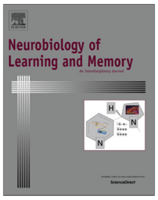
Neurobiology of Learning and Memory xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme



Multiple memory systems as substrates for multiple decision systems

Q1 Bradley B. Doll [a,b,*], Daphna Shohamy [b], Nathaniel D. Daw [a,c]

Q2 [a] Center for Neural Science, New York University, New York, NY, United States
[b] Department of Psychology, Columbia University, New York, NY, United States
[c] Department of Psychology, New York University, New York, NY, United States

ARTICLE INFO

Article history:
Received 19 November 2013
Revised 22 April 2014
Accepted 29 April 2014
Available online xxxx

Keywords:
Reinforcement learning
Relational memory

ABSTRACT

It has recently become widely appreciated that value-based decision making is supported by multiple computational strategies. In particular, animal and human behavior in learning tasks appears to include habitual responses described by prominent model-free reinforcement learning (RL) theories, but also more deliberative or goal-directed actions that can be characterized by a different class of theories, model-based RL. The latter theories evaluate actions by using a representation of the contingencies of the task (as with a learned map of a spatial maze), called an "internal model." Given the evidence of behavioral and neural dissociations between these approaches, they are often characterized as dissociable learning systems, though they likely interact and share common mechanisms.

In many respects, this division parallels a longstanding dissociation in cognitive neuroscience between multiple memory systems, describing, at the broadest level, separate systems for declarative and procedural learning. Procedural learning has notable parallels with model-free RL: both involve learning of habits and both are known to depend on parts of the striatum. Declarative memory, by contrast, supports memory for single events or episodes and depends on the hippocampus. The hippocampus is thought to support declarative memory by encoding temporal and spatial relations among stimuli and thus is often referred to as a relational memory system. Such relational encoding is likely to play an important role in learning an internal model, the representation that is central to model-based RL. Thus, insofar as the memory systems represent more general-purpose cognitive mechanisms that might subserve performance on many sorts of tasks including decision making, these parallels raise the question whether the multiple decision systems are served by multiple memory systems, such that one dissociation is grounded in the other.

Here we investigated the relationship between model-based RL and relational memory by comparing individual differences across behavioral tasks designed to measure either capacity. Human subjects performed two tasks, a learning and generalization task (acquired equivalence) which involves relational encoding and depends on the hippocampus; and a sequential RL task that could be solved by either a model-based or model-free strategy. We assessed the correlation between subjects' use of flexible, relational memory, as measured by generalization in the acquired equivalence task, and their differential reliance on either RL strategy in the decision task. We observed a significant positive relationship between generalization and model-based, but not model-free, choice strategies. These results are consistent with the hypothesis that model-based RL, like acquired equivalence, relies on a more general-purpose relational memory system.

1. Introduction

There can be multiple paths to a decision. For example, as we learn by trial and error about the value of different options, we can select among them based simply on how much reward has previously followed each of them, or instead flexibly reevaluate them in the moment by taking into account their particular expected consequences and the current value of those consequences. The latter strategy allows us to choose flexibly if our current needs or tastes have changed: for instance, if we progress from thirst to hunger. Two distinct classes of control systems developed in the engineering literature, called model-free and model-based reinforcement learning (RL), describe these computationally and representationally different approaches to value-based decision making (Sutton & Barto, 1998).

---

* Corresponding author. Address: 4 Washington Pl, Room 873B, New York, NY 10003, United States.
Q1 *E-mail address:* bradley.doll@nyu.edu (B.B. Doll).

http://dx.doi.org/10.1016/j.nlm.2014.04.014
1074-7427/

Much evidence in both humans and animals supports the idea that the brain implements both of these approaches (Doll, Simon, & Daw, 2012). In particular, model-free RL theories (Montague, Dayan, & Sejnowski, 1996) maintain an estimate of the net reward previously received following each action, updating it in light of each experience using a reward prediction error signal. These theories explain the behavioral phenomena of habits (Daw, Niv, & Dayan, 2005)—inflexible response tendencies that often arise after overtraining (Adams, 1982)—and, neurally, offer the predominant computational account of the reward prediction error-like phasic responses of dopamine neurons and of similar signals in human fMRI at striatal dopamine targets (Glimcher, 2011). The reward prediction error signal, the difference in reward received and reward expected, is the computational core of model-free RL that drives increases in action values following rewards and decreases following punishments. The phasic bursts and dips of midbrain dopamine neurons following rewards and punishments mirror this teaching signal (Montague et al., 1996). Dopaminergic projections to striatum modulate activation and plasticity (Reynolds & Wickens, 2002) of corticostriatal synapses, driving reward and punishment learning (Hikida, Kimura, Wada, Funabiki, & Nakanishi, 2010). Optogenetic work indicates a causal role for this signaling pathway in reward learning, consistent with the predictions of model-free RL (Steinberg et al., 2013; Tsai et al., 2009). Despite the success of this theory, model-free RL cannot explain more flexible, goal-directed actions that have been demonstrated experimentally using tasks such as latent learning or reward devaluation (Adams, 1982; Tolman & Honzik, 1930), nor the neural correlates of these behavioral effects (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Glascher, Daw, Dayan, & O'Doherty, 2010; Tricomi, Balleine, & O'Doherty, 2009), nor the correlates of in-the-moment evaluation of candidate actions (Pfeiffer & Foster, 2013; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010). These latter phenomena, instead, are well explained by model-based RL. Instead of net action values, model-based algorithms learn an "internal model" of the task—how different actions lead to resulting situations or outcomes, and how those outcomes map onto value—which then can be used to evaluate candidate actions' values through a sort of mental simulation or search at decision time.

Many experiments suggest that these two learning approaches trade off differentially, depending on task features such as the amount of training (Adams, 1982; Tricomi et al., 2009), dual-task interference (Otto, Gershman, Markman, & Daw, 2013, perhaps by promoting an advantage to striatal over hippocampal function, Foerde, Knowlton, & Poldrack, 2006), pharmacological interventions (Dickinson, Smith, & Mirenowicz, 2000; Wunderlich, Smittenaar, & Dolan, 2012) or brain lesions affecting areas associated with either system (Yin, Knowlton, & Balleine, 2004; Yin, Ostlund, Knowlton, & Balleine, 2005). Further, these learning approaches also differ spontaneously across individuals (Skatova, Chan, & Daw, 2013). All these results suggest that these two sorts of learning rely on neurally and cognitively dissociable systems (though interacting and sharing some common mechanisms, e.g. Daw et al. (2011)). However, although the cognitive and neural mechanisms supporting model-free RL are reasonably well understood, those supporting model-based RL are currently much less clear.

This profile of relatively better and worse understanding about the two systems is complementary to that for a similar dichotomy in another area of research, memory. Decades of work in cognitive neuroscience concerns the brain's multiple systems for memory. Traditionally, the study of memory systems has focused on a distinction between procedural and declarative memory, which are thought to differ in the kinds of representations they form, the contexts in which they are elicited, and the neural systems that support them (e.g. Squire, 1992; Gabrieli, 1998; Knowlton, Mangels, & Squire, 1996). The distinction between memory systems was initially dramatically supported by observations that declarative memory was impaired in patients with amnesia due to hippocampal damage, while procedural memory was relatively spared in the same patients (Corkin, 1968). This effect has subsequently been demonstrated in amnestics with varying etiologies in numerous incrementally acquired, feedback-driven procedural learning tasks, such as mirror-reading (Cohen & Squire, 1980; Gabrieli, Corkin, Mickel, & Growdon, 1993), the pursuit-rotor (Heindel, Salmon, Shults, Walicke, & Butters, 1989), and weather prediction tasks (Knowlton et al., 1996). In contrast, degeneration of the nigrostriatal dopamine system in Parkinson's disease and loss of striatal integrity in Huntington's disease impairs procedural memory, but leaves declarative memory relatively intact (Heindel et al., 1989; Knowlton et al., 1996; Martone, Butters, Payne, Becker, & Sax, 1984). The dissociable behavioral and neural characteristics of these memory systems are now well-described—procedural memory is thought to be inflexible, implicit, incremental, and reliant on striatum, whereas declarative memories are more flexible, relational, possibly subject to conscious access, and reliant on hippocampus.

Thus, in many ways, this distinction appears to be closely related to that between model-free and model-based RL, with model-free RL corresponding to procedural memory (Knowlton et al., 1996), and model-based corresponding to declarative memory (Daw & Shohamy, 2008; Dickinson, 1980). The computational and neural mechanisms formalized by model-free theories and tasks also explain classic procedural learning tasks, which feature incremental learning from trial-to-trial feedback, and implicate the striatum and its dopaminergic inputs (Knowlton et al., 1996; Shohamy et al., 2004). Evidence that distraction by cognitive load disrupts model-based, but not model-free RL (Otto, Gershman, et al., 2013) is mirrored by evidence that such interference disrupts declarative, but not procedural task learning (Foerde et al., 2006). Better understanding of the relationship between decision making and memory systems has the potential to shed light on both areas, in particular because in memory, much is known about the brain's systems for declarative memory, which might provide a crucial relational encoding mechanism underlying model-based RL. (Conversely, our relatively strong knowledge of the brain's mechanisms for model-free RL may fill in many gaps in our understanding of procedural memory.) Thus, in the present study we aimed to examine evidence for these correspondences by studying the relationship between tasks from the memory and decision literature.

A key feature of the hippocampal memory system, which is particularly relevant to model-based RL is the encoding of relations—associations between multiple stimuli or events (Cohen & Eichenbaum, 1993). A hallmark of relational memories that suggests a parallel to model-based RL is that multiple, previously learned associations can be combined in novel ways. A family of classic memory tasks including acquired equivalence (Honey & Hall, 1989), associative inference (Bunsey & Eichenbaum, 1996), and transitive inference (Davis, 1992) assess this feature of relational memory. In these tasks, subjects first learn sets of overlapping stimulus relationships and then generalize to or infer the relationships between never-before-seen stimulus combinations in a test phase. Neurally, successful performance of these relational memory tasks is associated with the hippocampus (Greene, Gross, Elsinger, & Rao, 2006; Myers et al., 2003; Preston, Shrager, Dudukovic, & Gabrieli, 2004; Shohamy & Wagner, 2008).

These features of relational memory suggest a correspondence to model-based RL. Unlike the response- or reward-related associations underlying model-free learning—which are clearly in the domain of procedural memory (Knowlton, Squire, & Gluck, 1994; Nissen & Bullemer, 1987)—model-based RL relies on learning a world model, that is, an arbitrary set of associations between stimuli or situations (as with a map of a spatial task), which can

Q1 Please cite this article in press as: Doll, B. B., et al. Multiple memory systems as substrates for multiple decision systems. *Neurobiology of Learning and Memory* (2014), http://dx.doi.org/10.1016/j.nlm.2014.04.014