



How Sequential Interactive Processing Within Frontostriatal Loops Supports a Continuum of Habitual to Controlled Processing

Randall C. O'Reilly^{1,2*}, Ananta Nair², Jacob L. Russin¹ and Seth A. Herd^{1,2}

¹ Computational Cognitive Neuroscience Lab, Department of Psychology, Computer Science, and Center for Neuroscience, University of California, Davis, Davis, CA, United States, ² eCortex, Inc., Boulder, CO, United States

We address the distinction between habitual/automatic vs. goal-directed/controlled behavior, from the perspective of a computational model of the frontostriatal loops. The model exhibits a continuum of behavior between these poles, as a function of the interactive dynamics among different functionally-specialized brain areas, operating iteratively over multiple sequential steps, and having multiple nested loops of similar decision making circuits. This framework blurs the lines between these traditional distinctions in many ways. For example, although habitual actions have traditionally been considered purely automatic, the outer loop must first decide to allow such habitual actions to proceed. Furthermore, because the part of the brain that generates proposed action plans is common across habitual and controlled/goal-directed behavior, the key differences are instead in how many iterations of sequential decision-making are taken, and to what extent various forms of predictive (model-based) processes are engaged. At the core of every iterative step in our model, the basal ganglia provides a “model-free” dopamine-trained Go/NoGo evaluation of the entire distributed plan/goal/evaluation/prediction state. This evaluation serves as the fulcrum of serializing otherwise parallel neural processing. Goal-based inputs to the nominally model-free basal ganglia system are among several ways in which the popular model-based vs. model-free framework may not capture the most behaviorally and neurally relevant distinctions in this area.

Keywords: habits, goals, controlled processing, automatic processing, computational modeling, frontal cortex, basal ganglia

OPEN ACCESS

Edited by:

John A. Bargh,
Yale University, United States

Reviewed by:

Ion Juvina,
Wright State University, United States
Terrence C. Stewart,
University of Waterloo, Canada

*Correspondence:

Randall C. O'Reilly
oreilly@ucdavis.edu

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 26 June 2019

Accepted: 18 February 2020

Published: 10 March 2020

Citation:

O'Reilly RC, Nair A, Russin JL and Herd SA (2020) How Sequential Interactive Processing Within Frontostriatal Loops Supports a Continuum of Habitual to Controlled Processing. *Front. Psychol.* 11:380. doi: 10.3389/fpsyg.2020.00380

INTRODUCTION

Since its inception, the field of psychology has been fascinated by the distinction between two types of behavior, one that leads us to act relatively automatically, according to well-worn habits, and another that allows us to act with intent and deliberation (James, 1890; Thorndike, 1911; Hull, 1943; Tolman, 1948). These two classes of thought and action have been referred to by several different sets of terminologies, each with slightly varying definitions, which has sown some confusion in the literature (Hassin et al., 2009; Kool et al., 2018; Miller et al., 2018, 2019). Historically, the first terminology applied to this intuitive distinction was *stimulus-response* vs. *cognitive-map guided*

(Thorndike, 1911; Tolman, 1948). This distinction was later replaced by *habitual vs. goal-directed behavior* (Tolman, 1948; Balleine and Dickinson, 1998; Dickinson and Balleine, 2002; Killcross and Coutureau, 2003; Balleine, 2005; Yin and Knowlton, 2006; Tricomi et al., 2009), which co-existed alongside *automatic vs. controlled* processing (Shiffrin and Schneider, 1977; Cohen et al., 1990; Miller and Cohen, 2001). More recently, a good deal of work has been directed at the distinction between *model-free* and *model-based* reinforcement learning (Sutton and Barto, 1998; Doya, 1999; Doya et al., 2002; Daw et al., 2005).

In this paper, we attempt to clarify the relationships among these terminological distinctions through the lens of a computational model of the underlying brain mechanisms. This model builds on detailed neural recording data available on animal action-selection. One of the major conclusions from this model is that these apparently distinct types of behavior may be manifestations of a core underlying neural system, which evaluates the relative cost/benefit tradeoffs of engaging in more time-consuming, deliberative processing using the same basic mechanisms that drive all the other behavioral decisions that an organism must make. Furthermore, we argue that the neural pathways that support the habitual stimulus-response level behavior are actually an integral part of the same system that supports deliberative, controlled processing. Thus, this framework provides a unified view of action selection and decision making from the most basic habitual level up to the most complex, difficult decisions that people face. In our theory, Type 2 (deliberative) decisions are essentially composed of many Type 1 (automatic) decisions. Thus, it offers a mechanistic explanation of the proposed continuum between them (Melnikoff and Bargh, 2018).

Goal-Driven/Controlled vs. Habitual/Automatic

We first establish some common ground by attempting to define a consensus view about the closely-related distinctions between *goal-driven vs. habitual*, and *controlled vs. automatic* processing. Of the two, controlled vs. automatic (Shiffrin and Schneider, 1977) is perhaps more clearly defined, by virtue of a history of computational models based on the idea that the prefrontal cortex (PFC) supports controlled processing by maintaining active working memory representations that drive top-down biasing of processing elsewhere in the brain (Cohen et al., 1990; Miller and Cohen, 2001; Herd et al., 2006; O'Reilly, 2006; O'Reilly and Frank, 2006). Cognitive control is needed to support novel, difficult, complex tasks, e.g., to overcome prepotent (i.e., habitual) response pathways in the widely-studied Stroop task. As a task or stimulus-response pathway becomes more strongly practiced, behavior becomes more automatic and free from the need for this top-down biasing support. Thus, automatic and habitual are closely related terms. The connection between goal-driven and controlled processing is somewhat less exact, as one could imagine behaving according to goals that do not require significant cognitive control (Bargh, 1989), and potentially even exerting cognitive control in the absence of clear

goal-driven motivations. Sustained active neural firing of goal-like representations, that can exert an ongoing biasing effect on behavior, is perhaps a more direct mechanistic connection between the two.

Phenomenologically, habitual behavior is typically characterized as being relatively insensitive to the current reward value of actions, and not as strongly under the control of active, conscious goal engagement (Wood and Runger, 2016). On the other hand, it remains a challenge to consider the nature of real-world behaviors that are characterized as habits, as they often involve extended sequences of actions coordinated over reasonably long periods of time (e.g., driving home from work, making coffee, etc.) – these do not seem to be entirely unconscious activities devoid of any cognitive control influences, or contextual sensitivity (Cushman and Morris, 2015). Furthermore, how can it be that subtle, unconscious factors can sometimes strongly shape our overt behavior (Bargh, 2006; Huang and Bargh, 2014)?

Our general answer to these questions, as captured in our computational modeling framework, is that both habits and more controlled, goal-driven behaviors emerge from a shared neural system, and both operate within a common *outer-loop* of overall cognitive control that pervasively shapes and modulates the nature of processing performed in the *inner-loops* associated with specific task performance. This is similar to the hierarchical control framework of Cushman and Morris (2015), except that we postulate a sequential, temporal organization of decision making and control, where the same neural systems iteratively process multiple steps over time, including periodic revisiting of the broader context and goals that we refer to as the *outer-loop*. Thus, habits only drive behavior when permitted by this outer-loop of cognitive control, and indeed the actual unfolding of behavior over time is usually at least somewhat coordinated by the outer-loop. Furthermore, as we'll elaborate below, a crucial factor across all behavior in our framework is a so-called *Proposer* system that integrates many different factors in a *parallel-constraint-satisfaction* system to derive a proposed plan of action at any point in time – the properties of this system may explain how unconscious factors can come to influence overt behavior in the course of solving the *reduction problem* of choosing one plan among many alternatives (Bargh, 2006; Huang and Bargh, 2014).

The Model-Free vs. Model-Based Dichotomy

Within the above context, how does the *model-based vs. model-free* (MBMF) framework fit in? This framework has engaged new enthusiasm by offering the promise of a more formal, precise definition of the relevant processes, and by leveraging the direct connection between reinforcement learning principles and properties of the midbrain dopamine system (Montague et al., 1996; Schultz, 2013). Specifically, the model-free component is typically defined as relying on learned, compiled estimates of future reward associated with a given current state (or potential actions to be taken in that state), which have been trained via phasic dopamine-like temporal difference signals, as in the classic

TD and Q-learning reinforcement learning frameworks (Sutton and Barto, 1998). By contrast, the model-based system adds an internal model that can simulate the evolution of the state of the world over multiple iterations, so that action selection can be based on those predicted states. As such, the model-free system is considered to be relatively inflexible to changes in the reward function, including changes resulting from internal state (e.g., not being hungry at the moment), whereas the model-based system can dynamically adjust its predictions based on goal changes and other changes, and is thus more flexible.

Thus, it is this key difference in the relative flexibility of the two systems that maps onto the existing notions of goal-driven vs. habitual behavior. However, there are various other aspects of the MBMF framework which map less well, creating significant confusion when people intend to characterize the goal-driven vs. habitual distinction, but using the MBMF terminology. At a very basic level, there is no principled reason why a model-free system should not have access, as inputs, to internal drive and goal states in addition to external environmental states. If it does, its behavior can also be goal-directed, and sensitive to internal bodily states such as hunger. In addition, model-based is not synonymous with goal-directed, as model-based is defined specifically in terms of models of the external environment. In our framework, a model-free-like system indeed receives internal state and goal inputs, and thus participates in goal-directed behavior. This illustrates an important mismatch between these two terminologies, which are often taken to be interchangeable. More generally, standard reinforcement-learning paradigms tend not to incorporate a significant goal-driven component, and instead generally assume a single overriding goal of maximizing a scalar-valued reward, which is delivered to an entirely externally-motivated agent (O'Reilly et al., 2014). Thus, aside from a few more recent examples (Berseth et al., 2018), standard reinforcement-learning models are not particularly well-suited for describing goal-driven processing in the first place.

Recent reviews by Miller et al. (2018, 2019) point out the following additional issues with the MBMF terminology. First, it is problematic that the model-free system relies on learned *value* estimates to drive action selection, whereas most existing data indicates that habitual behavior is specifically more insensitive to reward value (Wood and Runger, 2016). Second, the neural substrates associated with MBMF mechanisms are largely overlapping and hard to disentangle, involving the dopaminergic system, the basal ganglia (BG), and the prefrontal cortex (PFC). Whereas the BG was traditionally viewed as being primarily a habit-based motor area (e.g., Miller, 1981; Mishkin et al., 1984; Squire and Zola-Morgan, 1991; Squire, 1992; Packard and Knowlton, 2002) more recent evidence and theorizing suggest that, with the exception of the dorsal-lateral striatum, most of the BG is more clearly involved with non-habitual behavior and deliberative, controlled cognition in novel and challenging tasks (Pasupathy and Miller, 2005; Samejima et al., 2005; Yin et al., 2005; Balleine et al., 2007; Seger and Spiering, 2011; Pauli et al., 2012). Many authors nevertheless continue to assume the simple association of model-free with the BG, in keeping with the traditional habit-based ideas.

Furthermore, while the MBMF distinction is often considered to be dichotomous, more recent work has explored various combinations of these aspects to deal with the computational intractability of full model-based control, further blurring the lines between them (Pezzulo et al., 2013; Cushman and Morris, 2015). Likewise, there are many ways of approximating aspects of model-based predictions of future outcomes that may not fit the formal definition of iterative model-state updating, e.g., using predictive learning in the *successor-representation* framework (Dayan, 1993; Littman and Sutton, 2002; Momennejad et al., 2017; Russek et al., 2017; Gershman, 2018). This may be considered acceptable if the distinction is just that the model-free system has absolutely no model-like element, and the model-based system has any kind of approximation of a world model (Daw and Dayan, 2014), but this may end up straining the value of the distinction. For instance, a successor-representation model is otherwise quite similar to a standard model-free system, but it does use information about outcomes (although they do not usually explicitly predict an outcome).

The above considerations led Miller et al. (2018, 2019) to conclude that MBMF are both aspects of the goal-based, controlled-processing system, based on the prefrontal cortex/basal ganglia/dopamine circuits in the brain, while habitual, automatic processing is supported by an entirely separate system governed by a Hebbian, associative form of learning that strengthens with repetition.

Overview of the Paper

In the remainder of this paper, we present an alternative framework based on computational models of the basal ganglia/prefrontal cortex/dopamine system, which is consistent with the overall critique of MBMF by Miller et al. (2018, 2019), and provides a specific set of ways in which these brain systems can support a continuum of goal-directed, model-based forms of decision making and action selection. The original controlled vs. automatic distinction has always incorporated this notion that these are two poles along a continuum. Our framework goes further in describing how model-based and model-free elements interact in various ways and to varying degrees to provide a rich and multi-dimensional continuum of controlled, goal-driven cognition, which also supports varying degrees and shades of habitual or automatic elements.

This framework contrasts with several others that posit strongly dichotomous and internally homogenous habitual vs. goal-driven pathways, followed by an arbiter system that decides between the two (e.g., Daw et al., 2005; Miller et al., 2019). Instead, we propose that an outer-loop of goal-driven, but model-free, processing is itself essentially an arbiter of how much time and effort to invest in any given decision-making process. It controls the degree of engagement of a broader toolkit of basic decision-making computations to be deployed, as a function of their relative tradeoffs (c.f., Pezzulo et al., 2013). In particular, it controls whether to perform additional steps of predictive modeling down each given branch of the state-space model.

We also address a critical phenomenon for any model in this domain, which is the nature of the transition from controlled to automatic processing (Cohen et al., 1990; Gray et al., 1997;

Hikosaka and Isoda, 2010). Behaviorally, this transition occurs gradually over time and appears to reflect something like the strengthening of habit representations, which offer advantages in terms of speed, resistance to distraction, and the ability to do more in parallel, at the cost of flexibility and sensitivity to current goals – i.e., the fundamental underlying tradeoffs along this dimension. However, due to the multi-component nature of our goal-driven model, there are also various ways in which learning within this system can change these relative tradeoffs, leading to a richer picture of this process of habit formation.

THE PROPOSER-PREDICTOR-ACTOR-CRITIC MODEL

Our theoretical framework has been specified as a neural network model in the Leabra framework (O'Reilly, 1998; O'Reilly and Munakata, 2000; O'Reilly et al., 2016).

The *Proposer-Predictor-Actor-Critic* (PPAC) model (Figures 1, 2; Herd et al., 2019) leverages the prototypical loops descending from all areas of frontal cortex through the basal ganglia and converging back to modulate the function of matching areas of frontal cortex (Alexander et al., 1986; Haber, 2010, 2017; Sallet et al., 2013). Functionally, these BG/PFC loops support the ability to selectively activate and maintain neural activity (i.e., working memory) in the service of supporting top-down control representations (Miller and Cohen, 2001; Frank and O'Reilly, 2006; Herd et al., 2006; O'Reilly, 2006). As such, this system is critical for controlled, goal-driven processing. The PPAC model includes an important distinction among the nature of the cortical input representations into the BG: proposed actions vs. predicted outcomes. Critically, complex decision-making unfolds sequentially across multiple iterations in the model, each of which involves parallel operations across these circuits (i.e., a *serial-parallel* model, in which parallel computations are iterated serially).

In this theory, complex decision-making consists of a series of selections of internal “actions,” each of which consists of an update to working memory and/or episodic memory. Selecting a move in chess or choosing a plane ticket to purchase may each require a large number of belief updates (like “too expensive to fly direct in the afternoon”) and the selection of several new mid-level plans (like “try to threaten a more valuable piece instead of defending the knight”). Each of these can be stored in active memory, which executes controlled processing (by exerting top-down biasing of processing (Cohen et al., 1990; Miller and Cohen, 2001; Herd et al., 2006). Maintaining each plan or belief in working memory can also create an episodic memory trace for later recall and re-use. Our theory holds that each such representation is selected for maintenance (and therefore plan execution) much as motor representations are selected, by distinct but computationally and structurally analogous circuits.

Our theory expands on existing work on action selection in the basal ganglia, and addresses the contributions of cortex to this process. As such, we adopt the terminology of an actor-critic reinforcement learning architecture (Sutton and Barto, 1998; O'Reilly and Frank, 2006) to describe the computational roles

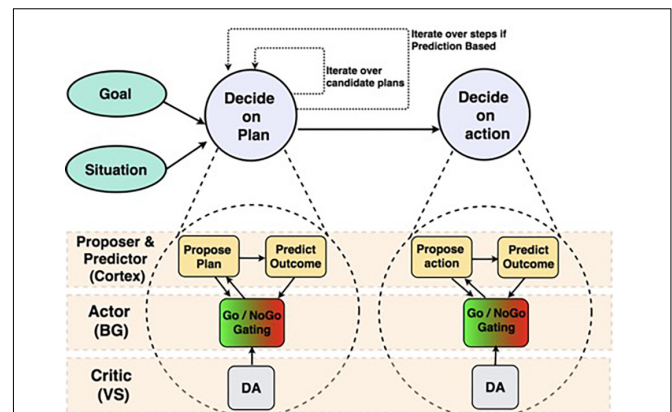


FIGURE 1 | Structure of Proposer-Predictor-Actor-Critic architecture (Herd et al., 2019) across frontal cortex and subcortical areas. We depict two parallel circuits with a hierarchical relationship. The top is a broad functional diagram, emphasizing the serially iterative and hierarchical nature of our proposed decision-making process. The bottom expands those functions, and identifies the brain areas that perform each function.

of basal ganglia and the dopamine system. The basal ganglia functions as an *Actor* that decides which action to take (or in our extended model, which plan to pursue). The dopamine release system, including amygdala, ventral striatum, and related areas, serves as a *Critic* by gauging the success of each action relative to expectations. Phasic dopamine release from this critic system serves as a reward prediction error learning signal for the basal ganglia actor system.

To this existing computational/biological theory we add two new computations, each made by participating regions of cortex. The first is a *Proposer* component. This system takes information about the current situation as input, and produces a single candidate plan representation. This proposer functional role may be less important for laboratory tasks, since they usually have a small set of actions (e.g., levers, yes/no responses), which can be learned thoroughly enough to process all options in parallel routes through the basal ganglia (e.g., Collins and Frank, 2014). However, dealing with unique real-world situations requires coming up with a potential approach before evaluating outcomes (e.g., different plausible routes in a trip planning context). This proposer system could use computations characterized as model-free, stimulus-response, constraint satisfaction, or model-based, depending on the complexity of the situation.

The other cortical addition is a *Predictor* component, which predicts the likely outcome of each proposed plan. In our model as currently implemented, this prediction always took place in two steps: predicting an “Outcome,” and from that outcome, predicting a “Result” or potential reward. We think that this type of prediction is actually performed by a variety of brain systems, using a variable number of steps for different types of decisions; but for the present purposes, it is adequate to simply think of this component as producing a prediction of an outcome by any means. This system’s computation is thus very much “model-based,” according to that terminology.

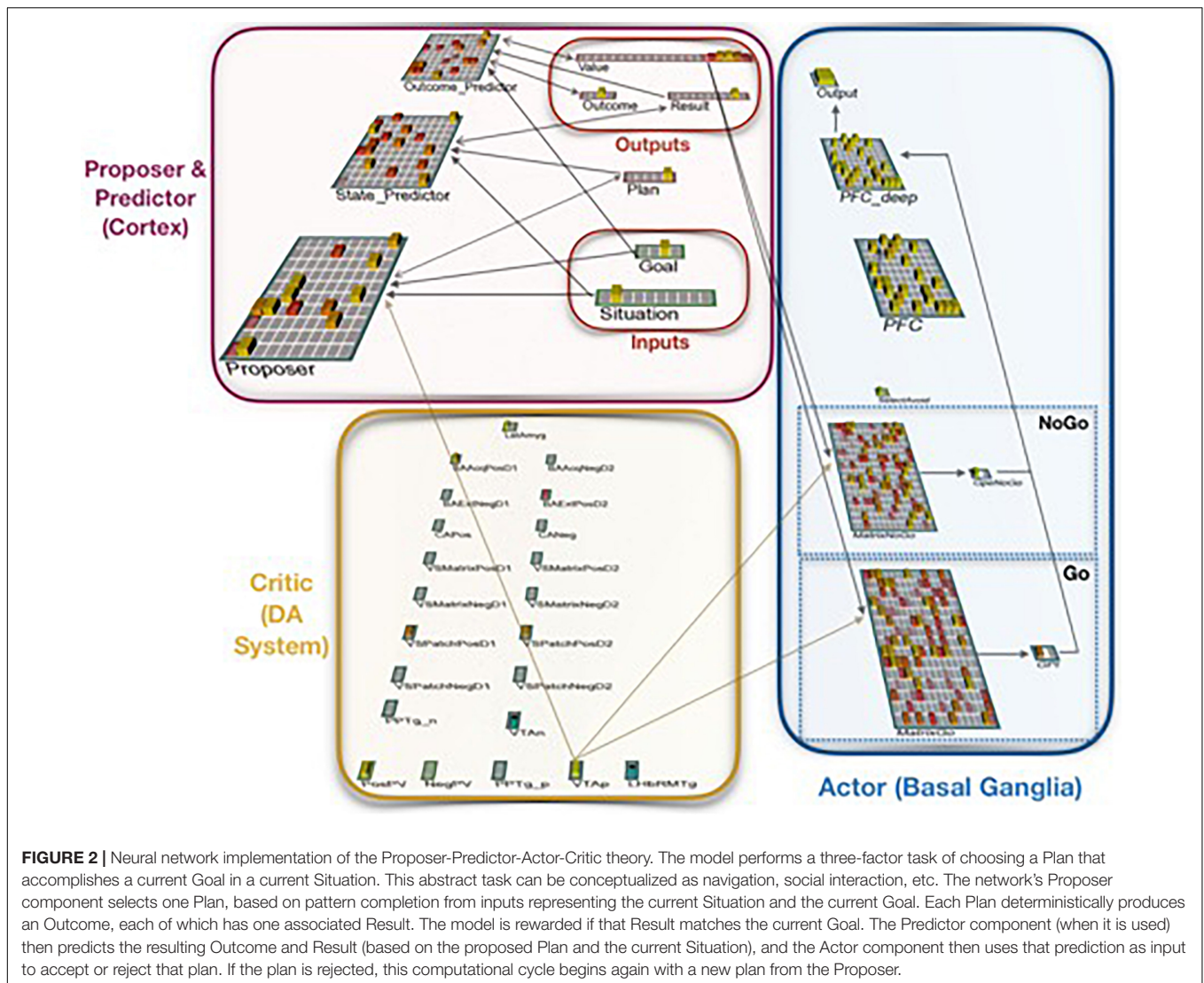


FIGURE 2 | Neural network implementation of the Proposer-Predictor-Actor-Critic theory. The model performs a three-factor task of choosing a Plan that accomplishes a current Goal in a current Situation. This abstract task can be conceptualized as navigation, social interaction, etc. The network's Proposer component selects one Plan, based on pattern completion from inputs representing the current Situation and the current Goal. Each Plan deterministically produces an Outcome, each of which has one associated Result. The model is rewarded if that Result matches the current Goal. The Predictor component (when it is used) then predicts the resulting Outcome and Result (based on the proposed Plan and the current Situation), and the Actor component then uses that prediction as input to accept or reject that plan. If the plan is rejected, this computational cycle begins again with a new plan from the Proposer.

In our system, the Actor uses the predicted outcome (when available) of the proposed plan to either accept or reject that plan. Having this specific outcome prediction greatly simplifies the computational task of the actor component; it need simply accept plans that are predicted to have rewarding outcomes, and reject those that do not. If the proposed plan is rejected, the Proposer component makes a new, different proposal, a new prediction is made by the Predictor, and the Actor again decides to accept or reject that newly proposed plan. This operation proceeds serially until a candidate plan is selected. The serial, one-at-a-time plan consideration is slow, but computationally helpful in making an accurate prediction of outcomes in novel, poorly learned domains. It allows the full power of the cortex to be directed toward each prediction, and avoids binding problems, as we address further in the Discussion section.

This computational approach can attack complex problem spaces by sequentializing a complex decision into many sub-decisions, and allowing the actor component to accept or reject each proposed sub-plan or sub-conclusion. We propose

that our ability to sequentialize a problem into sub-steps and make a binary decision for each is the source of humans' remarkable cognitive abilities relative to other animals. This method of simplification may, however, have particular inherent weaknesses that explain some of humans' notable cognitive biases.

Continuum of Controlled/Goal-Directed vs. Automatic/Habitual

Due to its sequential, hierarchical and multi-component nature, the model provides a mechanistic basis for a continuum of controlled/goal-directed vs. automatic/habitual behavior. At every sequential step, there is the potential for an outer-loop decision about what overall strategy to employ, e.g., whether to engage in further prediction, or iterate to another proposed plan, etc. Within that outer loop, there are more specific decisions regarding what factors to focus on, such as which branches to pursue in prediction, etc.

In cases of high urgency or low stakes, all of that complexity could be elided in favor of a quick thumbs-up (Go gating decision) from the Actor to the Proposer's initial suggestion. This optimization for speed could be created by reinforcement learning in the basal ganglia, with inputs that capture timing and relevant time pressures. We suggest that this may represent the majority of habitual or automatic responding – a fast path through the very same circuits, typically at the lower levels of the abstraction hierarchy (e.g., involving supplementary motor areas and the dorsolateral striatum). Thus, consistent with the continuum perspective, and a surprising difficulty in finding explicit claims and data about what neural substrates uniquely support habitual behavior (e.g., Wickens et al., 2007; Seger and Spiering, 2011), there may be no separate neural substrate associated with habitual behavior – it is just the simplest and fastest mode of processing through the entire decision-making apparatus.

If this is the case, then it would seem to challenge the various attempts to establish strong dichotomies between e.g., model-free vs. model-based, or even value-based vs. value-free or belief-based vs. belief-free (Miller et al., 2018, 2019). In short, even habitual behavior depends on a (usually implicit) decision to not engage in a more controlled form of behavior, and that decision likely depends on assessments of the relevant “stakes” (values or utilities) in the current context, and the estimated cost/benefit tradeoffs in engaging in more effortful levels of control (Pezzulo et al., 2013).

Thus, estimated value is always in play, even in the context of habitual behavior. To reconcile this idea with the finding that habitual behaviors are relatively insensitive to changes in reward, we would need to determine the relative cost/benefit tradeoff estimates associated with the alternative options that might have been taken instead of performing the habitual response. Certainly, if the habitual response would lead to imminent severe harm, and this was obvious to the individual, then we would expect them not to engage in it. Typically when clearly erroneous habitual responding occurs in the real world, it can be traced to a lack of attention being paid to the relevant factors, likely resulting from prior decisions to allocate that attention elsewhere. In other words, taken literally, a purely habitual response presumes that the person is otherwise somewhat of a zombie. Instead, we suggest, consistent with others (e.g., Cushman and Morris, 2015) that habitual responses occur within a broader context (i.e., the outer-loop) of at least some level of cognitive control.

The Model-Free Actor in the Loop

A central feature of our model is that the basal ganglia Actor system provides a value-based final Go/NoGo decision, even (and perhaps especially) under controlled, deliberative situations. The Actor fits the classic description of a model-free reinforcement learning system, and thus our framework says that there is an important model-free component to even high-level goal-driven and controlled behavior. This is consistent with a similar claim in the hierarchical model of Cushman and Morris (2015) and with their more recent experimental results (Cushman et al., this issue). Thus, whether one wants to call this Actor “model-free” or not, even when it receives all manner of

highly-processed goal, internal state, and prediction inputs from the cortex, further challenges the utility of this terminology. Furthermore, as we noted above, the availability of predicted outcome representations from the Predictor component can make the Actor's job very simple, and yet likely much more effective than a typical model-free system.

The Central Role of the Proposer

The function of the Proposer is particularly central to our overall framework, as it serves as the starting point for any action/plan initiation process. As noted, we think it functions through parallel, constraint-satisfaction processing to integrate a large number of different constraints, cues, and other contextual information to arrive at a plausible plan of action in a given situation (O'Reilly et al., 2014). It is precisely through this dynamic integration process that otherwise subtle, unconscious factors may be able to have measurable influences over our behavior (Bargh, 2006; Huang and Bargh, 2014). In addition, this property of the proposer enables even habit-based behavior to be somewhat flexible and capable of incorporating novel constraints from the current environmental state – even habitual actions are not purely ballistic and “robotic” in nature (Cushman and Morris, 2015; Hardwick et al., 2019).

Furthermore, as we'll see next, the incremental shaping of these Proposer representations over the course of learning plays a critical role in the automatization and habitization of behavior. Indeed, as the Proposer gets better and better at generating effective plans for increasingly well-known contexts, the Actor learns to essentially rubber-stamp these plans, thus resulting in fast, efficient habitual behavior. This happens through reinforcement learning shaping the weights from cortex to the basal ganglia Actor system; as the Actor sees more positive, rewarding examples, it becomes more biased toward a Go response. Along with its importance in habitual behavior, the Proposer component is also essential for coming up with plans in novel, challenging situations requiring controlled processing. Thus we argue that these functional distinctions may not have clear corresponding anatomical distinctions: the basal ganglia, Actor component is involved in all types of decisions, and that different areas of cortex may be recruited to play roles as Proposer, Predictor, and even to add more highly-processed inputs to model-free value predictions (Herd et al., 2019).

Transition From Slow and Controlled to Fast and Automatic Processing

One of the main results from our computational model (Herd et al., 2019) is shown in **Figure 3**, where the Proposer component gradually learned to choose a Plan appropriate for the current situation and goal. Initially, without relevant domain knowledge, the Proposer generates plans essentially at random, and a larger number of iterations are required to arrive at a Plan that the Actor approves of. Over the course of learning, the more appropriate initial plans generated by the Proposer reduces the number of iterations required, and thus the overall model gradually transitions from a more serial, iterative mode of processing to a more parallel mode of processing dominated

by the parallel constraint-satisfaction dynamic in generating plans in the Proposer system. This illustrates a continuum of habitization occurring over learning within the same overall system. Furthermore, the Proposer was able to learn only when the remaining systems chose to pursue a given plan; its learning was thus guided by the other systems, including the Predictor component.

Our initial model does not include the outer-loop ability to select which decision-making processes to engage in, so it did not have the ability to further optimize decision making by not engaging the Predictor at all, which would have resulted in even greater speedup, and corresponds with a more purely habitual response mode. We are currently working on a version of the model with this functionality.

Goal-Directed Behavior From a Model-Free System

In our model, the input to the Proposer system includes information about goals, so the behavior produced by this system qualifies as goal-directed, despite the relatively simple computations. Most computational work on model-free reinforcement learning systems addresses systems that do not include current goals as inputs. Those systems can only produce habitual behavior. However, there does not appear to be any strong justification for this assumption, and it seems more reasonable (as well as empirically justified) to assume that the relevant systems in the mammalian brain have access to a variety of useful information, including current goals. Indeed, there has been some discussion of goal-directed habits in other literature (Verplanken and Aarts, 1999; Aarts and Dijksterhuis, 2000).

When we assessed the accuracy with which the Proposer component produced a Plan which accomplished the current Goal (with Situation and Goal chosen at random from ten and four possibilities, respectively), we observed that this component displayed goal-directed behavior by matching Plans to Goals at an

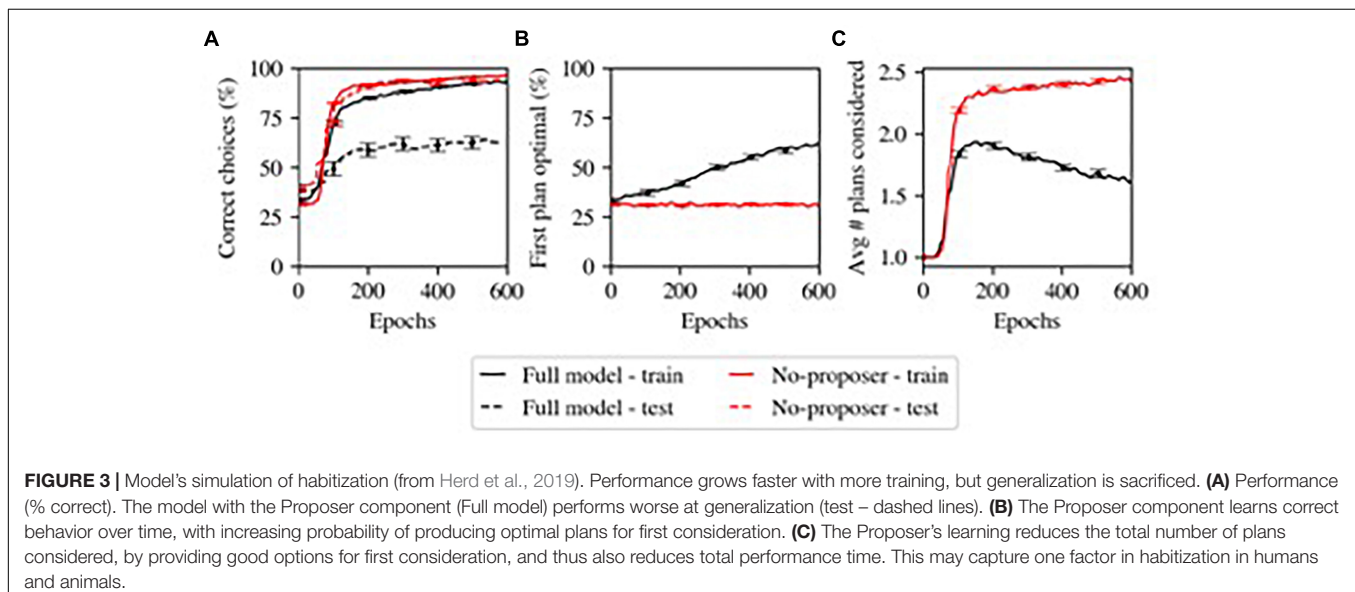
above-chance level, but learned slowly (Figure 3). This matches the slow transition from controlled to automatic processing (Gray et al., 1997) (note that we did not optimize parameters for Proposer learning in this task; some other parameterizations did produce better and faster learning).

Thus, our model illustrates one case in which goal-directed behavior results from thoroughly model-free computations.

Serial Processing Enables Coherent Predictions

A key advantage of the serial evaluation of different proposed plans in our model is that it allows many different brain areas to contribute to the evaluation process, without suffering from the *binding problem* that would otherwise arise from an attempt to evaluate multiple options in parallel. For example, if two options are considered together, and another brain area generates an activation associated with a prediction of difficulty, while another activates a prediction of relative ease, how do we know which prediction goes with which option?

This is analogous to the binding problem in visual search, where serial processing has also been implicated as a solution (Treisman and Gelade, 1980; Wolfe, 2003). For example, people cannot identify in parallel whether a display contains a particular *conjunction* of features (e.g., a red X among green Xs and red Os), whereas they can identify separable features in parallel (just Xs among Os, or just red among green). Likewise, the conjunction of options and their predicted consequences at many different levels in the brain, which likely depends on the current internal and external state, can be much more coherently evaluated by considering options one at a time. Furthermore, this serialization of the processing enables the *same* predictive and evaluative neural representations to be re-used across different situations and contexts, thus facilitating the transfer of knowledge to novel situations. In short, more complex model-based, predictive forms of control must involve serial processing mechanisms.



However, there are costs associated with serial processing, not only in terms of time, but also in terms of the coordination and control required to organize the serial processing itself. In addition, evaluating any one option relative to the predicted properties of other options requires some form of maintenance and comparison operations across these predictions, placing demands on working memory and other limited cognitive resources. Nevertheless, there are strong serial-order effects on decision-making, which such a serial model can naturally account for, so future modeling work will need to address these challenges in order to better address the complexities of these phenomena.

In summary, our sequential, integrated, systems-based approach provides some potentially novel perspectives on central questions about the nature of controlled/goal-driven vs. automatic/habitual behavior.

DISCUSSION

We have presented a computational systems-neuroscience approach to understanding the dynamics of decision making and action selection, which suggests that the classical dichotomy between habitual/automatic vs. goal-directed/controlled processing can be understood as different modes of functioning within a unitary system, operating fundamentally in a serial manner. The serial nature of the processing affords a natural incrementality to the continuum between these modes of processing – as the system iterates longer and engages more elaborated predictive and evaluative forms of processing, it shades more toward the goal-driven, controlled-processing end of the spectrum. By contrast, there is a fast track through the system where a proposed plan of action is derived rapidly through parallel constraint-satisfaction processing, which is then quickly approved by the basal ganglia Go/NoGo system – this corresponds to the habitual end of the spectrum. However, even this habitual level of behavior is contingent on an outer-loop of decision making that has established relevant thresholds and control parameters to enable the fast-track to be taken in the first place. Thus, habitual behavior still operates within an at-least minimally controlled context, in situations where the overall benefits of so behaving make sense compared to investing greater levels of control.

This framework contrasts with the dual-pathway model proposed by Miller et al. (2019) and similar models which suggest that habitual and controlled, goal-driven processing are subserved by parallel pathways that compete via an Arbiter system for control over behavior. It also contrasts with other models having a similar overall structure, but which use model-free and model-based components that likewise require an Arbiter system (e.g., Daw and Dayan, 2014). The framing of the interrelationship of habitual and controlled processing provided by Cushman and Morris (2015) is much more consistent with our framework, but further work is required to establish more detailed comparisons between their implemented models and our model. Likewise, the Pezzulo et al. (2013) model shares the central idea that model-based predictive mechanisms are only engaged when they yield additional value, and we will be working to relate

their computational-level mechanisms to the more biologically based framework we have developed.

Behaviorally, there are several important predictions that our model makes, which can be tested empirically. For example, consistent with a great deal of theory as well as folk psychology, we argue that habitual control is only enabled in either low-stakes or highly urgent situations. How does this outer loop of control interact with the various behavioral paradigms that have established the relative value-insensitivity of habitual behavior (Wood and Runger, 2016)? Can our model account for both this value-insensitivity but also the cases where relevant expected reward values shift the system to more controlled, goal-driven behavior? What behavioral paradigms can effectively test such dynamics? One recent result provides a nice confirmation of one of our model's core predictions: that habitization is primarily about rapid activation of a good proposed plan of action (i.e., the Proposer in our model), but there remains a final "goal-directed" process (the Actor in our model) responsible for actual action initiation (Hardwick et al., 2019).

Another fertile ground for testing the model is in the domain of serial order effects on decision-making. For example, the balloon analog risk task (Lejuez et al., 2003; White et al., 2008; Van Ravenzwaaij et al., 2011; Fukunaga et al., 2012; Fairley et al., 2019) involves making a long sequence of decisions about whether to keep pumping a simulated balloon, or cash out with a potentially sub-optimal level of reward, and it seems uniquely capable of capturing real-world individual differences in propensity toward risky behaviors (Lejuez et al., 2003; White et al., 2008). Various sources of evidence suggest that there is something about the sequential nature of this task that is critical for its real-world validity. Thus, we are actively exploring this question in terms of the serial processing present in our model. In addition, there are other well-established serial-order effects in decision making, including framing effects (Tversky and Kahneman, 1981; De Martino et al., 2006), and the widely-studied foot-in-the-door and door-in-the-face strategies (Pascual and Gueuen, 2005), which our serial model is particularly well-suited to explain.

AUTHOR CONTRIBUTIONS

All authors contributed to the development of the ideas and computational models described herein, and to the writing of the manuscript.

FUNDING

This work was supported by the ONR N00014-18-C-2067, N00014-18-1-2116, N00014-14-1-0670/N00014-16-1-2128, and NIH R01GM109996.

ACKNOWLEDGMENTS

We thank the other members of the CCN Lab for comments and discussion.

REFERENCES

- Aarts, H., and Dijksterhuis, A. (2000). Habits as knowledge structures: automaticity in goal-directed behavior. *J. Pers. Soc. Psychol.* 78:53. doi: 10.1037/0022-3514.78.1.53
- 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. doi: 10.1146/annurev.ne.09.030186.002041
- Balleine, B. W. (2005). Neural bases of food-seeking: affect, arousal and reward in corticostriatal limbic circuits. *Physiol. Behav.* 2005, 717–730. doi: 10.1016/j.physbeh.2005.08.061
- Balleine, B. W., Delgado, M. R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27, 8161–8165. doi: 10.1523/jneurosci.1554-07.2007
- Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419. doi: 10.1016/s0028-3908(98)00033-1
- Bargh, J. A. (1989). Conditional automaticity: varieties of automatic influence in social perception and cognition. *Unintended thought* 3, 51–69.
- Bargh, J. A. (2006). What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *Eur. J. Soc. Psychol.* 36, 147–168. doi: 10.1002/ejsp.336
- Berseth, G., Kyriazis, A., Zinin, I., Choi, W., and van de Panne, M. (2018). “Model-based action exploration for learning dynamic motion skills,” in *2018 IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS)*, (Piscataway, NJ: IEEE), 1540–1546.
- Cohen, J. D., Dunbar, K., and McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97:332. doi: 10.1037/0033-295x.97.3.332
- Collins, A. G., and Frank, M. J. (2014). Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psycho. Rev.* 121:337. doi: 10.1037/a0037015
- Cushman, F., and Morris, A. (2015). Habitual control of goal selection in humans. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13817–13822. doi: 10.1073/pnas.1506367112
- Daw, N. D., and Dayan, P. (2014). The algorithmic anatomy of model-based evaluation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130478. doi: 10.1098/rstb.2013.0478
- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 2005, 1704–1711. doi: 10.1038/nn1560
- Dayan, P. (1993). Improving generalization for temporal difference learning: the successor representation. *Neural Comput.* 5, 613–624. doi: 10.1162/neco.1993.5.4.613
- De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687. doi: 10.1126/science.1128356
- Dickinson, A., and Balleine, B. (2002). “The role of learning in the operation of motivational systems,” in *Learning, Motivation, and Emotion*, eds H. Pashler, and C. R. Gallistel, (New York, NY: John Wiley & Sons), 497–533.
- Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw.* 12, 961–974. doi: 10.1016/s0893-6080(99)00046-5
- Doya, K., Samejima, K., Katagiri, K., and Kawato, M. (2002). Multiple model-based reinforcement learning. *Neural Comput.* 14, 1347–1369. doi: 10.1162/089976602753712972
- Fairley, K., Parelman, J. M., Jones, M., and Carter, R. M. (2019). Risky health choices and the Balloon Economic Risk Protocol. *J. Econ. Psychol.* 73, 15–33. doi: 10.1016/j.joep.2019.04.005
- Frank, M. J., and O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav. Neurosci.* 120:497. doi: 10.1037/0735-7044.120.3.497
- Fukunaga, R., Brown, J. W., and Bogg, T. (2012). Decision making in the balloon analogue risk task (BART): anterior cingulate cortex signals loss aversion but not the infrequency of risky choices. *Cogn. Affect. Behav. Neurosci.* 12, 479–490. doi: 10.3758/s13415-012-0102-1
- Gershman, S. J. (2018). The successor representation: its computational logic and neural substrates. *J. Neurosci.* 38, 7193–7200. doi: 10.1523/JNEUROSCI.0151-18.2018
- Gray, J. A., Buhusi, C. V., and Schmajuk, N. (1997). The transition from automatic to controlled processing. *Neural Netw.* 10, 1257–1268. doi: 10.1016/s0893-6080(97)00058-0
- Haber, S. N. (2010). “Integrative networks across basal ganglia circuits,” in *Handbook of Behavioral Neuroscience*, ed M. S. Gazzaniga, (Amsterdam: Elsevier), Vol. 24, 535–552. doi: 10.1016/b978-0-12-802206-1.00027-1
- Haber, S. N. (2017). *Anatomy and Connectivity of the Reward Circuit*. In *Decision Neuroscience*, eds J.-C. Dreher and L. Tremblay (Cambridge, MA: Academic Press), 3–19.
- Hardwick, R. M., Forrence, A. D., Krakauer, J. W., and Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nat. Hum. Behav.* 3, 1252–1262. doi: 10.1038/s41562-019-0725-0
- Hassin, R. R., Bargh, J. A., and Zimerman, S. (2009). Automatic and flexible: the case of nonconscious goal pursuit. *Soc. Cogn.* 27, 20–36. doi: 10.1521/soco.2009.27.1.20
- Herd, S. A., Banich, M. T., and O'Reilly, R. C. (2006). Neural mechanisms of cognitive control: an integrative model of Stroop task performance and fMRI data. *J. Cogn. Neurosci.* 18, 22–32. doi: 10.1162/089892906775250012
- Herd, S. A., Krueger, K. A., Nair, A., and Mollick, J. A. (2019). Neural mechanisms of human decision-making. *Under Revis. Cogn. Affect. Behav. Neurosci.* (inpress).
- Hikosaka, O., and Isoda, M. (2010). Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends Cogn. Sci.* 14, 154–161. doi: 10.1016/j.tics.2010.01.006
- Huang, J. Y., and Bargh, J. A. (2014). The selfish goal: autonomously operating motivational structures as the proximate cause of human judgment and behavior. *Behav. Brain Sci.* 37, 121–135. doi: 10.1017/S0140525X13000290
- Hull, C. L. (1943). *Principles of Behavior: An Introduction to Behavior Theory*. New York, NY: Appleton-Century.
- James, W. (1890). *The principles of Psychology*. New York, NY: Henry Holt and Company.
- Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex.* 13, 400–408. doi: 10.1093/cercor/13.4.400
- Kool, W., Cushman, F. A., and Gershman, S. J. (2018). “Competition and cooperation between multiple reinforcement learning systems,” in *Goal-Directed Decision Making*, eds R. W. Morris, A. Bornstein, and A. Shenhav, (Cambridge, MA: Academic Press), 153–178. doi: 10.1016/b978-0-12-812098-9.00007-3
- Lejuez, C. W., Aclin, W. M., Zvolensky, M. J., and Pedulla, C. M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *J. Adolesc.* 26, 475–479. doi: 10.1016/s0140-1971(03)00036-8
- Littman, M. L., and Sutton, R. S. (2002). Predictive representations of state. *Adv. Neural Inform. Process. Syst.* 15:55–1561.
- Melnikoff, D. E., and Bargh, J. A. (2018). The mythical number two. *Trends Cogn. Sci.* 22, 280–293. doi: 10.1016/j.tics.2018.02.001
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Miller, K. J., Ludvig, E. A., Pezzulo, G., and Shenhav, A. (2018). “Realigning models of habitual and goal-directed decision-making,” in *Goal-Directed Decision Making*, eds R. W. Morris, A. Bornstein, and A. Shenhav, (Cambridge, MA: Academic Press), 407–428. doi: 10.1016/b978-0-12-812098-9.00018-8
- Miller, K. J., Shenhav, A., and Ludvig, E. A. (2019). Habits without values. *Psychol. Rev.* 126, 292–311. doi: 10.1037/rev0000120
- Miller, R. (1981). *Meaning and Purpose in the Intact Brain: a Philosophical, Psychological, and Biological Account of Conscious Processes*. Oxford: Oxford University Press.
- Mishkin, M., Malamut, B., and Bachevalier, J. (1984). Memories and habits: two neural systems. *Neurobiol. Learn. Mem.* 65–77. doi: 10.1523/JNEUROSCI.0840-16.2016
- Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D., and Gershman, S. J. (2017). The successor representation in human reinforcement learning. *Nat. Hum. Behav.* 1, 680. doi: 10.1038/s41562-017-0180-8
- Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947. doi: 10.1523/jneurosci.16-05-01936.1996

- O'Reilly, R. C. (1998). Six principles for biologically based computational models of cortical cognition. *Trends Cogn. Sci.* 2, 455–462. doi: 10.1016/s1364-6613(98)01241-8
- O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science* 314, 91–94. doi: 10.1126/science.1127242
- O'Reilly, R. C., and Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* 18, 283–328. doi: 10.1162/089976606775093909
- O'Reilly, R. C., Hazy, T. E., and Herd, S. A. (2016). The Leabra cognitive architecture: how to play 20 principles with nature and win! *Oxford Handb. Cogn. Sci.* 91, 91–116.
- O'Reilly, R. C., Hazy, T. E., Mollack, J., Mackie, P., and Herd, S. (2014). Goal-driven cognition in the brain: a computational framework. *arXiv preprint arXiv* [preprint]. Available at: <https://arxiv.org/abs/1404.7591> (accessed February 18, 2020).
- O'Reilly, R. C., and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press.
- Packard, M. G., and Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annu. Rev. Neurosci.* 25, 563–593.
- Pascual, A., and Guéguen, N. (2005). Foot-in-the-door and door-in-the-face: a comparative meta-analytic study. *Psychol. Rep.* 96, 122–128. doi: 10.2466/pr0.96.1.122-128
- Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433:873. doi: 10.1038/nature03287
- Pauli, W. M., Clark, A. D., Guenther, H. J., O'Reilly, R. C., and Rudy, J. W. (2012). Inhibiting PKM ζ reveals dorsal lateral and dorsal medial striatum store the different memories needed to support adaptive behavior. *Learn. Mem.* 19, 307–314. doi: 10.1101/lm.025148.111
- Pezzulo, G., Rigoli, F., and Chersi, F. (2013). The mixed instrumental controller: using value of information to combine habitual choice and mental simulation. *Front. Psychol.* 4:92. doi: 10.3389/fpsyg.2013.00092
- Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., and Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Computat. Biol.* 13:e1005768. doi: 10.1371/journal.pcbi.1005768
- Sallet, J., Mars, R. B., Noonan, M. P., Neubert, F. X., Jbabdi, S., O'Reilly, J. X., et al. (2013). The organization of dorsal frontal cortex in humans and macaques. *J. Neurosci.* 33, 12255–12274. doi: 10.1523/JNEUROSCI.5108-12.2013
- Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340. doi: 10.1126/science.1115270
- Schultz, W. (2013). Updating dopamine reward signals. *Curr. Opin. Neurobiol.* 23, 229–238. doi: 10.1016/j.conb.2012.11.012
- Seger, C. A., and Spiering, B. J. (2011). A critical review of habit learning and the basal ganglia. *Front. Syst. Neurosci.* 5:66. doi: 10.3389/fnsys.2011.00066
- Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 84:127. doi: 10.1037/0033-295x.84.2.127
- Squire, L. R. (1992). Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *J. Cogn. Neurosci.* 4, 232–243. doi: 10.1162/jocn.1992.4.3.232
- Squire, L. R., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science* 253, 1380–1386. doi: 10.1126/science.1896849
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Thorndike, E. L. (1911). *Animal Intelligence: Experimental Studies*. New York, NY: Macmillan.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55:189. doi: 10.1037/h0061626
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5
- Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232. doi: 10.1111/j.1460-9568.2009.06796.x
- Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science* 211, 453–458. doi: 10.1126/science.7455683
- Van Ravenzwaaij, D., Dutilh, G., and Wagenmakers, E. J. (2011). Cognitive model decomposition of the BART: assessment and application. *J. Math. Psychol.* 55, 94–105. doi: 10.1016/j.jmp.2010.08.010
- Verplanken, B., and Aarts, H. (1999). Habit, attitude, and planned behaviour: is habit an empty construct or an interesting case of goal-directed automaticity? *Eur. Rev. Soc. Psychol.* 10, 101–134. doi: 10.1080/14792779943000035
- White, T. L., Lejuez, C. W., and de Wit, H. (2008). Test-retest characteristics of the Balloon Analogue Risk Task (BART). *Exp. Clin. Psychopharmacol.* 16:565. doi: 10.1037/a0014083
- Wickens, J. R., Horvitz, J. C., Costa, R. M., and Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. *J. Neurosci.* 27, 8181–8183. doi: 10.1523/jneurosci.1671-07.2007
- Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends Cogn. Sci.* 7, 70–76. doi: 10.1016/s1364-6613(02)00024-4
- Wood, W., and Rünger, D. (2016). Psychology of habit. *Annu. Rev. Psychol.* 67, 289–314.
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919
- Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523. doi: 10.1111/j.1460-9568.2005.04218.x

Conflict of Interest: RO'R, AN, and SH were employed by company eCortex, Inc.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 O'Reilly, Nair, Russin and Herd. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.