavior to satiety is a common measure of habits, and so one might want to define these kinds of satiety-resistant behaviors as SR-based habits. Although it is fine to draw this conclusion, it might be more compelling to view these behaviors as things animals are engaged in and are motivated to do in their own right because, as the Morgan article discussed, they have their own history of utility. It is difficult to look at pet cats chasing laser pointers, or dogs chewing into squeaky toys, and describe them as SR robots; instead, we see the animals' behavior as something they have a desire to do and seem to enjoy, sometimes irrespective of any reward like food that may come from it. In short, an account of habits that considers actions as having an incentive value themselves may be worth consideration. If actions can be a goal, then it would make sense that they can be performed for no clear (i.e., reward-related) reason to experimenters. In the brain, it is worth studying a potential role for the DLS in this process.

4 Conclusion

SR accounts have helped science make terrific headway in understanding how behaviors can develop into habits, and how brain systems for habits function. However, on closer examination, SR accounts do not seem to have full explanatory power for understanding some aspects of behavior we consider to be hallmarks of habits, or for understanding the function of brain areas like the DLS. To fill this explanatory gap, we have explored additional frameworks (Fig. 1). These include considering habits as an active prospective process of performing actions, as a process of contextualizing old behaviors separate from new information, and/or as actions themselves becoming motivational targets. To complement the work being done on brain systems for goal-directed versus habitual policies of behavior, as well as work on behavioral markers of habits, it could be helpful to examine different possible conceptual frameworks for habits in experimental settings. In this vein, we propose it can be worth taking a brain structure that encourages habit-like persistence in behavior and asking *how* it does that.

Acknowledgments This work was supported by an NSF research grant to KSS (IOS 1557987), an NIH grant to KSS (R01DA04419), and an NIH grant to KAA (K00MH130162). We thank Youna Vandaele for insightful comments on an earlier version of this chapter.

References

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B, 34*, 77–98.

Ahrens, A. M., Meyer, P. J., Ferguson, L. M., et al. (2016). Neural activity in the ventral pallidum encodes variation in the incentive value of a reward cue. *The Journal of Neuroscience, 36*, 7957–7970.

Aldridge, J. W., Berridge, K. C., & Rosen, A. R. (2004). Basal ganglia neural mechanisms of natural movement sequences. *Canadian Journal of Physiology and Pharmacology, 82*, 732–739.

Amaya, K. A., Stott, J. J., & Smith, K. S. (2020). Sign-tracking behavior is sensitive to outcome devaluation in a devaluation context-dependent manner: Implications for analyzing habitual behavior. *Learning & Memory, 27*, 136–149.

Asem, J. S. A., & Holland, P. C. (2015). Dorsolateral striatum implicated in the acquisition, but not expression, of immediate response learning in rodent submerged T-maze. *Neurobiology of Learning and Memory, 123*, 205–216.

Bailey, K. R., & Mair, R. G. (2006). The role of striatum in initiation and execution of learned action sequences in rats. *The Journal of Neuroscience, 26*, 1016–1025.

Balleine, B. W., & Dezfouli, A. (2019). Hierarchical action control: Adaptive collaboration between actions and habits. *Frontiers in Psychology, 10*, 2735.

Balleine, B. W., Liljeholm, M., & Ostlund, S. B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behavioural Brain Research, 199*, 43–52.

Barnes, T. D., Kubota, Y., Hu, D., et al. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature, 437*, 1158–1161.

Barter, J. W., Li, S., Lu, D., et al. (2015). Beyond reward prediction errors: The role of dopamine in movement kinematics. *Frontiers in Integrative Neuroscience, 9*, 39.

Belin, D., Belin-Rauscent, A., Murray, J. E., & Everitt, B. J. (2013). Addiction: Failure of control over maladaptive incentive habits. *Current Opinion in Neurobiology, 23*, 564–572.

Bergstrom, H. C., Lieberman, A. G., Graybeal, C., et al. (2020). Dorsolateral striatum engagement during reversal learning. *Learning & Memory, 27*, 418–422.

Berke, J. D., Breck, J. T., & Eichenbaum, H. (2009). Striatal versus hippocampal representations during win-stay maze performance. *Journal of Neurophysiology, 101*, 1575–1587.

Berridge, K. C. (2021). Comment on Vandaele and Ahmed: Rethinking habits in addiction. *Neuropsychopharmacology, 46*(4), 687–688.

Bossert, J. M., Wihbey, K. A., Pickens, C. L., Nair, S. G., & Shaham, Y. (2009). Role of dopamine D_1-family receptors in dorsolateral striatum in context-induced reinstatement of heroin seeking in rats. *Psychopharm, 206*(1), 51–60.

Bouton, M. E., Allan, S. M., Tavakkoli, A., et al. (2021). Effect of context on the instrumental reinforcer devaluation effect produced by taste-aversion learning. *The Journal of Experimental Psychology: Animal Learning and Cognition, 47*, 476–489.

Carder, B., & Berkowitz, K. (1970). Rats' preference for earned in comparison with free food. *Science, 167*, 1273–1274.

Carelli, R. M., & West, M. O. (1991). Representation of the body by single neurons in the dorsolateral striatum of the awake, unrestrained rat. *The Journal of Comparative Neurology, 309*, 231–249.

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

Chang, S. E., & Smith, K. S. (2016). An omission procedure reorganizes the microstructure of sign-tracking while preserving incentive salience. *Learning & Memory, 23*, 151–155.

Colwill, R. M., & Rescorla, R. A. (1985). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology. Animal Behavior Processes, 11*, 520–536.

Corbit, L. H., & Janak, P. H. (2007). Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *The Journal of Neuroscience, 27*, 13977–13981.

Crego, A. C. G., Štoček, F., Marchuk, A. G., et al. (2020). Complementary control over habits and behavioral vigor by phasic activity in the dorsolateral striatum. *The Journal of Neuroscience, 40*, 2139–2153.

Crego, A. C. G., Amaya, K. A., Palmer, J. A., & Smith, K. S. (2023). Task history dictates how the dorsolateral striatum controls action strategy and vigor. *bioRxivorg*. https://doi.org/10.1101/2023.01.11.523640

Cromwell, H. C., & Berridge, K. C. (1996). Implementation of action sequences by a neostriatal site: A lesion mapping study of grooming syntax. *The Journal of Neuroscience, 16*, 3444–3458.

Crutcher, M. D., & DeLong, M. R. (1984). Single cell studies of the primate putamen. II. Relations to direction of movement and pattern of muscular activity. *Experimental Brain Research, 53*, 244–258.

Cunningham, P. J., Regier, P. S., & Redish, A. D. (2021). Dorsolateral striatal task-initiation bursts represent past experiences more than future action plans. *The Journal of Neuroscience, 41*, 8051–8064.

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.

Dayan, P., & Balleine, B. W. (2002). Reward, motivation, and reinforcement learning. *Neuron, 36*(2), 285–298.

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience, 14*(2), 473–492.

De Houwer, J. (2019). On how definitions of habits can complicate habit research. *Frontiers in Psychology, 10*, 2642.

Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *The European Journal of Neuroscience, 35*, 1036–1051.

Dickinson, A., & Weiskrantz, L. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 308*, 67–78.

Dickinson, A., Balleine, B., Watt, A., et al. (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior, 23*, 197–206.

DiFeliceantonio, A. G., & Berridge, K. C. (2016). Dorsolateral neostriatum contribution to incentive salience: Opioid or dopamine stimulation makes one reward cue more motivationally attractive than another. *The European Journal of Neuroscience, 43*, 1203–1218.

Dodson, P. D., Dreyer, J. K., Jennings, K. A., et al. (2016). Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. *Proceedings of the National Academy of Sciences of the United States of America, 113*, E2180–E2188.

Dudman, J. T., & Krakauer, J. W. (2016). The basal ganglia: From motor commands to the control of vigor. *Current Opinion in Neurobiology, 37*, 158–166.

Fentress, J. C. (1973). Development of grooming in mice with amputated forelimbs. *Science, 179*, 704–705.

Fraser, K. M., & Janak, P. H. (2017). Long-lasting contribution of dopamine in the nucleus accumbens core, but not dorsal lateral striatum, to sign-tracking. *The European Journal of Neuroscience, 46*, 2047–2055.

Garr, E., & Delamater, A. R. (2019). Exploring the relationship between actions, habits, and automaticity in an action sequence task. *Learning and Memory, 26*, 128–132.

Garrett, N., Allan, S., & Daw, N. D. (2023). Model based control can give rise to devaluation insensitive choice. *Addiction Neuroscience, 6*, 100070.

Glickman, S. E., & Schiff, B. B. (1967). A biological theory of reinforcement. *Psychological Review, 74*, 81–109.

Goodman, J., Ressler, R. L., & Packard, M. G. (2017). Enhancing and impairing extinction of habit memory through modulation of NMDA receptors in the dorsolateral striatum. *Neuroscience, 352*, 216–225.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience, 31*, 359–387.

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.

Haith, A. M., Reppert, T. R., & Shadmehr, R. (2012). Evidence for hyperbolic temporal discounting of reward in control of movements. *The Journal of Neuroscience, 32*, 11727–11736.

Holland, P. C. (2008). Cognitive versus stimulus-response theories of learning. *Learning & Behavior, 36*, 227–241.

Jackson, S. A. W., Horst, N. K., Axelsson, S. F. A., et al. (2019). Selective role of the putamen in serial reversal learning in the marmoset. *Cerebral Cortex, 29*, 447–460.

Jin, X., & Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature, 466*(7305), 457–462.

Jog, M. S., Kubota, Y., Connolly, C. I., et al. (1999). Building neural representations of habits. *Science, 286*, 1745–1749.

Jonkman, S., Pelloux, Y., & Everitt, B. J. (2012). Differential roles of the dorsolateral and midlateral striatum in punished cocaine seeking. *The Journal of Neuroscience, 32*, 4645–4650.

Kosaki, Y., Poulter, S. L., Austen, J. M., & McGregor, A. (2015). Dorsolateral striatal lesions impair navigation based on landmark-goal vectors but facilitate spatial learning based on a "cognitive map". *Learning & Memory, 22*, 179–191.

Kubota, Y., Liu, J., Hu, D., et al. (2009). Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *Journal of Neurophysiology, 102*, 2142–2160.

Künzle, H. (1977). Projections from the primary somatosensory cortex to basal ganglia and thalamus in the monkey. *Experimental Brain Research, 30*, 481–492.

Lerner, T. N. (2020). Interfacing behavioral and neural circuit models for habit formation. *Journal of Neuroscience Research, 98*(6), 1031–1045.

Locurto, C., Terrace, H. S., & Gibbon, J. (1976). Autoshaping, random control, and omission training in the rat. *Journal of the Experimental Analysis of Behavior, 26*, 451–462.

Malvaez, M., & Wassum, K. M. (2018). Regulation of habit formation in the dorsal striatum. *Current Opinion in Behavioral Sciences, 20*, 67–74.

Markowitz, J. E., Gillis, W. F., Jay, M., et al. (2023). Spontaneous behaviour is structured by reinforcement without explicit reward. *Nature, 614*, 108–117.

Mazzoni, P., Hristova, A., & Krakauer, J. W. (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *The Journal of Neuroscience, 27*, 7105–7116.

Morgan, M. J. (1974). Resistance to satiation. *Animal Behaviour, 22*, 449–466.

Naeem, M., & White, N. M. (2016). Parallel learning in an autoshaping paradigm. *Behavioral Neuroscience, 130*, 376–392.

Packard, M. G. (2009). Exhumed from thought: Basal ganglia and response learning in the plus-maze. *Behavioural Brain Research, 199*, 24–31.

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

Panigrahi, B., Martin, K. A., Li, Y., et al. (2015). Dopamine is required for the neural representation and control of movement vigor. *Cell, 162*, 1418–1430.

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research. Brain Research Reviews, 18*, 247–291.

Sage, J. R., & Knowlton, B. J. (2000). Effects of US devaluation on win-stay and win-shift radial maze performance in rats. *Behavioral Neuroscience, 114*, 295–306.

Schmitzer-Torbert, N., & Redish, A. D. (2004). Neuronal activity in the rodent dorsal striatum in sequential navigation: Separation of spatial and reward responses on the multiple T task. *Journal of Neurophysiology, 91*(5), 2259–2272.

Schreiner, D. C., Renteria, R., & Gremel, C. M. (2020). Fractionating the all-or-nothing definition of goal-directed and habitual decision-making. *Journal of Neuroscience Research, 98*, 998–1006.

Seger, C. A. (2018). Corticostriatal foundations of habits. *Current Opinion in Behavioral Sciences, 20*, 153–160.

Seiler, J. L., Cosme, C. V., Sherathiya, V. N., et al. (2022). Dopamine signaling in the dorsomedial striatum promotes compulsive behavior. *Current Biology, 32*, 1175–1188.e5.

Selemon, L. D., & Goldman-Rakic, P. S. (1985). Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *The Journal of Neuroscience, 5*, 776–794.

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? *The Journal of Neuroscience, 38*, 60–73.

Smedley, E. B., & Smith, K. S. (2018a). Evidence for a shared representation of sequential cues that engage sign-tracking. *Behavioural Processes, 157*, 489–494.

Smedley, E. B., & Smith, K. S. (2018b). Evidence of structure and persistence in motivational attraction to serial Pavlovian cues. *Learning & Memory, 25*, 78–89.

Smith, K. S., & Graybiel, A. M. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron, 79*, 608.

Smith, K. S., & Graybiel, A. M. (2014). Investigating habits: Strategies, technologies and models. *Frontiers in Behavioral Neuroscience, 8*, 39.

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.

Smith, K. S., Berridge, K. C., & Aldridge, J. W. (2011). Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. *Proceedings of the National Academy of Sciences of the United States of America, 108*, E255–E264.

Smith, K. S., Virkud, A., Deisseroth, K., & Graybiel, A. M. (2012). Reversible on-line 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*(46), 18932–18937.

Smith, A. C. W., Jonkman, S., Difeliceantonio, A. G., et al. (2021). Opposing roles for striatonigral and striatopallidal neurons in dorsolateral striatum in consolidating new instrumental actions. *Nature Communications, 12*, 5121.

Stalnaker, T. A., Calhoon, G. G., Ogawa, M., et al. (2010). Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. *Frontiers in Integrative Neuroscience, 4*, 12.

Tang, C., Pawlak, A. P., Prokopenko, V., & West, M. O. (2007). Changes in activity of the striatum during formation of a motor habit. *The European Journal of Neuroscience, 25*, 1212–1227.

Tang, C. C., Root, D. H., Duke, D. C., et al. (2009). Decreased firing of striatal neurons related to licking during acquisition and overtraining of a licking task. *The Journal of Neuroscience, 29*, 13952–13961.

Taylor, G. T. (1972). A limitation of the contrafreeloading phenomenon. *Psychonomic Science, 29*, 173–174.

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.

Tindell, A. J., Berridge, K. C., & Aldridge, J. W. (2004). Ventral pallidal representation of pavlovian cues and reward: Population and rate codes. *The Journal of Neuroscience, 24*, 1058–1069.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. Century/Random House UK.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208.

Turner, K. M., & Parkes, S. L. (2020). Prefrontal regulation of behavioural control: Evidence from learning theory and translational approaches in rodents. *Neuroscience and Biobehavioral Reviews, 118*, 27–41.

van Elzelingen, W., Warnaar, P., Matos, J., et al. (2022). Striatal dopamine signals are region specific and temporally stable across action-sequence habit formation. *Current Biology, 32*, 1163–1174.e6.

Vandaele, Y., & Ahmed, S. H. (2021). Habit, choice, and addiction. *Neuropsychopharmacology, 46*(4), 689–698.

Vandaele, Y., & Janak, P. H. (2023). Lack of action monitoring as a prerequisite for habitual and chunked behavior: Behavioral and neural correlates. *iScience, 26*, 105818.

Vandaele, Y., Mahajan, N. R., Ottenheimer, D. J., et al. (2019). Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training. *eLife, 8*. https://doi.org/10.7554/eLife.49536

Watson, P., O'Callaghan, C., Perkes, I., et al. (2022). Making habits measurable beyond what they are not: A focus on associative dual-process models. *Neuroscience and Biobehavioral Reviews, 142*, 104869.

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

Yin, H. H., Mulcare, S. P., Hilário, M. R. F., et al. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience, 12*, 333–341.