A critical review of habit learning and the basal ganglia

Carol A. Seger and Brian J. Spiering*

Department of Psychology, Colorado State University, Fort Collins, CO, USA

The current paper briefly outlines the historical development of the concept of habit learning and discusses its relationship to the basal ganglia. Habit learning has been studied in many different fields of neuroscience using different species, tasks, and methodologies, and as a result it has taken on a wide range of definitions from these various perspectives. We identify five common, but not universal, definitional features of habit learning: that it is inflexible, slow or incremental, unconscious, automatic, and insensitive to reinforcer devaluation. We critically evaluate for each of these how it has been defined, its utility for research in both humans and non-human animals, and the evidence that it serves as an accurate description of basal ganglia function. In conclusion, we propose a multi-faceted approach to habit learning and its relationship to the basal ganglia, emphasizing the need for formal definitions that will provide directions for future research.

Keywords: basal ganglia, habit learning, automaticity, reward

INTRODUCTION

The concept of habit learning has developed through the fruitful interaction of researchers in several intellectual domains, including animal learning, cognitive psychology, cognitive neuropsychology, and behavioral neuroscience. As a result, habit learning has taken on a variety of proposed definitions. In this paper, we will first describe the historical evolution of habit learning as a concept. We will then briefly describe the anatomical and functional roles of the basal ganglia that may underlie learning in general and habit learning in particular. Finally, we will revisit the defining features of habit learning and assess how well they characterize learning in the basal ganglia.

HISTORICAL EVOLUTION OF THE HABIT LEARNING CONCEPT

The term habit was used, but not explicitly defined, by William James in the seminal Principles of Psychology (James, 1890). It was used on occasion by early researchers studying animal learning, in particular Hull (1934a,b) and Lashley (1930, 1950). "Habit" roughly corresponded to the resulting motor behavior (e.g., Lashley referred to the "maze running habit"), and habit learning to acquisition of these behaviors in an instrumental learning context.

HIPPOCAMPAL RESEARCH: EARLY DEFINITIONS OF HABIT LEARNING

The earliest use of "habit learning" to refer to a specific form of learning came from researchers studying the effects of hippocampal damage in human and non-human animals. By the late 1960s it was clear that hippocampal damage affected learning on many, but not all, tasks. Hirsh (1974) first used the term "habit learning" to describe a particular type of memory or learning system. He defined the habit system as that "responsible for the learning of which hippocampally ablated animals are capable" (1974, p 421). Thus, from the beginning habit learning was defined negatively, in terms of what it was not (i.e., hippocampally based), rather than what it was. To Hirsch, the primary feature of hippocampal-based learning was contextual encoding (e.g., of the particular spatial and

temporal context at encoding) and retrieval of information that was contextually sensitive. He argued in contrast that habit learning was similar to the stimulus–response (S–R) learning processes proposed by earlier learning researchers, and that these S–R associations were specifically insensitive to context.

Miskin et al. (1984) extended Hirsch's concept of habit learning. Following Hirsch, they identified features of habit learning as the opposite those of hippocampally-based learning. One set of features was "rapid" versus "slow" learning. Rapid learning was defined as one-trial learning, which required the hippocampus, whereas slow learning required repeated trials and was preserved in amnesia. They immediately related the "rapid"–"slow" distinction to the distinction posed by Hirsh, which they referred to as flexible (contextually sensitive in Hirsh) versus inflexible learning. They proposed that there is "a trade-off between short-term flexibility afforded by the memory system and long-term reliability afforded by the habit system" (p. 73). Finally, they argued that habits were a relatively primitive form of learning that should therefore appear earlier in ontogeny as well as phylogeny, which they supported with developmental evidence from their lab.

Mishkin and colleagues were also the first to propose a crucial role for the basal ganglia in habit learning. The basis for their argument, which they termed "admittedly speculative" was the early development of the basal ganglia both in phylogeny and ontogeny, and the presence of widespread anatomical projections to the striatum from cortex that "provide a mechanism through which cortically processed sensory inputs could become associated with motor outputs generated in the pallidum and so yield the stimulus–response bonds that constitute habits" (p. 74).

COGNITIVE PSYCHOLOGY: HABIT AS IMPLICIT AND AUTOMATIC

The field of cognitive psychology did not use the term habit learning, but from the late 1960s through the 1980s several concepts were developed in this field that later were incorporated into theories of habit learning. These distinctions included unconscious,

or implicit, learning and memory (in contrast with conscious or explicit learning and memory), and automatic processing (in contrast with controlled processing). Both of these distinctions fall broadly within "dual process" theories of cognition that see one type of cognitive process as relatively unconscious, automatic, evolutionarily early, and similar across individuals, in contrast with a second type of cognitive process that is conscious, controlled, evolutionarily more recent, and subject to significant individual differences (see Evans, 2008 for a review).

Reber (1967) coined the term "implicit learning"; the concept was extended to "implicit memory" by Graf and Schacter (1985). The focus in both areas of research was on consciousness: identifying what could or could not be learned and/or retrieved without awareness. Implicit memory was defined as "when previous experiences facilitate performance on a task that does not require conscious or intentional recollection of those experiences" (Schacter, 1987, p. 501). Implicit memory tasks typically used priming paradigms in which improvement in accuracy and/or processing time was observed for repeated stimuli; priming was later divided into perceptual (repeated visual stimulus processing) and conceptual (repeated semantic processing) forms (Keane et al., 1991).

Seger (1994, p. 164) outlined three guidelines for implicit learning: (1) "the knowledge gained in implicit learning is not fully accessible to consciousness, in that subjects cannot provide a full … verbal account of what they have learned," (2) "information [learned] … is more complex than a single simple association or frequency count," and (3) "implicit learning does not involve processes of conscious hypothesis testing but is an incidental consequence of the type and amount of cognitive processing performed on the stimuli." Implicit learning was studied using several different tasks, most often the serial reaction time task (which measures improvement in reaction time when responding to stimuli when presented in a repeating sequence in comparison with stimuli presented in random order), and the artificial grammar task (which measures the ability of subjects to discriminate letter strings that follow a complex sequential pattern determined by a finite state automaton, or artificial grammar, from those that violate the pattern).

Another influential concept from cognitive psychology was that of automaticity, originally developed by Shiffrin and Schneider (1977) to account for different forms of attentional scanning. The concepts of automatic and controlled processing were widely adopted across various domains within cognitive psychology. Shiffrin and Schneider (1977) gave multiple criteria for considering a process to be automatic, including (1) automatic processes are not constrained by short-term memory capacity limitations and do not require attention; (2) automatic processes are generally performed too quickly to be consciously accessible and once initiated are completed regardless of subjects' intentions; (3) automatic processes require significant training, undergoing a gradual shift from controlled to automatic through the course of practice; and (4) automatic processes, once acquired, are difficult to modify. Criterion 1 led to an operational definition of automaticity as primary task performance not negatively affected by a parallel, short-term memory demanding task.

Through development of the process dissociation procedure, Jacoby (1991) related automaticity to implicit learning and memory. He argued that participants should be able to exert strategic control over conscious knowledge, and theorized that they should be able to control the behavioral expression of this knowledge in accordance with task instructions. Conversely, he argued that participants should be unable exert strategic control over unconscious knowledge and theorized that they might have difficulty controlling the behavioral expression of this knowledge. The critical feature of the process dissociation procedure is that participants are asked to demonstrate knowledge via both "inclusion" and "exclusion" instructions. Inclusion instructions demand that participants produce behavior in accordance with a learned structure, while exclusion instructions demand that participants produce discordant behavior. This approach led to a different operational definition of automaticity. Automatic processing is measured by calculating the intrusion of the previously learned material into the exclusion condition (false positives); controlled processing is defined as the difference between performance in the inclusion condition and the automatic processing measure.

COGNITIVE NEUROPSYCHOLOGY: HABIT AS A TYPE OF "NON-DECLARATIVE" MEMORY

Larry Squire and colleagues integrated the approaches taken by researchers examining hippocampal lesions in non-human animals, researchers in cognitive psychology, and researchers studying human patients with amnesia. Their theory developed across time. Cohen and Squire (1980) initially defined procedural learning as "operations governed by rules or procedures" in contrast to hippocampally based learning, which they characterized as "operations that depend on specific, declarative, data-based material." The term "procedural" was adapted from artificial intelligence research (Winograd, 1972; Anderson, 1982). Anderson's (1982) view was that all cognitive knowledge started by being represented declaratively, as individual "propositions," and procedural knowledge was formed by the compilation of groups of propositions into procedures. Procedural knowledge accounted well for the tasks known to be preserved in amnesia at that time, including pursuit rotor (Corkin, 1968), mirror drawing (Milner, 1962), and mirror reading (Cohen and Squire, 1980).

During the 1980s and early 1990s amnesic subjects were shown to have intact learning across a large number of novel tasks, primarily drawn from the implicit memory and learning literatures. These included perceptual priming (Graf and Schacter, 1985), the serial reaction time task (Nissen and Bullemer, 1987), artificial grammar learning (Knowlton et al., 1992), category learning using the Posner dot pattern task (Knowlton and Squire, 1993), and some aspects of learning on the Tower of Hanoi task (Cohen, 1984). It soon became clear that the term "procedural" was insufficient to characterize all the different types of non-hippocampal learning and memory. Squire and Zola-Morgan (1988) created the term "non-declarative" and defined it as "a heterogeneous collection of abilities: motor skills, perceptual skills, and cognitive skills (these abilities and perhaps others are examples of procedural memory); as well as simple classical conditioning, adaptation level effects, priming, and other instances where experience alters performance independently of providing a basis for the conscious recollection of past events" (p. 171). Non-declarative memory thus incorporated the cognitive psychology distinction between implicit and explicit memory with the result that hippocampal-based declarative

learning was now identified as memory that was accessible to consciousness, and the heterogeneous non-declarative memory systems as unconscious.

Squire and Zola-Morgan (1988, 1991) developed what was to become an often reprinted figure illustrating the types and subtypes of declarative and non-declarative memory (the 1991 version is shown in **Figure 1**). In Squire and Zola-Morgan (1988) the term "habit" isn't used; instead, several different "skills" are described including motor skills (pursuit motor, Corkin, 1968; serial reaction time, Nissen and Bullemer, 1987; mirror drawing Milner, 1962), perceptual skills (mirror reading; Cohen and Squire, 1980), and cognitive skills (Tower of Hanoi; Cohen, 1984. Hebb digits task: Brooks and Baddeley, 1976). By 1991, Squire and colleagues referred to this type of non-declarative memory as "skills and habits," as shown in **Figure 1**. They noted the basal ganglia as one potential neural system involved in habits and skills, along with the cerebellum.

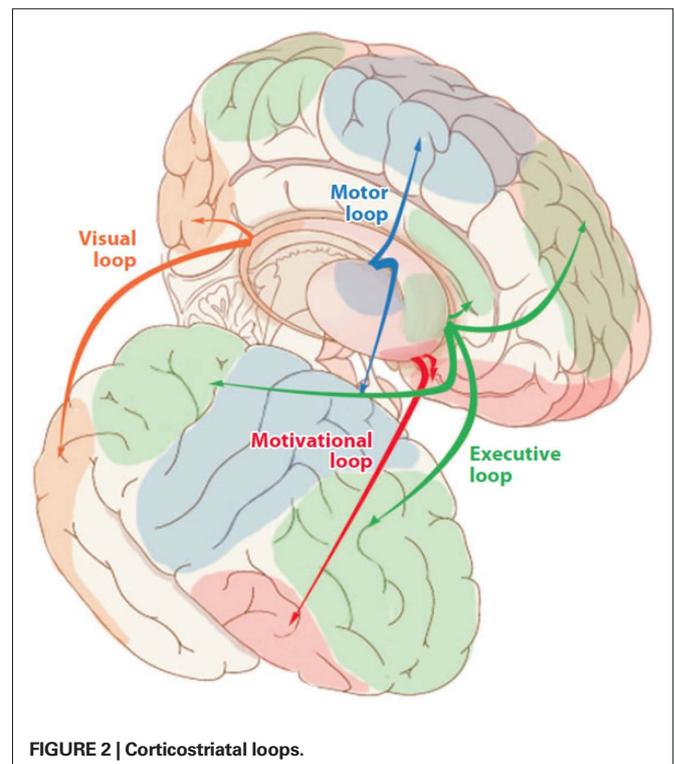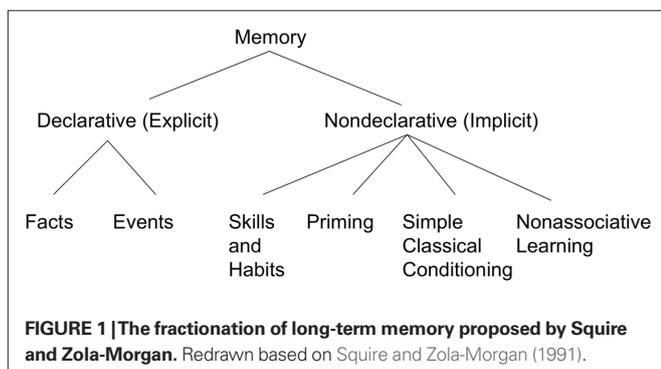### ANIMAL LEARNING: HABIT LEARNING AS ONE FORM OF INSTRUMENTAL CONDITIONING

In the 1980s, Dickinson (1985) proposed separate "goal-directed behavior" and "habit" instrumental learning systems, based on whether execution of the learned behavior is sensitive to the value of the reward or not, respectively. One typical manipulation is to devalue the reinforcer by satiating the animal before testing; the value of a food reward is greater when the animal is hungry than when it has recently fed. An animal will perform a habitual act to obtain food even when it has eaten to satiation. He contrasted habit with goal-directed behavior, which is sensitive to the motivational state of the animal. Subsequent neuroscience studies (Yin and Knowlton, 2006; Packard, 2009) found that the distinction between goal-directed and habitual learning corresponded with reliance on different parts of the basal ganglia: the dorsomedial rodent striatum (homologous to the primate anterior caudate nucleus), and the dorsolateral striatum (homologous to the primate posterior putamen) respectively.

Graybiel (2008) recently offered a broad definition of habit learning. "First, habits (mannerisms, customs, rituals) are largely learned; in current terminology, they are acquired via experience-dependent plasticity. Second, habitual behaviors occur repeatedly over the course of days or years, and they can become remarkably fixed. Third, fully acquired habits are performed almost automatically, virtually non-consciously, allowing attention to be focused elsewhere. Fourth, habits tend to involve an ordered, structured action sequence that is prone to being elicited by a particular context or stimulus. And finally, habits can comprise cognitive expressions

of routine (habits of thought) as well as motor expressions of routine" (Graybiel, 2008, p. 361). Like Squire's approach to non-declarative memory, Graybiel's definition brings together several features from previous work, including that habits are relatively automatic, and unconscious, and that habits can be inflexible and rigid (particularly well learned habits). Graybiel emphasizes two additional features of habit: first, that motor habits are sequential behaviors with complex structure, going beyond a simple concept of a "response," and second, that habits can extend beyond motor behaviors to include cognitive processes.

THE BASAL GANGLIA AND LEARNING

The basal ganglia are a group of subcortical nuclei, including the striatum, globus pallidus, substantia nigra, and subthalamic nucleus in humans. The basal ganglia interact with cerebral cortex via corticostriatal loops, in which information projects from cortex to the striatum, to the basal ganglia output nuclei, to the thalamus, and from there back to cortex (Alexander et al., 1986; Seger, 2008). The functions of the basal ganglia are supported by three pathways from the striatum to the thalamus, termed the "direct," "indirect," and "hyperdirect" pathways (Frank, 2005; Cohen and Frank, 2009). Broadly, the three pathways together implement a balance between regulating tonic inhibition in cortex as well as selective activation or gating of particular representations. The representations that the basal ganglia act upon is determined by the region of cortex within each corticostriatal loop. Although projections are continuous and there are no firm dividing lines between loops, it is useful for practical purposes to identify functionally different loops. Our approach includes four distinct loops (Seger and Cincotta, 2005; Seger, 2008) and is illustrated in **Figure 2**. They are the motor loop,

FIGURE 1 | The fractionation of long-term memory proposed by Squire and Zola-Morgan. Redrawn based on Squire and Zola-Morgan (1991).

FIGURE 2 | Corticostriatal loops.

which connects motor and premotor cortexes with the putamen; the executive loop, which connects lateral and medial prefrontal regions with the anterior caudate; the visual loop, which connects inferior temporal regions with the posterior caudate, and the motivational loop, which connects ventromedial prefrontal regions with the ventral striatum (including the nucleus accumbens and ventral caudate and putamen). Given the broad patterns of cortical projections to the basal ganglia, it is not surprising that the basal ganglia are associated with a large variety of functions, including motor control (Redgrave et al., 2010), cognitive coordination (Stocco et al., 2010), and emotional functions (Nakano et al., 2000).

The basal ganglia are involved in learning through a variety of inherent plasticity mechanisms. The best studied is N-Methyl-D-aspartate (NMDA) modulated long-term potentiation (LTP) at the corticostriatal synapse. Corticostriatal synapses also receive dopaminergic input and LTP is highly sensitive to the presence of dopamine (Pawlak and Kerr, 2008). Dopamine projections come from the midbrain, including the ventral tegmental area and portions of the substantia nigra. Some dopamine neuron activity is sensitive to reward expectation and is computationally well-described by reward prediction error (Schultz, 2002; Bromberg-Martin et al., 2010). This dopamine signal is well-suited to serve as a learning signal indicating the presence of unexpected rewards, thus the organism is more likely to repeat the behavior leading to the reward in the future.

The basal ganglia are particularly important in learning the relationship between sensory information and motor responses on the basis of trial by trial feedback (Seger, 2008; Shohamy et al., 2008). Computational models of dopamine-mediated plasticity within the direct pathway (Ashby and Ennis, 2006), and across pathways (Frank, 2005; Cohen and Frank, 2009) do an excellent job of accounting for learning in this type of task. Convergent evidence from a variety of species and techniques supports the view that the basal ganglia are critical for learning in these tasks (Yin and Knowlton, 2006; Graybiel, 2008; Balleine et al., 2009; Packard, 2009; Seger, 2009). Most habit learning tasks follow the same stimulus–response–reward/feedback task structure (Seger, 2009), and thus it is reasonable to propose that the basal ganglia should be important in habit learning.

REASSESSMENT OF HABIT LEARNING'S DEFINING CHARACTERISTICS

As the concept of habit learning developed, a number of different defining features were proposed. The following features were most commonly cited: inflexible, slow, unconscious, automatic, and insensitive to reinforcer devaluation. Here we revisit each of these defining criteria, asking the following questions about each: Why was it proposed? How precisely is it defined, and are there different definitions in use in different research areas? How accurately does this feature describe basal ganglia related learning? And, if relevant, how practical is the criterion for use with both human and non-human animals?

INFLEXIBLE

The characterization of habit learning as "inflexible" comes from Hirsh (1974), in contrast with flexible, context-dependent learning that was subserved by the hippocampus. Miskin et al. (1984) also included inflexible in their definition of habit learning, as

did Squire and colleagues in their development of the concept of non-declarative memory. Habit learning was independently characterized as inflexible by Dickinson (1985), who defined inflexibility in contrast to goal oriented behaviors, later shown to rely on prefrontal and dorsomedial striatal systems (Yin and Knowlton, 2006).

Flexible or inflexible has not been formally defined. The working definitions of these terms differ depending on whether habit learning is contrast with the hippocampal or prefrontal system. Within the hippocampal system, flexibility is often thought to be a consequence of individual memories formed by the hippocampus that can be applied to new situations. A commonly task thought to require hippocampally mediated flexibility is transitive inference, in which subjects are taught a set of ordinal relations, e.g., $A > B$, $B > C$, and then tested on whether they can infere that $A > C$ (Eichenbaum and Fortin, 2009). Some researchers have argued that the basal ganglia are limited to learning the individual ordinal relations and cannot support transitive inference or related phenomena (Myers et al., 2003; Shohamy et al., 2006). However, other research has found an opposite pattern of results, in which transitive inference relies on corticostriatal dopaminergic systems and is actually enhanced when the hippocampus is inhibited (Frank et al., 2003). Similar findings of hippocampal independence on other tasks thought to reflect flexibility (e.g., novelty transfer, Driscoll et al., 2004) indicate that this concept needs to be reassessed. Research is currently underway in a number of labs to better characterize what specific computational roles are played in inference tasks (Moustafa et al., 2010; Shohamy and Adcock, 2010).

Habit learning "inflexibility" is also defined in contrast with the sorts of flexibility enabled by executive functions subserved by the prefrontal cortex. In fact, executive functions were originally defined in clinical neuroscience as the ability to deal with novel or non-routine situations (Shallice, 1982). Prefrontal cortex enables flexible behavior through a variety of mechanisms involved in planning (setting goals, hypothesis formation, and testing), working memory (holding information online for several seconds), and cognitive control (the ability to execute plans in the face of distractions or other forms of interference; O'Reilly et al., 2010). Some have argued that the basal ganglia implement an inflexible learning process limited to past experience which then interacts with the flexible representations in prefrontal cortex.

Daw et al. (2005) argue that the basal ganglia select behaviors on the basis of the previous history of reinforcement, whereas the prefrontal cortex enables "model-based" control based on theories or strategies. Activity in the basal ganglia can be predicted by measures taken from reinforcement-learning modeling, specifically reward prediction (the estimate of the expected reward associated with choosing a particular behavior in the current state) and reward prediction error (the difference between the predicted and actually received reward). In this sense, the basal ganglia is inflexible because it is constrained to act in accordance with past reinforcement history. However, some studies have found patterns of basal ganglia activity that cannot be completely accounted for by reinforcement-learning models (Lopez-Paniagua and Seger, 2011).

One limitation of reinforcement-learning models is that they model the environment as a finite set of repeating states. In reality, organisms face situations that vary continuously, and need

to be able to generalize to similar but not identical situations. The basal ganglia are active in categorization tasks that require generalization to related, but novel, stimuli, indicating at least some flexibility (for review of some possible mechanisms, see Seger, 2008). It is unclear what the limits are to generalization in habit learning, and what role the basal ganglia may play in generalization.

SLOW OR INCREMENTAL

Habit learning was first characterized as slow or incremental by Miskin et al. (1984). As with "inflexible," this criterion was defined on the basis of learning in hippocampally ablated animals, in which learning required multiple trials. In contrast, animals with an intact hippocampus can show one-trial learning. The terms "slow" and "incremental" are often interpreted as requiring hundreds or thousands of trials, but this is not well established. Standard approaches from cognitive psychology involve examining learning curves for accuracy and reaction time, and potentially then habit learning can be thought to be complete when asymptote is achieved (see **Figure 3**, bottom section). Attempts to formalize learning rates come from reinforcement learning and state space modeling approaches. Reinforcement-learning approaches result in two common measures: reward prediction error, which is the measure of how unexpected the received reward is, and value, which is the expected reward associated with the current stimulus and associated action. When learning is the fastest, RPE is the highest and value rapidly changes. As a task is learned, RPE reduces to zero and value asymptotes toward its maximum (**Figure 3**, middle section).
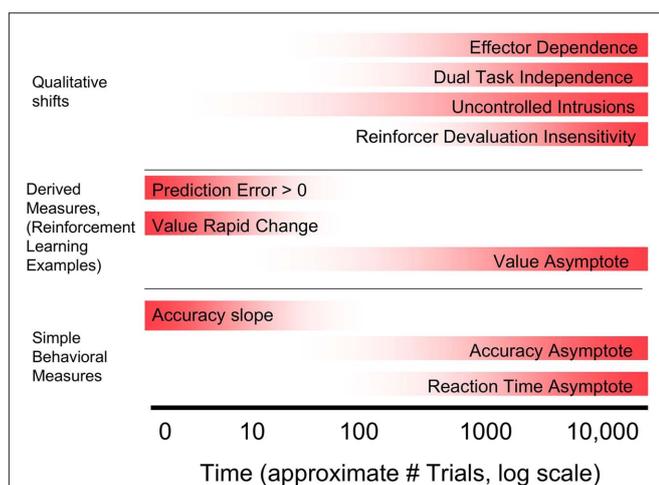

FIGURE 3 | Comparison of various possible criteria for habit learning that develop across learning on the basis of approximate point in training at which learning becomes habitual. Note that only criteria that can develop across training are included; criteria that are required across the entire time course of learning (unconscious, inflexible) are not. Top section: Qualitative criteria including two operational definitions of automaticity, and reinforcer devaluation. Middle and Bottom section: possible operational definitions of slow or incremental. Middle: Criteria based on computational modeling. This approach is illustrated using reinforcement-learning measures (reward prediction error, and value, or reward prediction) although other approaches can also be used. Bottom: Commonly used simple behavioral measures of learning.

Determining whether basal ganglia dependent learning is slow will depend on the operational definition of slow. However, it should be noted that basal ganglia dependent learning tasks vary greatly in how many trials it takes for subjects to reach maximal performance. Cromer et al. (2011) found that activity in the head of the caudate reached asymptote after five trials in a rule-learning paradigm. Delgado et al. (2005) found greatest caudate activity in an fMRI study during early learning (the first 8 repetitions of each stimulus) in comparison with later learning. Notably, these results are all from the caudate nucleus. Some researchers argue that the putamen should primarily subserve habitual learning. Studies that examine putamen activity during learning often find slower increases than in the caudate. However, activity levels in the putamen often follow behavior: activity reaches its maximum as behavioral accuracy reaches asymptote (Brasted and Wise, 2004; Williams and Eskandar, 2006), or as reinforcement-learning measures of learning, e.g., reward prediction, reach their maximum (Seger et al., 2010). Regional differences in learning speed are discussed further in the Conclusion.

If habit learning is acquired gradually, then when is performance fully habitual? Some people have argued that habits continue to develop even beyond the point at which behavioral measures cease to change, e.g., accuracy and reaction time reach their asymptotes. Grol et al. (2006) found continued practice related change in basal ganglia activity during these time points. Helie et al. (2010) and Waldschmidt and Ashby (2011) examined learning related changes long past the point at which accuracy reached asymptote, and found that basal ganglia activity continued to change and ultimately decreased to baseline levels. A more formal computational approach to measuring learning rates would be particularly helpful in this regard, as well as theories that can account for different levels of expertise and their neural correlates.

UNCONSCIOUS

Unconsciousness, as a defining feature of habit learning, stems from the inclusion of habit learning as a subtype of non-declarative memory by Squire and Zola-Morgan (1988, 1991). In their theory, declarative memory was accessible to consciousness, whereas non-declarative memory was not.

Consciousness can be difficult to define both on a practical and theoretical level. It is difficult to assess the degree of conscious access to knowledge in non-human animals, and impossible to assess verbalizable knowledge. Even with humans, there is debate about which measures of awareness are best for assessing whether there is conscious access to knowledge or not (Seth et al., 2008). Assessing awareness during task performance can affect the subject's strategic approach to the task, whereas assessing awareness after a task can easily miss information that might have been accessible to awareness during performance. On a more theoretical level, it is not always clear whether the relationship between awareness and learning is a necessary one; one logical possibility is that awareness is epiphenomenal and does not play a causal role in learning.

Recent research has found that the basal ganglia are involved in a wide variety of learning tasks, both ones in which learning is inaccessible to consciousness (Pessiglione et al., 2008), and in tasks in which subjects are aware of what they have learned, such as rule-learning tasks, arbitrary visuomotor learning tasks, and simple unstructured categorization tasks (Seger et al., 2011). Basal ganglia

recruitment is similar for relatively simple categorization tasks associated with high levels of verbalizable knowledge and more complex tasks associated with little verbalizable knowledge (Seger, 2008). Thus, basal ganglia do not seem to be exclusively associated with either conscious or unconscious learning. Furthermore, in recent research consciousness has proved to be a less reliable sign of hippocampal involvement in memory. The hippocampus has been shown to be required in several implicit learning tasks. These include contextual cuing, in which subjects become faster at searching repeated stimulus arrays (Greene et al., 2007), and some sequential relationships in the serial reaction time task (Schendan et al., 2003; Ergorul and Eichenbaum, 2006; Wilkinson et al., 2009).

AUTOMATIC

The concept of automaticity was developed in cognitive psychology by Shiffrin and Schneider (1977). It is itself a complex concept with four main characteristics. Three of these characteristics have already been discussed as potential defining features of habit learning: that automatic performance is unconscious, that the knowledge applied automatically is rigid or inflexible, and that automatic processes are acquired slowly and incrementally. The remaining characteristic is that automatic processes do not require the limited capacity cognitive mechanisms involved in short-term memory and selective attention. This leads to an operational definition that automatic tasks should be able to be performed in a dual task situation along with a demanding task that requires short-term memory and selective attention processes.

Although this definition is on the surface clear, in practice it is hard to know whether a particular dual task actually monopolizes appropriate limited capacity cognitive mechanisms. Recently, the concept of controlled processing has undergone extensive revision; there is no longer support for a simple modal model of memory, with a single limited capacity short-term memory store, though evidence suggests there are some general purpose or shared resources (Lavie, 2010). The modern view of short-term, or working, memory, and executive function includes qualitatively different short-term stores for different materials (Linden, 2007), and instead of a single attentional mechanism it includes a wide variety of cognitive control mechanisms (Banich et al., 2009; Braver et al., 2009). Learning in basal ganglia dependent tasks is often less affected by dual tasks than comparison tasks (Zeithamova and Maddox, 2006).

Dual task independence is also problematic when considering whether the basal ganglia are involved in habit learning, because the basal ganglia are in addition important for executive functions involved in task switching and selection. Thus, in a dual task situation any basal ganglia activity could be due to demands on the basal ganglia for coordinating the dual tasks, rather than for either of the tasks individually (Poldrack et al., 2005). Nevertheless, some researchers have used dual task methodologies successfully, such as Foerde et al. (2006) who found greater reliance on the basal ganglia for classification learning during dual task conditions in comparison with single task conditions. Interestingly, they found greater reliance on the putamen during dual task learning, which raises the possibility that dual tasks may load some corticostriatal networks more than others.

An alternative operational definition was proposed by Jacoby: that an automatic process will be performed regardless of a person's intentions, and thus will affect performance on a task even when the subject is attempting to not be affected (an exclusion task in Jacoby's terminology). This operational definition has not often been used in examining the basal ganglia in habit learning, though some researchers studying motor sequence learning have found that the striatum is recruited during automatic performance and is affected by prefrontal cortical mechanisms when subjects attempt to suppress the automatic performance (Destrebecqz et al., 2005).

REINFORCER REVALUATION INSENSITIVITY

The requirement that habit learning be insensitive to reinforcer revaluation comes from the field of animal learning. It has the advantage of being well defined, and it is clear how to apply this criterion experimentally, at least with non-human animal subjects. It is also clear how this criterion relates to learning in basal ganglia dependent tasks. This criterion dissociates the dorsomedial from dorsolateral striatum, with only the latter involved in habitual action.

The criterion does have some practical disadvantages. It requires two manipulations: first, the subject's value for the reinforcer must be changed (typically via feeding to satiation), and second, the behavior must be tested under conditions of extinction. It is unclear how effectively this procedure can be used with human subjects, who are more likely to notice that they are no longer being rewarded and change their behavior strategically (though some studies with humans have been published; Valentin et al., 2007). Second, it is unclear how the shift to reinforcer value independence corresponds to other meaningful transitions in the development of expertise, such as reaching asymptotic behavioral performance, or the emergence of dual task independence (Ashby et al., 2010; see also **Figure 3**).

CONCLUSION

As surveyed above, no single defining feature completely captures all the commonly-held beliefs about habit learning. Furthermore, the combination of these features are not always compatible. For example, the criterion of reinforcer devaluation and dual task independence each imply that learning should be considered habitual at a different point in training. We draw three main conclusions from our examination of habit learning and the basal ganglia. First, we provide a taxonomy of criteria for habit learning and divide them into two primary classes. Second, we examine patterns within the different corticostriatal loops and argue that the loops differ in the degree to which they meet criteria for habit learning, with the motor loop qualifying on more criteria than the executive loop. Third, we argue that the basal ganglia and corticostriatal systems interact with other neural systems and therefore that habit learning should not be assumed to exclusively require the basal ganglia, and describe some ways that these systems may interact.

The criteria for habit learning discussed above fall into two types. One type are criteria that can apply at any stage of learning, early or late. In particular, the criteria that learning is unconscious and inflexible were traditionally meant to characterize learning at all stages. Another type of criteria is based on the view

that habit learning develops across time and emerges as learning progresses. There have been various behavioral hallmarks of learning that have been proposed. **Figure 3** illustrates these hallmarks and indicates approximate points in time across training that they may be achieved. Broadly, behavioral landmarks can be divided into three subtypes. First, those based on simple analyses of behavior such as accuracy and reaction time, in which learning is defined as habitual when the measure reaches asymptote, or when the task is "overlearned" via continued training past the point of asymptote. Second, those that apply computational modeling techniques to extract latent parameters thought to characterize learning. The most commonly used approach is from reinforcement learning, in which there are two relevant parameters: reward prediction error and value, or reward prediction itself. Learning can be considered habitual when prediction error approaches zero, and reward prediction approaches asymptote. Both simple behavioral and model-based approaches provide potential operational definitions for the criterion of "slow or incremental." Third, qualitative criteria that are achieved at some point in learning. These include reinforcer devaluation insensitivity, automaticity defined as dual task insensitivity, and automaticity defined as inability to consciously control habitual knowledge. In addition, the field of motor learning suggests an additional possible criterion: the emergence of motor effector specificity. Across training, motor learning begins with relatively abstract representations that are accessible to multiple motor effectors, but learning become specific to the motor effector across training (Abrahamse et al., 2010).

The multiplicity, and at times incommensurability, of the different criteria for habit learning reinforces our belief that the field would benefit by moving towards more precise definitions of the various habit learning features. Formal mathematical or computational models will clarify exactly what is meant by slow and fast, flexible and inflexible learning and will allow for clear testable predictions. Formal models also have the advantage that they provide insight into potential underlying neural mechanisms; for example, reinforcement-learning modeling is particularly useful because it can be related to the firing patterns of dopaminergic neurons and the effects of dopamine on synaptic plasticity in the basal ganglia (Cohen and Frank, 2009; Moustafa and Gluck, 2011).

Another important lesson is that the basal ganglia is not a single unitary structure that is limited to a single cognitive domain. As described above, the basal ganglia and cortex interact in corticostriatal loops that implement different cognitive functions depending on the cortical regions involved. In **Table 1** we summarize evidence for whether each corticostriatal loop meets criteria for being habitual. Broadly, regions participating in the motor loop (putamen and motor cortex) meet most criteria for habit learning, whereas regions participating in the visual loop are not well studied, and regions participating in the executive loop have a mixed pattern of results, meeting criteria for habit learning on some dimensions, but missing it on many more. The results summarized in **Table 1** broadly support arguments made by researchers studying rodents who argue that the putamen (rodent dorsolateral striatum) is the neural substrate for habit learning and that the caudate (dorsomedial striatum) is involved in non-habitual goal-directed learning.

Table 1 | Habit learning criteria within dorsal corticostriatal loops.

Criterion	Executive	Visual	Motor
QUALITATIVE DIFFERENCES			
Inflexible defined as reinforcement based[1]	Mixed	?	Yes
Unconscious[2]	Mixed	?	Mixed
Automatic: dual task independent[3]	No	?	Yes
Automatic: PDP exclusion intrusions[4]	No	?	Yes
Reinforcer devaluation[5]	No	?	Yes
SLOW OR INCREMENTAL: COMPUTATIONAL MEASURES[6]			
Reward prediction error	Strong	Weak	Weak
Value	Weak	Strong	Strong
SLOW OR INCREMENTAL: SIMPLE BEHAVIORAL MEASURES			
Learning rate (slope)[7]	Strong	Weak	Weak
Learning asymptote	Weak	Strong	Strong
Changes beyond asymptote[8]	Mixed	?	Yes

[1]Daw et al. (2005) *and* Lopez-Paniagua and Seger (2011).
[2]Seger et al. (2011) *and* Pessiglione et al. (2008).
[3]Poldrack et al. (2005), Foerde et al. (2006), *and* Waldmann and Ashby (2011).
[4]Destrebecqz et al. (2005).
[5]Balleine et al. (2009) *and* Yin and Knowlton (2006).
[6]Haruno and Kawato (2006) *and* Seger et al. (2010).
[7]Brasted and Wise, 2004, Williams and Eskandar (2006), *and* Seger et al. (2010).
[8]Grol et al. (2006), Helie et al. (2010, *and* Waldschmidt and Ashby (2011).

Finally, it is important to avoid equating the behaviorally-defined habit learning system with the neurally-defined basal ganglia system. Given the complexity of habit learning, it likely recruits a number of neural systems in healthy, intact organisms. Neuroimaging studies of skill and habit learning tasks typically find learning related plasticity in several neural systems (Poldrack and Gabrieli, 2001; Poldrack et al., 2005). Probably the most studied system is the medial temporal lobe. However, other neural systems such as the cerebellum have an effect on learning and interact with basal ganglia system (Doyon et al., 2009). Exactly how these systems interact during habit learning is an open area of research. One approach is to postulate that habit learning and other systems learn independently and in parallel; the system that ultimately controls behavior is determined by competitive interactions between the systems (Ashby et al., 1998; Poldrack and Packard, 2003; Packard, 2009). Another approach assumes that initial learning is accomplished by a non-habit learning system, but that knowledge is transferred to the habit system across training (Ashby et al., 2010). When the basal ganglia and hippocampus systems are examined, some experimental results find antagonism, some cooperation, and some complete independence (see Seger and Miller, 2010, for a review). Among researchers studying human learning, an emerging view is that the hippocampus is recruited the first time a stimulus is seen in order to set up a memory representation of that stimulus, and that the basal ganglia can then utilize this representation when learning relations between the stimulus and the response (Meeter et al., 2008; Shohamy et al., 2008; Seger et al., 2011). Between the basal ganglia and prefrontal systems, the traditional view that the basal ganglia subserves habit learning led initially to arguments that cortical activity should precede activity in the basal ganglia. However,

more recent theories argue that the basal ganglia are active primarily during learning, and that well established habits are represented cortically (Pasupathy and Miller, 2005; Seger and Cincotta, 2006; Ashby et al., 2007, 2010).

REFERENCES

Abrahamse, E. L., Jiménez, L., Verwey, W. B., and Clegg, B. A. (2010). Representing serial action and perception. *Psychon. Bull. Rev.* 17, 603–623.

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychol. Rev.* 89, 369–406.

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

Ashby, F. G., Ennis, D. M., and Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychol. Rev.* 114, 632–656.

Ashby, F. G., and Ennis, J. M. (2006). The role of the basal ganglia in category learning. *Psychol. Learn. Mem.* 1–36.

Ashby, F. G., Turner, B. O., and Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci. (Regul. Ed.)* 14, 191–232.

Balleine, B. W., Lijeholm, M., and Ostlund, S. B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* 199, 43–52.

Banich, M. T., Mackiewicz, K. L., Depue, B. E., Whitmer, A. J., Miller, G. A., and Heller, W. (2009). Cognitive control mechanisms, emotion and memory: a neural perspective with implications for psychopathology. *Neurosci. Biobehav. Rev.* 33, 613–630.

Brasted, P. J., and Wise, S. P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur. J. Neurosci.* 19, 721–740.

Braver, T. S., Paxton, J. L., Locke, H. S., and Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7351–7356.

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834.

Brooks, D. N., and Baddeley, A. D. (1976). What can amnesic patients learn? *Neuropsychologia* 14, 111–122.

Cohen, M. X., and Frank, M. J. (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav. Brain Res.* 199, 141–156.

Cohen, N. J. (1984). "Preserved learning capacity in amnesia: evidence for multiple memory systems," in *The Neuropsychology of Memory*, eds. N. Butters and L. Squire (New York: Guilford), 83–103.

Cohen, N. J., and Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* 210, 207–210.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia* 6, 255–265.

Cromer, J. A., Machon, M., and Miller, E. K. (2011). Rapid association learning in the primate prefrontal cortex in the absence of behavioral reversals. *J. Cogn. Neurosci.* 23, 1823–1828.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Delgado, M. R., Miller, M. M., Inati, S., and Phelps, E. A. (2005). An fMRI study of reward-related probability learning. *Neuroimage* 24, 862–873.

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., Van Der Linden, M., Cleeremans, A., and Maquet, P. (2005). The neural correlates of implicit and explicit sequence learning: interacting networks revealed by the process dissociation procedure. *Learn. Mem.* 12, 480–490.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–78.

Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., Lehéricy, S., and Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behav. Brain Res.* 199, 61–75.

Driscoll, I., and Sutherland, R. J., Prusky, G. T., and Rudy, J. W. (2004). Damage to the hippocampal formation does not disrupt representational flexibility as measured by a novelty transfer test. *Behav. Neurosci.* 118, 1427–1432.

Eichenbaum, H., and Fortin, N. J. (2009). The neurobiology of memory based predictions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1183–1191.

Ergorul, C., and Eichenbaum, H. (2006). Essential role of the hippocampal formation in rapid learning of higher-order sequential associations. *J. Neurosci.* 26, 4111–4117.

Evans, J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.* 59, 255–278.

Foerde, K., Knowlton, B. J., and Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11778–11783.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J. Cogn. Neurosci.* 17, 51–72.

Frank, M. J., Rudy, J. W., and O'Reilly, R. C. (2003). Transitivity, flexibility, conjunctive representations, and the hippocampus. II. A computational analysis. *Hippocampus* 13, 341–354.

Graf, P., and Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *J. Exp. Psychol. Learn. Mem. Cogn.* 11, 501–518.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* 31, 359–387.

Greene, A. J., Gross, W. L., Elsinger, C. L., and Rao, S. M. (2007). Hippocampal differentiation without recognition: an fMRI analysis of the contextual cueing task. *Learn. Mem.* 14, 548–553.

Grol, M. J., de Lange, F. P., Verstraten, F. A., Passingham, R. E., and Toni, I. (2006). Cerebral changes during performance of overlearned arbitrary visuomotor associations. *J. Neurosci.* 26, 117–125.

Haruno, M., and Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J. Neurophysiol.* 95, 948–959.

Helie, S., Roeder, J. L., and Ashby, F. G. (2010). Evidence for cortical automaticity in rule-based categorization. *J. Neurosci.* 30, 14225–14234.

Hirsh, J. (1974). The hippocampus and contextual retrieval of information from memory: a theory. *Behav. Biol.* 12, 421–444.

Hull, C. L. (1934a). The concept of the habit-family hierarchy and maze learning: part I. *Psychol. Rev.* 41, 33–54.

Hull, C. L. (1934b). The concept of the habit-family hierarchy and maze learning: part II. *Psychol. Rev.* 41, 134–152.

Jacoby, L. L. (1991). A process dissociation framework: separating automatic from intentional uses of memory. *J. Mem. Lang.* 30, 513–541.

James, W. (1890). *Principles of Psychology*. New York: Henry Holt.

Keane, M. M., Gabrieli, J. D., Fennema, A. C., Growdon, J. H., and Corkin, S. (1991). Evidence for a dissociation between perceptual and conceptual priming in Alzheimer's disease. *Behav. Neurosci.* 105, 326–342.

Knowlton, B. J., Ramus, S., and Squire, L. R. (1992). Intact artificial grammar learning in amnesia: dissociation of classification learning and explicit memory for specific instances. *Psychol. Sci.* 3, 172–179.

Knowlton, B. J., and Squire, L. R. (1993). The learning of categories: parallel brain systems for item memory and category knowledge. *Science* 262, 1747–1749.

Lashley, K. S. (1930). Basic neural mechanisms in behavior. *Psychol. Rev.* 37, 1–24.

Lashley, K. S. (1950). "In search of the engram," in *Society of Experimental Biology Symposium*, Vol. 4 (Cambridge: Cambridge University Press), 454–480.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Curr. Dir. Psychol. Sci.* 19, 143–148.

Linden, D. E. (2007). The working memory networks of the human brain. *Neuroscientist* 13, 257–267.

Lopez-Paniagua, D., and Seger, C. A. (2011). Interactions within and between corticostriatal loops during component processes of category learning. *J. Cogn. Neurosci.* 23, 3068–3083.

Meeter, M., Radics, G., Myers, C. E., Gluck, M. A., and Hopkins, R. O. (2008). Probabilistic categorization: how do normal participants and amnesic patients do it? *Neurosci. Biobehav. Rev.* 32, 237–248.

Milner, B. (1962). "Les troubles de la memoire accompagnant les 1esions hippocampiques bilaterales," in *Physiologic de l'Hippocampe* (Paris: Centre National de la Recherche Scientifique), 257–272.

Miskin, M., Malamut, B., and Bachevalier, J. (1984). "Memories and habits: two neural systems," in *Neurobiology of Learning and Memory*, eds. G. Lynch, J. L. McGaugh, and N. M. Weinberge (New York: Guilford), 65–67.

Moustafa, A. A., and Gluck, M. A. (2011). A neurocomputational model of dopamine and prefrontal-striatal interactions during multicue category

learning by Parkinson patients. *J. Cogn. Neurosci.* 23, 151–167.

Moustafa, A. A., Keri, S., Herzallah, M. M., Myers, C. E., and Gluck, M. A. (2010). A neural model of hippocampal-striatal interactions in associative learning and transfer generalization in various neurological and psychiatric patients. *Brain Cogn.* 74, 132–144.

Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Kluger, A., Ferris, S., Golomb, J., Schnirman, G., and Schwartz, R. (2003). Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J. Cogn. Neurosci.* 15, 185–193.

Nakano, K., Kayahara, T., Tsutsumi, T., and Ushiro, H. (2000). Neural circuits and functional organization of the striatum. *J. Neurol.* 247, V1–V15.

Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.* 19, 1–32.

O'Reilly, R. C., Herd, S. A., and Pauli, W. M. (2010). Computational models of cognitive control. *Curr. Opin. Neurobiol.* 20, 257–261.

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

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

Pawlak, V., and Kerr, J. N. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neurosci.* 28, 2435–2446.

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., and Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59, 561–567.

Poldrack, R. A., and Gabrieli, J. D. (2001). Characterizing the neural mechanisms of skill learning and repetition priming: evidence from mirror reading. *Brain* 124, 67–82.

Poldrack, R. A., and Packard, M. G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41, 245–251.

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., and Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *J. Neurosci.* 25, 5356–5364.

Reber, A. S. (1967). Implicit learning of artificial grammars. *J. Mem. Lang.* 6, 855–863.

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., Lehericy, S., Bergman, H., Agid, Y., DeLong, M. R., and Obeso, J. A. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat. Rev. Neurosci.* 11, 760–772.

Schacter, D. L. (1987). Implicit memory: history and current status. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 501–518.

Schendan, H. E., Searl, M. M., Melrose, R. J., and Stern, C. E. (2003). An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* 37, 1013–1025.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263.

Seger, C. A. (1994). Implicit learning. *Psychol. Bull.* 115, 163–196.

Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neurosci. Biobehav. Rev.* 32, 265–278.

Seger, C. A. (2009). "The involvement of corticostriatal loops in learning across tasks, species, and methodologies," in *The Basal Ganglia IX*, eds H. J. Groenewegen, P. Voorn, H. W. Berendse, A. B. Mulder, and A. R. Cools (New York: Springer-Verlag), 25–39.

Seger, C. A., and Cincotta, C. M. (2005). The roles of the caudate nucleus in human classification learning. *J. Neurosci.* 25, 2941–2951.

Seger, C. A., and Cincotta, C. M. (2006). Dynamics of frontal, striatal, and hippocampal systems during rule learning. *Cereb. Cortex* 16, 1546–1555.

Seger, C. A., Dennison, C. S., Lopez-Paniagua, D., Peterson, E. J., and Roark, A. A. (2011). Dissociating hippocampal and basal ganglia contributions to category learning using stimulus novelty and subjective judgments. *Neuroimage* 55, 1739–1753.

Seger, C. A., and Miller, E. K. (2010). Category learning in the brain. *Annu. Rev. Neurosci.* 33, 203–219.

Seger, C. A., Peterson, E. J., Cincotta, C. M., Lopez-Paniagua, D., and Anderson, C. W. (2010). Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *Neuroimage* 50, 644–656.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., and Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn. Sci. (Regul. Ed.)* 12, 314–321.

Shallice, T. (1982). Specific impairments of planning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298, 199–209.

Shiffrin, W., and Schneider, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychol. Rev.* 84, 1–66.

Shohamy, D., and Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends Cogn. Sci. (Regul. Ed.)* 14, 464–472.

Shohamy, D., Myers, C. E., Geghman, K. D., Sage, J., and Gluck, M. A. (2006). L-dopa impairs learning, but spares generalization, in Parkinson's disease. *Neuropsychologia* 44, 774–784.

Shohamy, D., Myers, C. E., Kalanithi, J., and Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neurosci. Biobehav. Rev.* 32, 219–236.

Squire, L. R., and Zola-Morgan, S. (1988). Memory: brain systems and behavior. *Trends Neurosci.* 11, 170–175.

Squire, L. R., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science* 253, 1380–1386.

Stocco, A., Lebiere, C., and Anderson, J. R. (2010). Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol. Rev.* 117, 541–574.

Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.

Waldschmidt, J. G., and Ashby, F. G. (2011). Cortical and striatal contributions to automaticity in information-integration categorization. *Neuroimage* 56, 1791–1802.

Wilkinson, L., Khan, Z., and Jahanshahi, M. (2009). The role of the basal ganglia and its cortical connections in sequence learning: evidence from implicit and explicit sequence learning in Parkinson's disease. *Neuropsychologia* 47, 2564–2573.

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

Winograd, T. (1972). Understanding natural language. *Cogn. Psychol.* 3, 1–191.

Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476.

Zeithamova, D., and Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Mem. Cognit.* 34, 387.