Reward-related learning via multiple memory systems

Mauricio R. Delgado and Kathryn C. Dickerson
Department of Psychology & Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ 07102

Abstract

The application of a neuroeconomic approach to the study of reward-related processes has provided significant insights in our understanding of human learning and decision-making. Much of this research has primarily focused on the contributions of the cortico-striatal circuitry, involved in trial and error reward learning. As a result, less consideration has been allotted to the potential influence of different neural mechanisms such as the hippocampus, or to more common ways in human society in which information is acquired and utilized to reach a decision, such as through explicit instruction rather than trial and error learning. This review examines the basic and applied value of examining the individual contributions of multiple learning and memory neural systems and their interactions during human decision-making in normal individuals and neuropsychiatric populations. Specifically, the anatomical and functional connectivity across multiple memory systems are highlighted to suggest that probing the role of the hippocampus and its interactions with the cortico-striatal circuitry via the application of model-based neuroeconomic approaches may provide novel insights into several neuropsychiatric populations who suffer from damage to one of these structures, and as a consequence have deficits in learning, memory, or decision-making.

Keywords

Striatum; hippocampus; prediction error; dopamine; reinforcement learning; memory systems

Introduction

There is an emerging literature investigating the neural basis of decision-making that encompasses the field of neuroeconomics (1). Across a variety of techniques and interdisciplinary efforts, neuroeconomics research has emphasized two key points: a) choice behavior is influenced by learning systems involved in trial and error learning (such as instrumental conditioning), and b) cortico-striatal neural systems are involved in the computation of basic principles of conditioning and reinforcement learning that underlie simple decision-making. This cross-species account of decision-making represents a significant contribution to the field, and informs various influential models of decision-making (2–4), but it does not fully reflect the complexity of learning in humans, where the existence of multiple memory systems has been well documented (5, 6). Given the intricacy of the environment, it would be peculiar if one neural system was primarily responsible for learning signals that contribute to decision-making. In fact, choices made in everyday human society are not solely driven by experience through trial and error learning, as we also learn and allow our decisions to be influenced by communication (e.g., talking to others), knowledge of rules (e.g., when to cross the street) and facts (e.g., this person is...
Thus, consideration of how distinct neural circuits interact during reward-related learning to influence choice behavior can enhance the contributions of neuroeconomics research and contain potential clinical applications.

In this review, we focus on the distinct and common contributions of the basal ganglia (BG; involved in learning through trial and error), and the medial temporal lobes (MTL), particularly the hippocampus (involved in declarative learning of rules and facts) during learning and how they translate to decision-making in both healthy and clinical populations.

We first discuss the concept for and evidence supporting multiple memory systems, followed by subsections describing relevant learning and decision-making research involving the BG and MTL, along with potential interactions between these structures. Finally, we evaluate how neuroeconomics research may be expanded upon in the context of multiple memory systems and synthesized to help inform neuropsychiatric disorders that involve compromised BG and MTL structures.

**Multiple memory systems**

The values we assign to different stimuli and prior experiences tend to influence how we make decisions. For instance, an individual may choose to shop at the local family-owned grocery store versus a new larger grocery store in town because of prior positive experiences at the local establishment (Figure 1). Within this example, one explanation of how the choice is made is based on a greater value assigned to the local shop because of reinforcement learning. An impressive collection of studies has highlighted the role of cortico-striatal circuits along with modulation of dopamine neurons as integral for this type of learning (7–9). Not surprisingly, individuals with pathologies that involve BG dysfunction, such as Parkinson’s disease (PD), have difficulties learning by trial and error and updating learned contingencies (10–15), which may manifest as perseverative choices in some contexts (16). Defined by a loss of the majority of dopaminergic neurons in the midbrain, PD consists of several motor impairments (e.g., akinesia) as well as cognitive deficits (17, 18). Importantly, patients afflicted with PD can learn through different neural mechanisms, such as MTL structures involved in declarative learning, which can also help guide choices (19) and highlights PD as a candidate disorder for investigating multiple memory systems.

Over the past few decades, an increasingly supported theory suggests that memory is not a unitary process, but rather consists of multiple processes which engage distinct neural substrates (5, 6, 20). A major dissociation is between declarative and nondeclarative memory. Declarative memory consists of knowledge of facts and events; this type of memory may be rapidly formed and later readily recalled, is flexible in nature and relies on the integrity of the MTL (21). In contrast, nondeclarative memory encompasses several different types of learning including habit learning, procedural, conditioning, and priming (20). Such learning is thought to involve regions of the BG, along with cortical structures and the amygdala, and is characterized by a slower time course of acquisition and less flexibility with respect to updating, as in the case of habits (21, 22).

Support for this theory comes from a vast array of research ranging from non-human animal investigations to studies with clinical populations (19). In a classic paradigm, a rodent is trained to navigate a maze to retrieve a food reward (23, 24). A clear dissociation between multiple memory systems is apparent by poor performance in a spatial learning version of a radial maze (the “win-shift” task) in rats with lesions to the fimbria-fornix – the output of the MTL – whereas rats with lesions in the BG are unimpaired at remembering the previously rewarded locations. In contrast, MTL lesioned rats show intact performance in a stimulus-response version of the maze (the “win-stay” task) where approach behavior is
dictated by a conditioned cue, and is impaired in rats with lesions in the dorsal striatum, an input structure in the BG (23–25). This dissociation between the MTL and BG with respect to declarative and nondeclarative-like learning tasks respectively is mirrored across different discrimination related paradigms, such as delayed-nonnatching to sample (26) or concurrent discrimination (27) which suggests that lesions to the MTL adversely affect a monkey’s ability to successfully complete discrimination tasks (27, 28), but not skill learning (28). Additionally, patients suffering from MTL amnesia have extremely compromised declarative memory, while nondeclarative types of learning (e.g., skill learning) remain intact (29–31). Likewise, patients afflicted with PD have difficulty with some forms of trial and error learning, while declarative knowledge persists (30, 32).

There are several examples where this dissociation is challenged, however, primarily based on disagreements on what constitutes declarative versus nondeclarative memory (33), in part as evidenced by conflicting results in patient populations (34–39). For instance, successful probabilistic trial and error learning has been observed in patients with PD (39), leading to the argument that deficits observed may be more nuanced than originally hypothesized, dependent on the type of task (e.g., skill learning versus perceptual) and the locus of impairment (e.g., cortico-striatal circuits versus distributed neocortical association regions) (35, 38). The emergence of neuroeconomics research promises to shed light on this debate by introducing new paradigms and quantitative model-based approaches. A popular method budding from cognitive and computational neuroscience research (40, 41), that is rooted in classical learning theories (42), is the application and characterization of model-based learning signals (9). Widely employed in the interdisciplinary study of neuroeconomics, these model-derived signals can be used to quantify how a cognitive process is implemented in a specific brain region and can be expanded across different memory systems to explore how such systems interact in the context of goal-directed behaviors.

The basal ganglia & reward learning

Learning what choice provides the best outcome can be an arduous process often accomplished by trial and error. In an experimental setting, an animal will learn to predict positive outcomes after repeated successful pairings of a reward with a conditioned stimulus, such as a light or a lever, leading to increased approach behavior towards the stimulus. Early in the decision-making process, therefore, the animal learns to integrate the action-outcome association (e.g., lever-pressing leads to liquid reward) with the value attributed to such outcome (e.g., high value for liquids when thirsty), in a manner that allows constant updating of action-contingencies (e.g., when lever does not yield predicted outcome, a prediction error is coded, leading to updating of current contingencies). These types of studies illustrate how basic learning principles of conditioning can shape current decision-making and future behavior. Across species and through different methodologies, research has converged on the role of the cortico-striatal circuitry in the computation of basic learning principles of conditioning and reinforcement underlying simple decision-making (9, 43–46).

The BG complex is comprised of multiple cortico-striatal loops which exert distinct functions (8). The striatum, in particular, has been linked to various types of learning involving procedural skills (47), habits (30, 48, 49), and reward learning (43, 44, 50). Functional dissociations between subdivisions of the striatum have also been proposed, with the dorsomedial striatum involved in processing action-contingency (43, 51), goal-directed learning (52), and instrumental conditioning (46), and the dorsolateral striatum linked with habit and motor learning (49). Activity in the ventral striatum has been primarily associated with reward-related processing, demonstrating sensitivity to changes in subjective value (9, 53) and prediction-based learning (46). Not surprisingly, these striatum subdivisions are also
distinguished with respect to cortical connectivity, with the dorsomedial striatum connected
to an associative network with the prefrontal and parietal association cortices, the
dorsolateral striatum connected to sensorimotor regions and the ventral striatum connected
with ventral portions of the prefrontal cortex (e.g., orbitofrontal cortex) (44, 49).

One of the most thought-provoking findings with respect to BG circuitry and affective
learning concerns the role of midbrain dopaminergic neurons in coding prediction error
signals (54). It was observed that dopaminergic neurons in non-human primates respond to
unexpected rewards (e.g., drop of juice) but no longer respond to the juice once its delivery
becomes predictable. Instead, the neurons fire to the earliest predictor of a reward (e.g., tone
predicting juice). When a predicted reward fails to occur, dopamine neurons show a
depression in neuronal firing, thought to be a prediction error signal used to update
contingencies during affective learning. Using functional magnetic resonance imaging
(fMRI), prediction error learning signals have correlated with blood oxygen level dependent
(BOLD) responses in the human midbrain (55), but primarily in dopaminergic targets such
as the striatum using a variety of rewards (e.g., juice, money) and paradigms (e.g., classical
and instrumental conditioning) where predictions and contingencies are typically acquired
and updated through trial and error (46, 56, 57).

The human striatum has been shown to be involved in various facets of reward-related
processing that encompass affective learning and inform future decision-making (9, 44) –
from anticipating or predicting a potential reward (58) to coding the differential response
between positive and negative outcomes (59), particularly when a contingency between an
action and an outcome is perceived (51). Thus, it is not surprising that BOLD responses in
the human striatum correlate with prediction error learning signals during conditioning
paradigms (46, 60), with stronger correlations suggesting greater learning rates (61) and
exploitative, rather than explorative, decision-making (62). More recently, such model-based
approaches have begun to be applied to clinical populations with BG dysfunction, namely
PD patients. Consistent with the pathology of the disorder, prediction error signals have
been found to be preserved in more ventral regions of the striatum, but impaired in
dorsolateral regions relative to healthy controls (13) suggesting a potential explanation for
why PD patients are impaired in probabilistic trial and error paradigms (30, 32, 39, 63).

The medication status of PD patients during scientific studies (e.g., on or off dopaminergic
medicine) has also become a formidable way of investigating the potential role of a
dopaminergic learning signal and its influences on decision-making. The mere enhancement
or reduction of dopaminergic function by pharmacological treatment (L-DOPA or
haloperidol respectively) in healthy participants modulates the magnitude of a prediction
error learning signal, suggesting that increases in dopaminergic function lead to choices
being perceived as more rewarding (64). Similar manipulations in PD patients hint at
specific impairments in learning from positive outcomes when in an unmedicated state (but
intact learning from negative outcomes) and normal learning from positive (but not
negative) outcomes when treated with dopaminergic agonists (11, 13, 16). Interestingly, in
some individual cases of PD, administration of dopaminergic agonists targeting the dorsal
striatum can lead to enhanced impulse control impairments such as increased gambling
behavior (65–67), posited to be due to an overflow of dopamine in the ventral striatum (68–
70). Neuroeconomic approaches of modeling dopaminergic learning signals raise interesting
possibilities for future clinical research involving BG dysfunction (13, 16) which allow for
extensions beyond the classification of a particular impairment (e.g. learning deficits in
feedback-based tasks in PD) to identification of the mechanism underlying such impairment
(e.g., learning deficits due to prediction error impairments and medication status). An
exciting future direction is the incorporation of multiple learning signals arising from

Biol Psychiatry. Author manuscript; available in PMC 2013 July 15.
distinct structures, such as the MTL, that underlie different forms of learning signals which participate in goal-directed behavior.

**Medial temporal lobe, declarative learning & reward-related contributions**

When deciding to try out the new grocery store in town, an individual may rely not only on past experiences with the local shop, but also declarative information obtained regarding the new store. One may read a positive review about the wide array of food available at the new store or be informed by a friend about the great selection of produce (Figure 1). This type of declarative information, thought to be dependent on the integrity of the MTL, is a major influence on our decisions. The MTL consists of many substructures, including the hippocampus, entorhinal cortex, the parahippocampus and perirhinal cortices (71). Collectively, these substructures support multiple types of declarative learning including memory for facts and events (21), novelty detection (72, 73), mediating flexible transfer of knowledge (31, 74, 75), as well as spatial learning and navigation (76, 77). Studies in non-human animals and patient populations with damage to the MTL have demonstrated the importance of this structure in forming new declarative memories and flexibly applying knowledge in new scenarios (6, 14, 31, 78).

More recently, the hippocampus has been highlighted in reward-related studies (79–86). In comparison to the substantial existing evidence of the involvement of distinct MTL structures in various aspects of declarative memory, however, there is still less known about the role of the hippocampus in reward-related learning and decision-making. One important consideration with regards to the hippocampus and reward processing is its connectivity with cortico-striatal circuits (44, 83). Specifically, the hippocampus is anatomically connected with the striatum (87, 88) and it has been suggested that cortical mechanisms may mediate the interaction between them (89). The potential functional significance of these interactions is illustrated in specific learning contexts, such as spatial learning (81). In such context, distinct but parallel contributions to successful path navigation have been proposed for the hippocampus, involved in prospection or the representation of future paths, the ventral striatum, involved in the valuation of future rewards, and the dorsal striatum, tracking action-related task components, that together aid decision-making (77).

The hippocampus also receives projections from midbrain dopaminergic structures (90, 91) and an alternative, perhaps complementary view, is that dopaminergic modulation provides a learning signal for both the hippocampus and the striatum (75, 89). Whether these dopaminergic signals exert similar or distinct influences upon both these structures is an intriguing question for neuroeconomics currently. A popular hypothesis posits that a distinct function of dopaminergic input into the hippocampus facilitates the entry of information into long-term memory by novelty detection (73). Consistent with this idea, dopamine has been thought to play a key role in adaptive memory and determining what information is learned and subsequently remembered (e.g., those associated with a high reward value (79, 83, 86)). This connectivity between the hippocampus and midbrain dopaminergic centers is also involved in mediating the association between different environmental contexts and reward (92), and the ability to generalize prior learned associations into future decisions (74). Research on neuropsychiatric populations has specified a double dissociation between patients with PD and those with hippocampal atrophy during initial learning (impaired in PD) and the subsequent ability to transfer (impaired in hippocampal patients) during an associative learning task (31), highlighting the role these regions play in learning and flexibly using information. Interestingly, hippocampal damage does not typically correlate with deficits in performance in some learning tasks associated with dopamine and striatal function, such as trial and error feedback-based learning (14, 93). Moreover, during
instrumental conditioning tasks with delayed reinforcement, rats with lesions of the hippocampus actually display better learning than control animals (94), perhaps because such lesions hinder the formation of context-outcome associations, focusing instead on action-outcome contingencies dependent on cortico-striatal systems.

A potential similar function of dopaminergic modulation in both the hippocampus and striatum may be the encoding of prediction error learning signals during specific learning contexts. In associative learning tasks, learning-related spiking activity is correlated with a reward-predictive value signal in monkey hippocampal neurons (82), while in human probabilistic learning paradigms, activity in the hippocampus has been suggested to track differences between positive and negative outcomes (95) potentially encoding a feedback-based model-derived reward prediction error signal (96) in parallel with cortico-striatal systems. A challenge for neuroeconomics currently is understanding the functional and temporal significance of this learning signal as the human hippocampus has also been associated with coding for the uncertainty of cue-outcome associations (97), a novelty response (98, 99), and a mismatch signal (100, 101). The few reports that exist thus far advocating an association between the hippocampus and outcome prediction (95, 96, 102) do not necessarily claim that the hippocampal activity exclusively correlates with a scalar prediction error, since plausible alternative models, such as a mismatch signal or an unsigned prediction error, were not tested. Rather, these initial reports point to neuroeconomic approaches as a potential tactic for advancing these questions for future consideration. It is likely that the nature of a learning signal coded in the hippocampus depends on the learning context (e.g., probabilistic task with declarative and nondeclarative elements), and the anatomical location and functional connectivity of loci of activation (e.g., CA1 subfield of the hippocampus has been linked with tracking associative mismatches (100)). Nevertheless, understanding the contributions of the dopaminergic learning signal encoded in the hippocampus may present new directions in decision-making research and translational applications for neuropsychiatric disorders by highlighting how humans flexibly use information in novel contexts to guide future choices.

**Interactive memory systems during decision-making**

It is theorized that the inadequacy of a single learning and memory system’s ability to deal with a complex situation may have led to the evolution of multiple memory systems (5). To illustrate this in the context of decision-making and neuroeconomics more specifically, one can consider “the trust game”, a classic paradigm in behavioral economics used to investigate risky decision-making and social interactions (103). Briefly, in its simplest form the game involves a one shot interaction between two players, where the first player, the proposer, is given an initial sum of money (e.g., $20) and decides whether to keep it and end the game or invest some or all of that money with the second player, the responder. Any amount invested gets tripled (e.g., $60), and the responder now decides whether to share (or not) some of the investment back to the proposer (e.g., each player gets $30). It is a risky proposition, but a potentially profitable one for the proposer who has to show “trust” in their partner. Adaptations of this game to neuroimaging settings involving repeated interactions evoke trial and error learning and suggest that reputation building is dependent on learning signals being processed in the striatum (104–106), with a shift in the magnitude of the BOLD signal in the striatum to the earliest prediction of a trustworthy interaction based on the development of reputation (106), akin to the shift observed in reinforcement learning models of prediction error signals (54, 57).

The concept of multiple memory systems can be applied to these complex learning and decision-making situations where other types of information can influence decisions. For instance, a proposer may have either good or bad priors with respect to the responder, such
as knowledge about their ethical and business practices. In this situation, the additional information affects the belief about probability of reciprocation and decisions to trust (104, 105, 107). Furthermore, learning signals in the striatum involved in updating predictions based on current outcomes are diminished when prior declarative information is present, as if a proposer places less weight on the current experience or outcome due to the value assigned to the prior information (105).

One hypothesis is that declarative knowledge may bias decision-making in this case, potentially overriding exploration of alternatives favored by trial and error learning mechanisms, and suggesting dynamic, perhaps competitive, interactions between the MTL and BG that can affect future choices. This is consistent with evidence from probabilistic learning studies which propose that these memory systems may compete during some learning scenarios (108, 109). An alternative hypothesis is that MTL and BG may at times cooperate during learning situations, thus allowing compensation for a faulty system in a state where the brain is diseased (110, 111) or when such an interaction is behaviorally relevant (112).

A third idea is that the nature of the interactions between MTL and BG involves parallel processing of information, with the interactions being competitive or cooperative depending on task demands (25, 113, 114). This debate across the theory of multiple memory systems regarding the nature of the interactions between MTL and BG can profit from neuroeconomics research and the setting conjectured by experimental paradigms such as the trust game. For example, in the trust game the dynamic interactions between the MTL and BG could be perceived as competitive, however reinforcement learning signals have been observed during probabilistic learning tasks in both structures (95, 96), conceivably suggesting parallel engagement or even a cooperative interaction between these regions. Introducing instructed knowledge to a probabilistic learning paradigm diminishes the reinforcement learning signals in these areas, concurrent with an increase in activity in the dorsolateral prefrontal cortex (95), posited to mediate the interaction between the MTL and BG (89). Thus, parallel processing within these distinct neural circuits in the MTL and BG, which are influenced by dopaminergic input and shaped by cortical modulation, may underlie learning and decision-making processes that occur during complex social interactions such as the trust game.

**Further considerations**

Several models posited to underlie decision-making processes have been highlighted in the literature and merit brief discussion (3, 4, 9, 115). One dual system account of decision-making, for example, distinguishes between a cognitively-based, controlled system (deliberative system) and an automatic, emotional system (impulsive system), acutely important during inter-temporal choices (116–119). Deliberative processes have been posited to involve prefrontal and parietal cortices, while automatic processes have been linked with subcortical regions such as the striatum ((4, 119); although see (116)). Another influential model highlights goal-directed and habitual learning as two key systems involved in guiding choice behavior (115, 120), with goal-directed learning involving the anterior caudate nucleus and medial prefrontal regions, and habitual learning associated with the posterior putamen (48, 49, 115). Current work is also exploring a distinction between reinforcement learning error signals related to reward prediction, referred to as model-free, and expectations regarding current states given previous choices, referred to as model-based (121, 122).

Integrating these models in a cohesive account of how distinct neural systems involved in learning, such as the striatum (typically highlighted across models) and the hippocampus...
(observed less often), can inform decision-making is an important future endeavor. One recent proposal posits four key decision-making systems in the brain: goal-directed (dorsolateral PFC and dorsomedial striatum), habitual (dorsolateral striatum), episodic (MTL), and Pavlovian (123). These distinct systems help guide decisions independently, but may also interact to steer choices depending on the context of the decision.

As this review primarily focused on the potential contributions of learning signals associated with the BG and MTL to decision-making, less discussion was attributed to other important regions that contribute to learning and decision-making, namely the amygdala and the prefrontal cortex (124, 125). The amygdala is a critical structure for aversive learning (126) and emotional memory (127) that has implications for decision-making in particular contexts, such as when decisions involve ambiguity (128) or are framed as gains or losses (129). Interestingly, the amygdala is another potential link between interactions of the MTL and BG, as it has been shown to modulate activity in these structures during specific learning tasks (e.g., win-shift and win-stay respectively; (130)). Thus, a clear future direction is how learning signals originating from the amygdala can influence the MTL and BG during decision-making processes.

In contrast, the involvement of the PFC in decision-making processes has been well documented (4, 9, 115). Both dorsolateral and ventromedial prefrontal cortical regions, in particular, have been associated with the representation of goal values (9) and integrating reward-related information along with a striatal and dopaminergic network (131), consistent with the previously highlighted theories of cortical modulation of MTL and BG interactions (89). Future research using computational models of learning signals will prove useful in further understanding the relationship between PFC regions and the striatum during action-outcome learning (95) as potentially informed by cortico-hippocampal connections that mediate rule governance (132).

Conclusion

Understanding how neural structures underlying distinct learning and memory systems interact has profound potential to inform neuropsychiatric disorders, especially in populations where decision-making is adversely affected by the disease or disease treatment (e.g., pathological gambling and PD (65–67)). The advent of fMRI has allowed for extension of elegant animal models and patient studies in support of multiple learning and memory systems – although a debate remains with respect to how underlying neural structures, namely the BG and MTL, interact to facilitate learning and decision-making. The emergence of neuroeconomic approaches, such as the application of signals derived from models of cognitive processes that are used to more carefully identify neural mechanisms, promises to shed light on this debate and extend this knowledge beyond basic principles of learning that inform decision-making to other important forms of learning common to humans (e.g., instructed and observational learning (95, 133)). The evolution of neuroeconomics based models, paradigms, and initial efforts of their application in special neuropsychiatric populations bodes well for future advances in understanding the connectivity and interactions of neural mechanisms involved in learning, memory, and decision-making to promote more efficient diagnosis and treatment options.

Acknowledgments

This work was supported by a National Institute of Mental Health grant (MH-084041) and a National Science Foundation Grant (0718153).
References


_Biol Psychiatry_. Author manuscript; available in PMC 2013 July 15.
Figure 1.
Multiple memory systems during human decision-making. In everyday life, humans are required to make hundreds of decisions, ranging from the complex, e.g., what career path to follow, to the simple, such as where to buy groceries. One may rely on past experiences (positive or negative) in order to make future decisions, as illustrated in the figure by positive past experiences at a local Mom & Pop Grocery. Repeated good experiences at one store may lead to continually shopping at the establishment, a form of learning and decision-making thought to involve the basal ganglia (BG), particularly the input unit of the BG – the striatum. However, decisions can be influenced by distinct neural systems underlying multiple learning and memory systems in addition to the striatum, including the hippocampus, part of the medial temporal lobe (MTL) which may store declarative information about one’s environment. For example, speaking with a friend about a new larger grocery store in town or reading a favorable review in the newspaper may lead one to make the decision to try out the new store. In this manner multiple memory systems may interact in everyday decision making, via either direct connections between these regions (MTL to striatum in BG) or via midbrain dopaminergic influences stemming from the ventral tegmental area and substantia nigra (VTA/SN), hypothesized to facilitate learning and memory processes in the MTL and BG. Abbreviations: MTL, medial temporal lobe; BG, basal ganglia; VTA, ventral tegmental area; SN, substantia nigra.