Chapter 17

Computational models of motivated frontal function

RANDALL C. O'REILLY*, JACOB RUSSIN, AND SETH A. HERD
Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, United States

Abstract

Computational models of frontal function have made important contributions to understanding how the frontal lobes support a wide range of important functions, in their interactions with other brain areas including, critically, the basal ganglia (BG). We focus here on the specific case of how different frontal areas support goal-directed, motivated decision-making, by representing three essential types of information: possible plans of action (in more dorsal and lateral frontal areas), affectively significant outcomes of those action plans (in ventral, medial frontal areas including the orbital frontal cortex), and the overall utility of a given plan compared to other possible courses of action (in anterior cingulate cortex).

Computational models of goal-directed action selection at multiple different levels of analysis provide insight into the nature of learning and processing in these areas and the relative contributions of the frontal cortex versus the BG. The most common neurologic disorders implicate these areas, and understanding their precise function and modes of dysfunction can contribute to the new field of computational psychiatry, within the broader field of computational neuroscience.

INTRODUCTION

The frontal lobes are well deserving of all the attention in this volume and the broader scientific literature, due to their outsized role in so many central aspects of human cognition and behavior. However, the frontal lobes do not work alone: there is increasing evidence that frontal cortex depends critically on the basal ganglia (BG) and parietal lobes, for example. Thus, a more complete understanding of frontal function likely requires a systems-level framework that integrates the contributions of all these brain areas. A computational modeling approach can play an essential role in this context, by helping to understand how different brain systems can work together while each contributes a distinct function. In this chapter, we review some of the major computational frameworks for understanding frontal function within a larger systems perspective. To narrow the scope and provide a more concrete, substantive treatment within the very broad range of frontal functions, we focus on frontal contributions to motivated decision-making, where the ventral and medial frontal areas

(e.g., orbitofrontal cortex, OFC; anterior cingulate cortex, ACC) play central roles. Evolutionarily, it seems that these brain areas are the oldest, primary frontal areas. For example, in rodents, analogs of these areas (using primate terminology) are clearly present, whereas analogs of dorsolateral prefrontal cortex (PFC) areas are less obviously developed (Ongür and Price, 2000; Brown and Bowman, 2002; Uylings et al., 2003). Thus a focus on these areas may shed light on the essential forces shaping frontal function, which later adapted to support higher-level cognitive functions in primates and humans.

In this context, we evaluate the hypothesis that a common set of core specializations, which can be understood through mechanistic computational models of the underlying neural systems, underlies the unique frontal contributions across all frontal areas. Of particular relevance to this volume, the ventral and medial areas of frontal cortex are widely recognized as the most important for understanding common disorders of clinical relevance, including major depressive disorder, anxiety disorders, attention deficit hyperactivity disorder

---

*Correspondence to: Randall C. O'Reilly, Department of Psychology and Neuroscience, University of Colorado Boulder, 345 UCB, Boulder, CO, 80309, United States. Tel: +1-303-492-0054, Fax: +1-303-492-2967, E-mail: randy.oreilly@colorado.edu

(ADHD), obsessive–compulsive disorder (OCD), etc. (Ongür and Price, 2000; Itami and Uno, 2002; Fitzgerald et al., 2005; Kim et al., 2011; Haber and Behrens, 2014). These can all be considered disorders of motivation and affect in one way or another, highlighting the importance of understanding the neural mechanisms of motivation and how the frontal cortex and BG work together to support goal-directed behavior, and also how these systems can go awry in these disorders.

We begin with an integrated framework that synthesizes a number of different ideas from various computational models of frontal function, which then provides a roadmap for a subsequent discussion of specific models. This should provide the clinician with a broad overall sense of the kinds of insights and approaches that computational models offer and good places to look for further reading. We then broaden our discussion to include more of the unique contributions of dorsolateral PFC (DLPFC) and frontopolar PFC areas, which are most significantly enlarged in primates and humans, to provide a more complete picture of frontal function in humans, including the functions of robust and flexible working memory, higher-level cognitive function, and behavioral inhibition. Finally, we discuss the promise and pitfalls of the relatively new field of computational psychiatry, within the broader field of computational neuroscience.

A SKETCH OF AN INTEGRATED, SYSTEMS–NEUROSCIENCE COMPUTATIONAL MODEL

To situate the many different specific computational models in the literature, we begin with an overall sketch of a biologically realistic framework for affectively motivated decision-making across the frontal cortex and BG systems. Historically, the BG were originally considered to be specialized for habit learning, with the frontal lobes being more important for higher-level, "smarter" cognitive functions (Miller, 1981; Mishkin et al., 1984; Squire, 1992; Packard and Knowlton, 2002). In the domain of motivated decision-making, the distinction between a goal-directed process that takes into account subsequent outcomes of actions, versus a more habitual, stimulus–response process that does not, has been an important dimension for dissociating behaviors and underlying brain systems (e.g., Tolman, 1948; Balleine and Dickinson, 1998; Yin and Knowlton, 2006; Tricomi et al., 2009). For example, there is long-standing evidence that the OFC is important for anticipating the specific affectively relevant outcomes of previously acquired actions, so that behavioral action–selection accurately tracks manipulations such as the selective devaluation of particular rewards (Balleine and Dickinson, 1998).

OFC lesions impair this devaluation-appropriate behavior, putatively by causing the system to fall back on the conditioned habit of performing actions that lead to the now-devalued reward.

A similar distinction has been developed in computational models of reinforcement learning (RL), where model-based RL is distinguished from model-free RL (Daw et al., 2005; Daw and Dayan, 2014). Computationally, a model-free RL system memorizes (caches) the value (reward or punishment) of outcomes associated with different states of the environment, whereas a model-based RL system employs an internal model that can simulate the outcomes of different potential courses of action, to make action choices informed by their likely outcomes.

More recently, there has been increasing evidence that the BG are also important for flexible, higher-level cognitive functions (Pasupathy and Miller, 2005), and there is now a general consensus that the habit-learning role of the BG is specifically associated with the dorsolateral striatum, in conjunction with associated motor cortex areas, while other areas of the striatum play a critical role in goal-directed behavior (Samejima et al., 2005; Yin and Knowlton, 2006; Pauli et al., 2012a). Functionally, the FC is well suited for generating and evaluating a menu of possible action plans, which the BG then help to select among. The selection process in BG drives updating of FC to reflect the selected option, which is then maintained to support subsequent behavior, consistent with the classic top-down biasing framework (Desimone, 1996; O'Reilly et al., 1999; Miller and Cohen, 2001).

For this type of FC/BG dynamic, it is more natural to think in terms of articulating differences among different FC/BG loops, rather than between the FC and BG themselves. The classic FC/BG loops identified by Alexander et al. (1986) provide a suitable coarse-grained grouping of such areas (Fig. 17.1). Each of these loops can be associated with a broad category of different information: affectively significant outcomes in OFC, sensory-motor plans of action in DLPFC, and overall affectively oriented integration of this OFC and DLPFC information in the ACC. The ACC effectively provides a net-utility level of evaluation of different potential plans of action (as represented by DLPFC) in terms of their ability to achieve desired outcomes (in OFC), befitting its location and extensive interconnectivity between these two frontal areas. Consistent with a motivational/activational understanding of ACC (e.g., Stuss and Alexander, 2007; Holroyd and Yeung, 2012), when a high-utility plan is available, the ACC could then provide a "motivating" activation to energize accomplishment of that plan.

Biologically, monkey electrophysiology and other data are broadly consistent with this overall framework for understanding different contributions of OFC,
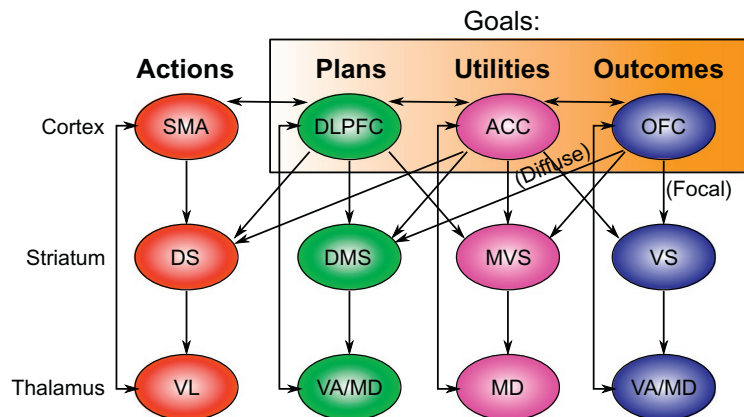
Fig. 17.1. Schematic for how four different anatomically defined loops through the frontal cortex and basal ganglia (Alexander et al., 1986) can each represent information relevant for making appropriately motivated decisions taking into account affectively relevant outcomes (orbitofrontal cortex, OFC), sensory-motor plans (dorsolateral prefrontal cortex, DLPFC), and the integration of these factors, along with potential effort and other costs in terms of overall utility (anterior cingulate cortex, ACC). All of these factors, combined together, enable the organism to act in a goal-driven manner. In contrast, the loop through the supplementary motor areas (SMAs) is thought to reflect a lower level of motor action selection, which can be shaped by the goal-driven areas but can also operate independently based on current sensory inputs, in which case it reflects longer-term habit learning of the form that has been classically ascribed to the BG.

ACC, and DLPFC (Rushworth et al., 2007; Kennerley et al., 2011; Walton et al., 2014; Rich and Wallis, 2016; Hunt et al., 2018). At a broad level, each of these areas has neurons tuned for relevant information such as overall value, specific outcomes, plans, etc., consistent with the extensive interconnectivity among these areas. However, instead of thinking of this as a mass of undifferentiated tissue, more sensitive population-coding analyses show systematic differences consistent with the preceding characterization (Hunt et al., 2018). Specifically, Hunt et al. (2018) found that OFC had both a clear encoding of the current stimulus identity, along with a more graded, linear representation of reward value, whereas the ACC had a more decision-like encoding of value, divided sharply between high and low values, and it, along with DLPFC, encoded the action plan. Moreover, the OFC was unique in encoding values of prior options, appropriately modulated by attention, whereas ACC and DLPFC only encoded the value of the currently attended option. Thus, OFC representations can support the ability to compare relative outcomes with each other, while DLPFC and ACC are more focused on evaluating a current plan of action and whether the current evidence supports or contradicts that plan. Furthermore, Hunt et al. (2018) found that a dissociable component of ACC activity ramped up to a threshold level just prior to action initiation, consistent with the idea that ACC activation is key for motivating the selected plan.

Although these data are focused on frontal activation, there are comparable data about value and action-selection coding in corresponding areas of striatum

(Samejima et al., 2005; Yin and Knowlton, 2006; Redgrave et al., 2010; Pauli et al., 2012a), and our overall framework again suggests that, whereas the frontal areas are encoding signals in a high-dimensional, dynamic fashion, the BG are evaluating these signals and helping to select appropriate cognitive steps to take in focusing attention and further processing on the available options, and then helping to select the final course of action chosen.

Overview of computational mechanisms for motivated decision-making

Within the preceding biologically motivated framework, we can now consider how different computational frameworks can help clarify the underlying mechanisms of motivated decision-making. As noted previously, the model-based versus model-free distinction has been very influential in the context of understanding how goal-directed behavior might work. However, the notion of model-based RL has been somewhat more narrowly defined within the computational models, as compared to how many empirical scientists think about these processes and as compared to the broader space of relevant computational models. Specifically, the model-based RL framework is focused on an iterative, simulation-like planning process that projects events forward in time to determine the outcomes of actions. Typically, multiple such steps of iterative simulation and evaluation are taken, much like a classic lookahead process in chess or other strategy games. By contrast, there are many simpler types of information and

processes that can also contribute to goal-directed behavior, and we are skeptical that rodents or even most other mammals engage in much of this kind of sequential, iterative planning process. Indeed, even humans are likely to favor quicker, simpler heuristic kinds of processes rather than engaging in effortful extended planning. In chess, for example, studies show that experts rely extensively on perceptual and memory-based processes, to a much greater extent than typical AI planning models (Gobet and Simon, 1996). Thus, while we are clearly capable of this kind of elaborative planning process, it is likely reserved for only the most novel, important situations, while everyday decision-making is likely dominated by other simpler processes.

From the neural recording data reviewed previously, for example, we can see how flexible attentional modulation of different value representations in OFC, together with a utility-focused evaluation of the current plan in ACC, can support effective goal-directed decision-making. The underlying value, plan, and utility representations can be learned over time and generalized from prior relevant experiences, instead of requiring a fully detailed simulation of the environment. These kinds of learned representations can be thought of as various levels of state abstraction in the broader RL framework, where the current sensory state is recoded in various ways that emphasize particularly decision-relevant factors, and support more systematic generalization of prior experience into novel situations. For example, many different computational models have been proposed for learning about future outcomes of prior situations and actions, at many different levels, and these learned predictions can provide a relatively quick and easy way of taking into account likely future consequences of actions, instead of attempting to actually simulate the current situation. A particularly simple yet powerful form of such a model learns a successor representation (Dayan, 1993) that encodes the discounted full set of future expected states likely to follow from the current state, and this aggregated state matrix provides the input to the action-selection system, which benefits from being able to generalize over states that all lead to similar subsequent outcomes.

Outside of the formal RL framework, Alexander and Brown (2011) proposed that medial PFC areas learn to encode various types of outcomes based on actions taken and showed how such a model can explain a range of electrophysiological and other data about these areas. In addition, the widely publicized work of the Google DeepMind group has pioneered the coupling of RL models with deep neural network systems that are particularly good at learning powerful state abstractions, which have been shown to perform well at various simpler video games (Mnih et al., 2015). This same framework,

augmented with an optimized model-based RL-like lookahead planner system, subsequently was able to beat the best human players at the very challenging game of Go (Silver et al., 2016). Thus, the broad framework of state-abstraction learning in support of goal-directed behavior represents a critical contribution of computational models for understanding the power of frontal function, which would be hard to appreciate and understand without the concrete, implemented models that demonstrate the impact of different forms of abstraction. We review these models in somewhat greater detail later.

Relatively less computational modeling work has been devoted to understanding how different frontal/BG areas might work together dynamically to support motivated decision-making. For example, we are not aware of biologically realistic models of the iterative planning process underlying formal model-based RL models, nor how such a system could be learned in a plausible way. Most existing models of frontal/BG action selection are focused on a single parallel competition among different options, as represented in different competing subsets of neurons in the BG pathways (Barto, 1995; Houk and Wise, 1995; Gurney et al., 2001; Doya, 2003; Frank, 2005, 2011; Humphries et al., 2006), but it is not clear how these parallel pathways might be learned for relatively novel decision-making situations, nor how the BG might participate in a more complex goal-directed decision-making process involving coordinated information across the OFC, ACC, and DLPFC.

To address some of these issues, we have recently developed an integrated computational model that captures many aspects of this framework and simulates the basic processes behind motivated decision-making (Herd et al., 2019). This model incorporates OFC, DLPFC, and ACC frontal areas, representing outcome, plan, and integrated utility information about a space of different situations where different action plans lead to different outcomes. These outcomes in turn have different net value or utility depending on internal state (e.g., hunger, thirst) and other factors. This model shows that, by encoding these different factors across different brain areas, it can successfully generalize to novel combinations of situations, action plans, and internal states, through the appropriate combination of representations like those found in the OFC, DLPFC, and ACC. A major contribution of this model is to demonstrate how this generalizable, flexible decision-making ability depends on sequentially evaluating different possible courses of action, with the BG evaluating the overall utility of each in turn for action selection. This sequential evaluation process is more plausible for complex, novel decision-making situations, compared to the parallel selection process that most BG models envision. It seems unlikely that dedicated BG channels would be available,

with appropriate learned synaptic connections from relevant frontal state abstractions, to support the kinds of novel, arbitrary decisions that humans often confront in daily life. Instead, it seems more plausible that these parallel models of BG action selection are applicable to well-learned, consolidated, "habitual" response pathways. In our overall framework, this would correspond to the SMA FC/BG loop of Alexander et al. (1986) (Fig. 17.1). However, even these pathways are likely strongly influenced by maintained activations across higher frontal areas, consistent with the spiral connectivity of these areas (Haber et al., 2000).

COMPUTATIONAL MODELS OF MOTIVATED DECISION-MAKING

In this section, we provide a somewhat more detailed description of selected computational models that illuminate important components of the motivated decision-making process, and consider how they fit within the larger framework articulated previously.

First, the model-free side of RL has been extensively explored through fundamental RL learning mechanisms such as the temporal differences (TD) actor-critic framework, where reward prediction errors (in the critic) train up a competitive action–selection process that chooses actions that are more likely to lead to higher levels of future reward (Sutton and Barto, 1981, 1998). The correspondence between elements of this framework and different parts of the brain are shown in Fig. 17.2, with the BG playing the role of the actor (in concert with its projections through the frontal motor cortex), and midbrain phasic dopamine signals playing the role of a TD-like critic (Montague et al., 1996), driving
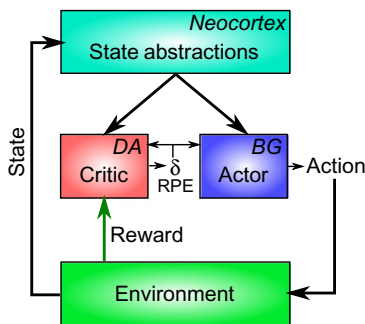

Fig. 17.2. Mapping of basic reinforcement learning (RL) functions in a temporal-differences (TD) actor-critic architecture onto brain areas (DA = midbrain dopamine system including ventral tegmental area and substantia nigra pars compacta and their driving afferents; BG = basal ganglia). The critic trains itself and the actor via reward prediction errors (RPEs), carried by dopamine and symbolized by δ in standard RL terminology. Abstractions of the environmental state, in the neocortex, play a critical role in enabling systematic, goal-directed action plans.

dopamine-modulated learning in the BG (Gerfen, 2001; Frank, 2005; Gerfen and Surmeier, 2011). The neocortex abstracts the environmental state input to support more powerful function in both the critic and the actor.

Computationally, there are different ways of organizing the learning and action selection on the actor side (i.e., learning a policy in RL terminology). For example, the basic TD learning framework learns an estimate of the expected future value associated with different states of the environment, and then chooses actions as a function of the estimated value of the states that each action leads to. This represents a simple one-step lookahead planning function, even though it is typically categorized as a model-free process, because the environmental model is very simple. The Q-learning algorithm (Watkins and Dayan, 1992) is a variant that learns the value of state-action pairs, so it does not require this lookahead process and can choose actions directly based on their learned value estimates, and it can also handle situations where actions themselves might have different values. Both of these algorithms can be used with lookup tables that just directly encode these learned expected values, or with various types of function approximators (e.g., a neural network) that learn these values in a more compact, efficient manner than the full lookup tables. Furthermore, there are many important issues concerning the trade-off between exploiting (choosing the action with the highest value) versus exploring other options—early in learning, exploration should be prioritized, as the value estimates are unlikely to be very accurate, but later in learning, it can make more sense to optimize exploitation (with the risk of only finding a locally optimal solution instead of a possibly more globally optimal one) (Sutton and Barto, 1998; Cohen et al., 2007).

The essential learning process in these models involves chaining prediction errors backward in time—for example if the system fails to accurately anticipate a reward (or punishment), this then updates its estimate for the value of the state that immediately preceded the current one. When the system is next in that state, the updated value estimate will in turn trigger an update of the state that preceded that one, and so on backward in time/states. This iterative process can be rather slow, so often a graded eligibility trace over prior states (typically parameterized by the symbol λ) can be used to push updates back across multiple prior states at once. This trace is important for capturing the properties of dopamine firing in the brain, which does not exhibit the characteristic backward chaining dynamic as predicted by TD without this trace (Pan et al., 2005). Other more biologically based models show how the different brain systems driving phasic dopamine firing produce this and

other characteristic firing patterns without requiring an eligibility trace (Hazy et al., 2010).

In a biologic context, the extensive cortical inputs into the BG action selection system play the role of state abstractions. As explained earlier, the nature of these state abstractions can play a huge role in determining how effective the overall system is. An early compelling demonstration of this point was provided by the backgammon-playing success of an RL model using a simple neural network on top of carefully hand-crafted input states (Tesauro, 1995; this was the "DeepMind" model of its era). One way of understanding why humans are so adaptable is that we have such a large and powerful state abstraction learning system in our highly enlarged neocortex, which can extract systematic information useful for making good strategic decisions. In comparison, animals with less elaborated neocortex must rely on more literal stimulus-driven state representations, which do not support sophisticated goal-directed behavior very well and instead drive a simpler stimulus–response level of learning.

The recent use of deep neural network models for discovering powerful abstract state representations, exemplified by the work of the Google DeepMind group (Mnih et al., 2015; Silver et al., 2016), demonstrates the importance of state abstractions. These models combine relatively standard Q-learning mechanisms (augmented with a few key tricks) together with powerful error back-propagation through deep layers of simple neuron-like processing elements, to shape abstract encodings of the raw environmental state in ways that support impressive overall performance on a range of different tasks. One of the key tricks here was avoiding the positive-feedback loops that tend to develop when a system's behavior (policy) is based on what it has learned so far, causing it to overly focus on exploiting learned policies instead of acquiring new ones. The DeepMind model randomly reshuffles small sequences of learning experience to avoid these positive-feedback loops. They motivated this idea in part by theories of hippocampal replay of experiences through later recall (McClelland et al., 1995), although their implementation would likely place unrealistic demands on the extent and capacity of such a hippocampal replay mechanism.

State abstractions are often thought of in terms of recoding the current environmental state, but considerable advantage can be gained by learning representations that integrate across time as well. The classic TD algorithms described previously are constrained to treat each point in time separately, continually choosing actions at each step based on their predicted future outcomes. However, it has been suggested that human planning takes place on more temporally abstract representations, where a whole sequence of actions and its cumulative outcome may be represented holistically (Botvinick and Weinstein, 2014). For example, there is extensive evidence of anticipatory motor planning in speech and other domains (e.g., Dell et al., 1997). This allows us to develop coordinated, high-level action plans, for example. A number of intersecting lines of neuroimaging and neuropsychologic research suggests that PFC may be an important brain structure underlying this kind of processing (Fuster, 1991; Koechlin and Jubault, 2006; Badre and D'Esposito, 2007; Badre, 2008; Botvinick, 2008).

Temporal abstraction has been formally explored in a series of computational frameworks collectively referred to as hierarchical RL (Barto and Mahadevan, 2003). For example, in the options framework, at each time step the agent can choose from the set of available low-level actions as in traditional RL, but it can also choose from a set of options, which are composed of subpolicies defining whole sequences of low-level actions to be executed (Sutton et al., 1999). More recently, the hierarchical RL approach has been explored with a variety of different deep neural network algorithms that allow hierarchical representations to be discovered by the agent itself rather than being hand-coded by the programmer (Kulkarni et al., 2016). A more cognitively motivated version of such models was proposed by Collins and Frank (2013), which was applied to explain a range of human behavioral data on how task representations are learned from experience.

An interesting potential upshot of temporal abstraction has been suggested by Botvinick et al. (2009) and Botvinick and Weinstein (2014): in what the latter term saltatory model-based hierarchical RL, an agent with temporally abstract representations of the state-space would be able to predict the environmental consequences of its actions beyond one time-step in the future. This capacity would greatly reduce the computational complexity of generating a temporally extended action-plan; rather than requiring a full tree-search through potential future actions at each time-step, the agent could jump directly to the predicted outcome of a sequence of prelearned (habitual) actions. For example, rather than predicting the consequences of every low-level muscle-movement in the sequence of actions required to make a cup of coffee, the outcome of the high-level plan could itself be considered in order to motivate the decision. Thus, the agent could effectively simulate the experience of having a cup of coffee in hand to help it make the decision to execute the required sequence of actions.

Another form of temporal abstraction, the successor representation mentioned earlier (Dayan, 1993), involves associating each state with a temporally abstract representation of the states that are expected to follow

from it. The original motivation behind this work was that TD-learning of a value function becomes easier when distributed representations of states with similar expected future rewards are close together (in Euclidean distance). When successor representations are used, the value function is easier to learn because states with similar successors will also have similar future rewards.

More recently, the successor representation has been reconceptualized as an intermediate between model-based and model-free algorithms, because it encodes useful information about the dynamics of the environment without requiring a full model of every possible transition probability (Momennejad et al., 2017). In a series of behavioral experiments, the authors showed that a hybrid between the successor representation algorithm and the usual model-based TD-learning best captured the basic features of human behavior on tasks involving changes to environmental contingencies (i.e., reward devaluation and changes to transition probabilities).

In related work, human participants exhibited event segmentation of visual stimuli based on temporal community structure: the stimuli that tended to follow one another (and would therefore have similar successor representations) were grouped together so that transitions within community groups were identified by participants as belonging to a single event, whereas transitions between such groups were seen as the natural starting points of new events (Schapiro et al., 2013). The fMRI results of this study offered support to the hypothesis that visual stimuli with similar expected successor items had greater overlap in their neural representations, and specifically identified left inferior frontal gyrus as an important brain area for representing temporal structure. Although these models focused on bottlenecks as an important characteristic for segmenting regions of state-space, other work has emphasized points of maximum prediction error or maximum entropy where people naturally segment events and other temporally extended stimuli (e.g., Reynolds et al., 2007). Thus it may well be that the hippocampus and other medial temporal lobe areas play a critical role in helping to organize the continuous stream of environmental states into distinct episodes (e.g., Ranganath and Hsieh, 2016; Schapiro et al., 2017) that then provide a basis for more systematic goal-directed action selection in the future (Schacter et al., 2007).

The PRO model of Alexander and Brown (2011) provides a related form of temporal abstraction learning, where the ACC is hypothesized to learn the expected subsequent outcomes of different actions, along multiple different dimensions at the same time. Critically, this framework suggests that cortical representations are high-dimensional and not reduced to a single scalar value, as in standard RL models. Furthermore, the temporal predictions of subsequent outcomes give rise to two kinds of prediction errors: absences of expected outcomes, and occurrences of unexpected outcomes— these are tracked separately in the model, and are critical for accounting for a range of neural data, including phenomena otherwise attributed to conflict coding in the ACC (Botvinick et al., 2001). The PRO model is consistent with the broad framework outlined previously, in terms of ACC connecting DLPFC motor plans with OFC outcome representations and playing a critical role in encoding the net utility of different plans under consideration. Furthermore, this model suggests how the ACC might learn based on prediction errors. A more elaborated hierarchical version of this model specifically connects the ACC with a model of DLPFC and provides a more comprehensive framework for goal-directed motor control in these areas (Alexander and Brown, 2015). Meanwhile, the original proponents of the conflict model have adopted a largely consistent, broader framework for understanding the ACC in terms of an overall net-utility construct, evaluating the cost and benefits of exerting cognitive control (Shenhav et al., 2013).

Finally, the notion that the cortical representations are fundamentally formed through predictive learning mechanisms is consistent with a detailed model of the role of the higher-order thalamus and neocortical deep layers, organized over time at the alpha frequency (10 Hz) (O'Reilly et al., 2017). Specifically, this model suggests that the neocortex is generating predictions over the thalamus (acting as a kind of projection screen for coordinating predictions across multiple areas), and every 100 ms it is driven by phasic bust firing from neocortical deep-layer projections that represent the bottom-up sensory ground truth for what actually did happen. Learning throughout the cortex occurs based on the difference between the prediction and this ground-truth outcome signal, as broadcast from the thalamus broadly back up to cortex. This same circuitry exists throughout the neocortex and could provide the neural basis for predictive learning in frontal areas (ACC, OFC, DLPFC, etc.), along the lines envisioned by many of the previous models. One specific idea is that OFC directly learns to predict amygdala, insular, and other affectively driven representations, whereas ACC learns to predict these OFC representations, conditional on plan-based inputs from the DLPFC (i.e., it learns to anticipate the affective outcomes of given high-level action plans, as in the PRO model of Alexander and Brown, 2011). The DLPFC, meanwhile, is driven by learning to predict sensory-motor outcomes, building on lower levels of such learning occurring in the parietal lobe, but also with influence from the ACC representations.

In summary, computational models of frontal/BG function in motivated decision-making provide many

important ideas for how to solve the fundamental problem of deciding what course of action to take to achieve desired future outcomes (goals). There are many different ways in which predicted future states can be anticipated, with one extreme being the classical model-based RL framework, using explicit iterative simulations of the environment. Short of this, there are many simpler "heuristic" approaches involving various forms of predictive learning that can anticipate potential future outcomes without requiring a full-blown simulation. From a cognitive and behavioral perspective, these different computational mechanisms provide different levels of explicit access to the relevant predictive knowledge, and abilities to generalize to novel situations.

For example, with a full-blown simulation, you likely have extensive access to the reason why you chose a given action plan, in terms of the predicted effects on future outcomes, whereas the simpler predictive learning approaches may correspond to more of a "hunch" or intuition about likely outcomes, without as much explicit reasoning behind it. These different strategies can also vary in their dependence on prior experience in the relevant domain and ability to generalize systematically to novel situations. Many of the computational models excel in well-trained domains (e.g., the DeepMind models play millions upon millions of games to develop their expertise), but they fail to capture that essential ability of human intelligence to rapidly adapt and extrapolate to novel tasks. A critical construct here is the spectrum of controlled versus automatic processing (Shiffrin and Schneider, 1977; Cohen et al., 1990): people use deliberate, explicit controlled processing in novel situations, and progressively develop automatic forms of processing as they gain greater familiarity. Controlled processing depends on juggling active neural firing information in PFC (Miller and Cohen, 2001; O'Reilly, 2006), whereas automatic processing relies on shaping synaptic weights through learning. Thus, advancing the frontal abilities of computational models represents an important direction for improving their overall abilities to handle novel situations.

GENERALIZING TO BROADER FRONTAL FUNCTIONS

In this section, we attempt to integrate the previous more-focused consideration of the affective decision-making aspects of frontal function with the broader space of frontal functions. We argue that this approach recapitulates the course of evolution: frontal cortex was initially specialized for supporting the core affective action selection processes considered earlier, and larger mammalian and especially primate PFC areas evolved from that foundation. In brief, aside from everything being more richly represented across larger brain areas,

it appears that the DLPFC and frontal pole areas are most uniquely enlarged in primates and especially humans (Semendeferi et al., 2001; Badre and D'Esposito, 2009; Thiebaut de Schotten et al., 2012). In the preceding framework, these areas would support the ability to represent more complex, and especially deeply hierarchical, action plans. Many have argued for various different hierarchical organizations of PFC areas (e.g., Koechlin et al., 2003; Bunge, 2004; Bunge et al., 2005; Badre, 2008; Badre and D'Esposito, 2009), but there are also indications that things might be more mixed and complex than the pure form of any of these models (Reynolds et al., 2012). Nevertheless, the highly enlarged area 10 is anatomically connected in a manner consistent with its observed role in enabling reasoning involving multiple different relationships—it only interconnects with other, more posterior frontal areas, whereas these other frontal areas are directly interconnected with posterior cortex (Badre and D'Esposito, 2009). Furthermore, there is evidence that the medial and lateral areas of dorsal frontal cortex work together along the rostral-caudal axis (Kouneiher et al., 2009), consistent with the idea that these dorsal medial areas are evaluating the affective outcomes and overall net utility of potential action plans at different levels of abstraction.

The broader dorsal versus ventral organization of frontal cortex is also consistent with the core distinctions in affective frontal areas between the OFC (ventral) and ACC (dorsal) (O'Reilly, 2010) (Fig. 17.3). Specifically, OFC and ventral-lateral PFC (VLPFC) areas are more clearly driven by aspects of stimulus identity, presumably originating in projections from the inferotemporal cortex (IT) where neurons first encode object-level information robustly (i.e., the *What* visual pathway; Ungerleider and Mishkin, 1982). In contrast, ACC and DLPFC more clearly encode action plans and related information, consistent with their projections from parietal areas that encode sensation-for-action information (i.e., the *How* visual pathway; Goodale and Milner, 1992). For example, the comparison between coding in OFC versus ACC reviewed earlier (Hunt et al., 2018) is consistent with this overall organization, with OFC showing clear encoding of stimulus information, while ACC more strongly encoded the current action plan.

In the following sections, we consider various computational models that have advanced our understanding of different mechanisms of frontal function more broadly, in relation to the overall organization sketched here.

Robust active maintenance, working memory, and BG gating

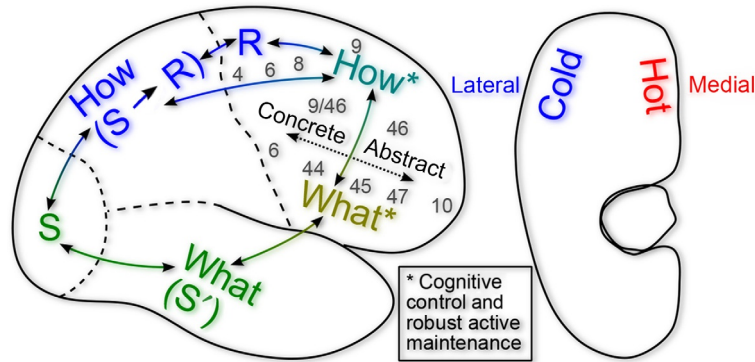A long-standing framework for understanding frontal function centers around the apparently unique capacity

Fig. 17.3. Schematic of cortical functional organization along dorsal–ventral, medial–lateral, and anterior–posterior dimensions (O'Reilly, 2010). The ventral visual pathway encodes stimulus identity information in inferotemporal cortex (IT), and this carries forward into frontal cortex, in terms of OFC representations of the affective value and associations of different stimuli, and more lateral ventral areas driving top-down control over IT representations of verbal and semantic knowledge. The dorsal pathway through parietal cortex transforms visual and other sensory input to support action planning and execution, and this carries through to dorsal lateral PFC high-level planning and control of action and corresponding evaluation of action plans in ACC.

for robust maintenance of neural firing patterns in frontal cortex (e.g., Fuster and Alexander, 1971; Kubota and Niki, 1971; Miller and Desimone, 1994; Goldman-Rakic, 1995; Sommer and Wurtz, 2000). According to the previous framework, it is plausible that this capacity for active maintenance evolved to allow goal representations to be sustained over time to guide behavior toward accomplishing selected goals and plans, and perhaps to also enable learning about action/outcome connections that span a longer temporal duration. Computationally, there have been a number of different models at different levels of analysis exploring how neurons can sustain a pattern of firing robustly over time, from more biologically detailed models to more abstract mathematical attractor models (e.g., Brunel and Wang, 2001; Durstewitz and Seamans, 2002; Goldman et al., 2002; Koulakov et al., 2002; Brody et al., 2003; Durstewitz, 2009; Barak and Tsodyks, 2014; Lim and Goldman, 2014; Zylberberg and Strowbridge, 2017).

These kinds of working memory systems face an important trade-off: to the extent that they are better at sustaining neural activity in the face of interference, it then becomes more difficult to rapidly update when new information must be maintained. A gating system can resolve this trade-off by dynamically modulating the system, opening the gate when new information must be updated, and closing the gate to support robust maintenance of existing information. One possibility is that dopamine neuromodulation drives this kind of gating dynamic (Braver and Cohen, 1999, 2000; O'Reilly et al., 1999; Durstewitz and Seamans, 2002). However, in contrast to the more global effects of dopamine, BG-mediated gating of frontal cortex can provide a more fine-grained gating signal, allowing some information to be updated while other information is robustly

maintained (Frank et al., 2001; O'Reilly, 2006; O'Reilly and Frank, 2006). This biologically based gating idea coincides with the computationally motivated long-short-term-memory (LSTM) model of gating in working memory (Hochreiter and Schmidhuber, 1997), which is currently experiencing a resurgence of interest along with other neural-network learning mechanisms, and demonstrates the computational power of this gating framework. This gating framework also corresponds to a partially observable Markov decision process (Todd et al., 2008), which has been studied extensively in machine learning and computer science, and other approaches have also been used to model this gating process (Dayan, 2007, 2008).

The LSTM model and later versions of the BG-gating model (Hazy et al., 2007) have included an output gating mechanism in addition to maintenance gating—this form of gating allows information to be actively maintained in a form that does not influence ongoing processing until the output gate is opened. There is evidence for such a dynamic operating in the human brain (Chatham et al., 2014), and it can support a powerful form of indirection (Kriete et al., 2013) along with other benefits (Kriete and Noelle, 2011). Our current understanding of the interaction between BG and frontal cortex now incorporates the differential roles and connectivity of the superficial and deep layers of neocortex, along with a central role of predictive learning (O'Reilly et al., 2012, 2017). In brief, we now think of BG gating as mediating the transfer of activation from a more labile, transient superficial-layer representation to a more robustly maintained deep-layer representation, which, via excitatory loops through the thalamus, is ultimately responsible for robust active maintenance. This framework also unifies the functions of maintenance gating and output gating: both involve

the activation of deep-layer representations, with the primary difference being the subsequent projections from these deep layers. Output gating corresponds to activation of deep-layer representations with extensive direct connections to other cortical areas to effect motor or cognitive actions, while maintenance gating corresponds to representations that primarily project to other PFC areas that can then serve as output gating areas for this maintained information. These distinctions are clearly graded, and any given area may reflect a mix of both patterns of connectivity. This is in contrast to the LSTM framework, which has a stricter segregation between maintenance and output gating.

This overall BG-gating framework has the potential to connect the extensive literatures on the involvement of ventral striatum (including the nucleus accumbens, NAcc), and ventral/medial frontal cortex (including OFC, ACC), in similar domains and disorders. Despite the fact that these areas are extensively interconnected, many studies examine each of these areas separately. Instead, it may make more sense to consider them as a system, with the NAcc playing a critical role in driving a gating-like updating of robust active maintenance in OFC, ACC, etc. For example, when NAcc responds phasically to conditioned stimuli (CS) as a result of dopamine-modulated learning there, this could drive updating of CS and associated unconditioned stimuli (US) outcome information in OFC. The OFC can then robustly maintain this information to track progress toward achieving desired outcomes. This overall framework is consistent with a variety of different types of data, including the relative impact of OFC on reversal learning (Stalnaker et al., 2007; Pauli et al., 2012b), and selective deactivation of dorsal versus ventral striatum (Atallah et al., 2006; Pauli et al., 2010).

Furthermore, it may potentially help understand the differences between sign-tracking versus goal-tracking behavior (Flagel et al., 2010). Some recent data shows the promise of this approach for understanding psychiatric disorders (Burguière et al., 2013).

Symbolic cognition and the nature of PFC representations

One active, unresolved question in the field is the extent to which the PFC is important for enabling abstract, symbolic cognition, and if there are particular signatures of this in the nature of PFC representations (e.g., Miller et al., 2003; Nieder, 2009; O'Reilly et al., 2014). Some single-cell monkey electrophysiological data comparing PFC versus posterior cortical representations have concluded that the PFC representations are more crisply tuned to the task-relevant dimensions, e.g., showing sharp task-defined visual categorical boundaries

(Freedman et al., 2003) or more symbol-like number representations (Diester and Nieder, 2007). A computational model demonstrated how sustained activate maintenance can shape PFC representations to have these more symbol-like representations (Rougier et al., 2005). On the other hand, other recordings have demonstrated mixed selectivity in PFC representations (Mante et al., 2013; Rigotti et al., 2013; Fusi et al., 2016), where individual neurons have significant tuning for multiple different stimulus and task dimensions. Mixed selectivity is a hallmark of posterior-cortical representations and has been discussed in terms of coarse-coded distributed representations in earlier neural network literature (e.g., Rumelhart and McClelland, 1986; O'Reilly and Busby, 2002). It allows efficient representation of high-dimensional information, and different downstream pathways can extract different features from these representations as needed. However, a critical limitation of these representations is in generative, combinatorial tasks, when different representations need to be flexibly and systematically combined, without any additional opportunity for tuning up new representations (i.e., single-trial generalization). In this case, higher-dimensional distributed representations can interfere with each other, significantly impairing this ability for systematic generalization, while more crisply categorical, symbol-like representations enable more flexible recombination of elements (Brousse, 1993; O'Reilly, 2001).

This kind of instant generative behavior is generally considered to be something that only humans truly excel at, with limited evidence of it outside of our species. Thus there is always the possibility that the PFC in other primates has only the basic precursors of the kinds of truly systematic representations present in the human brain, but this is obviously a difficult hypothesis to test, given the limitations of current recording techniques. Furthermore, even in the macaque monkey, some of the differences noted previously may be due to the nature of the tasks involved: the mixed-selectivity tasks were strongly perceptually driven and lacked any significant demand for abstraction or flexibility (e.g., same-different matching of complex picture stimuli, or motion versus color discrimination). By contrast, the more categorical representations were observed in a task with systematic stimulus variation along two dimensions, only one of which was task-relevant, and when systematically associating visual representations of number with arabic digits. Thus, it would be interesting to more systematically explore the effects of task factors on the nature of the learned representations in PFC, particularly in tasks that require truly generative single-trial generalization.

There is a further question regarding the nature of the influence of BG gating on PFC activation patterns and representations. In an extreme case, BG gating should

impart a strong discretization over time and space on corresponding activity in the PFC, as specific subsets of PFC neurons are gated on and off. In our original models, we hypothesized that topographically organized anatomic features called stripes in PFC might correspond to independently gatable subsets of PFC neurons (Frank et al., 2001; O'Reilly and Frank, 2006). However, it is not clear if the relevant projections between the PFC and thalamus, and through the BG subareas, actually obey a strong topographic ordering constraint, as opposed to being more diffusely projecting. In this latter case, it would be difficult to trace the influence of BG gating on a coherent subset of PFC neurons if those neurons were widely scattered. Furthermore, it is not clear to what extent a topographic organization is functionally beneficial for the gating function—further computational studies are necessary to explore this issue. Given the number of synapses involved in these BG/PFC loops, it would be technically challenging to trace the topography all the way through, but we are hopeful that future anatomic studies may help resolve some of these questions. Meanwhile, it does seem to be generally true that PFC neurons exhibit relatively sharp transitions over time, particularly when they deactivate after a response has been made—this is consistent with a gating-like dynamic. The potential influence of BG gating on PFC activations could be tested through selective optogenetic manipulation studies.

Inhibition

From a clinical perspective, it has long been useful to think of the frontal cortex, and its interconnectivity with the BG, as supporting behavioral inhibition: for example, in the early observation of utilization behavior in patients with damage to a variety of prefrontal areas (Lhermitte et al., 1986), and the observation of a close linkage between the right inferior frontal cortex (rIFC) and performance on stop-signal tasks, in which a prepared response must be withheld at the last possible moment (Aron et al., 2004). This latter framework is also a nice example of how frontal–BG circuits might work together in coordinating the initiation and control over motor action. An important question that computational models have impacted is whether this overall behavioral inhibition reflects literal neural inhibition emanating from the frontal cortex, or whether it is more accurate to understand the frontal cortex as maintaining and supporting desired behaviors via excitatory connections, which then has the consequent effect of inhibiting unwanted behaviors (Cohen and Servan-Schreiber, 1992; Miller and Cohen, 2001; Stedron et al., 2005; Herd et al., 2006; Munakata et al., 2011).

In the context of understanding the functional contributions of the rIFC, recent studies have shown that it is equally activated when a signal triggers pressing a button twice instead of withholding a response, in an otherwise identical paradigm (Chatham et al., 2012) and even when detecting a salient stimulus when no external response is required (Hampshire et al., 2010). These data and others are consistent with the alternative view that rIFC is important for attentional monitoring of the signal (consistent with the broader perspective of IFC as top-down control over ventral pathway visual processing; O'Reilly, 2010), rather than driving literal inhibition of the response itself. Another potential framing would be to think of the rIFC, and right frontal cortex more generally, as maintaining alternative action plans, while the left frontal cortex is engaged with the currently active, dominant plan (Charron and Koechlin, 2010). Although proponents of the direct inhibition framework have argued that rIFC activation always causes some kind of breaking or interruption of motor execution (Aron et al., 2014), that may not be exclusive to this area, and instead may reflect the consequences of any form of update to the current action plan in frontal cortex, via the hyperdirect pathway from frontal cortex to the subthalamic nucleus (Frank, 2006). Unraveling the precise contributions of different brain areas in a complex interacting circuit such as this is an ideal target for computational modeling.

COMPUTATIONAL PSYCHIATRY AND FRONTAL/BG MODELS

The frontal cortex/BG system is strongly implicated in most of the common mental disorders, including major depressive disorder, anxiety disorders, OCD, schizophrenia, addiction, and ADHD. This fact, combined with the evident complexity of these systems and relative limitations of existing pharmacologic treatments, has led to a recent focus on the potential of computational frameworks to impact the next generation of treatment approaches, i.e., computational psychiatry (Montague et al., 2012; Wang and Krystal, 2014; Huys et al., 2016). The interest in supporting this approach from the NIH and other funding agencies is obviously exciting to proponents of the computational modeling approach, but it also carries obvious risks. It is likely that the extensive comorbidity between these disorders reflects common failure modes arising from many different etiologies, within a highly complex interactive system. These failure modes reflect the dominance of negative affective states such as fear, which has a clear evolutionary survival value, even if it also leads to some level of dysfunction in modern human society. Thus, effective treatment requires fighting against millions of years of

entrenched evolutionary forces, and our expectations of probability of success should be adjusted accordingly. Nevertheless, the degree of fragmentation of empirical research along many different lines (e.g., animal versus human, BG versus frontal cortex) provides an important opportunity for computational modeling approaches to integrate across these different areas and provide coherent frameworks for thinking about how these systems function at many different levels. Furthermore, aside from any particular theoretic advances, computational and machine learning techniques have demonstrated potential to improve diagnosis and select appropriate treatments based on an individual's specific profile (Huys et al., 2016).

Computational psychiatry can be considered a subset of the broader discipline of computational neuroscience, which uses computational models to understand how the brain works more generally. Many of the computational models discussed in this chapter are important examples of computational neuroscience, and in particular systems–neuroscience level models. One of the enduring challenges in this field is figuring out how to evaluate and test these often complex models, and understanding exactly what these models have contributed to the broader field of neuroscience and how these contributions can be enhanced going forward. In many ways, models function like more specific versions of the kinds of theories that drive all of science: they provide ways of interconnecting different phenomena and levels of analysis into a more coherent and self-consistent explanatory account. Like theories, their ultimate value is in driving further experimental tests. Furthermore, once a model is sufficiently well understood, its fundamental insights can often be conveyed without going into much of the specific details (as we have attempted to do here in this review), leading nonmodelers to often wonder about whether all that effort developing the model in the first place was really necessary. Instead of being discouraged, one could regard this as a compliment: all major scientific advances eventually seem obvious in retrospect, and all the false leads and misguided ideas along the way are quickly forgotten. So it is with computational modeling: it is a process that some people find useful for figuring out how complex systems might work.

If modelers can do a good job of conveying their insights in ways that lead to useful experiments, and the field as a whole progresses toward a shared understanding of how these brain systems function together in ways that are consistent with both the experimental data and coherent computational models, then we can likely agree that progress is being made. Indeed, an optimistic view suggests that our existing understanding of the large-scale systems–neuroscience-level organization of the brain has already converged on a stable, self-consistent understanding of many of the major functions. Nevertheless, much work remains to be done, and many of the issues raised here represent some of the most challenging and important directions for future work in this field.

REFERENCES

Alexander WH, Brown JW (2011). Medial prefrontal cortex as an action–outcome predictor. Nat Neurosci 14 (10): 1338–1344.

Alexander WH, Brown JW (2015). Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. Neural Comput 27 (11): 2354–2410.

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

Aron AR, Robbins TW, Poldrack RA (2004). Inhibition and the right inferior frontal cortex. Trends Cogn Sci 8: 170–177.

Aron AR, Robbins TW, Poldrack RA (2014). Inhibition and the right inferior frontal cortex: one decade on. Trends Cogn Sci 18 (4): 177–185.

Atallah HE, Lopez-Paniagua D, Rudy JW et al. (2006). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. Nat Neurosci 10: 126–131.

Badre D (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn Sci 12: 193–200.

Badre D, D'Esposito M (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. J Cogn Neurosci 19: 2082–2099.

Badre D, D'Esposito M (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? Nat Rev Neurosci 10 (9): 659–669.

Balleine BW, Dickinson A (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37: 407–419.

Barak O, Tsodyks M (2014). Working models of working memory. Curr Opin Neurobiol 25: 20–24.

Barto AG (1995). Adaptive critics and the basal ganglia. In: JC Houk, JL Davis, DG Beiser (Eds.), Models of information processing in the basal ganglia, MIT Press, Cambridge, MA, pp. 215–232.

Barto AG, Mahadevan S (2003). Recent advances in hierarchical reinforcement learning. Discrete Event Dyn Syst 13 (4): 341–379.

Botvinick MM (2008). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cogn Affect Behav Neurosci 7: 356–366.

Botvinick M, Weinstein A (2014). Model-based hierarchical reinforcement learning and human action control. Philos Trans R Soc B 369 (1655), pii: 20130480.

Botvinick MM, Braver TS, Barch DM et al. (2001). Conflict monitoring and cognitive control. Psychol Rev 108: 624–652.

Botvinick M, Niv Y, Barto AC (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. Cognition 113 (3): 262–280.

Braver TS, Cohen JD (1999). Dopamine, cognitive control, and schizophrenia: the gating model. Prog Brain Res 121: 327–349.

Braver TS, Cohen JD (2000). On the control of control: the role of dopamine in regulating prefrontal function and working memory. In: S Monsell, J Driver (Eds.), Control of cognitive processes: attention and performance XVIII, MIT Press, Cambridge, MA, pp. 713–737.

Brody CD, Romo R, Kepecs A (2003). Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. Curr Opin Neurobiol 13: 204–211.

Brousse O (1993). Generativity and systematicity in neural network combinatorial learning, Technical Report CU-CS-676-93, University of Colorado at Boulder, Department of Computer Science, Boulder, CO.

Brown VJ, Bowman EM (2002). Rodent models of prefrontal cortical function. Trends Neurosci 25: 340–343.

Brunel N, Wang XJ (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. J Comput Neurosci 11: 63–85.

Bunge SA (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. Cogn Affect Behav Neurosci 4 (4): 564–579.

Bunge SA, Wendelken C, Badre D et al. (2005). Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. Cereb Cortex 15 (3): 239–249.

Burguière E, Monteiro P, Feng G et al. (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. Science 340 (6137): 1243–1246.

Charron S, Koechlin E (2010). Divided representation of concurrent goals in the human frontal lobes. Science 328: 360–363.

Chatham CH, Claus ED, Kim A et al. (2012). Cognitive control reflects context monitoring, not motoric stopping, in response inhibition. PLoS One 7: e31546.

Chatham C, Frank M, Badre D (2014). Corticostriatal output gating during selection from working memory. Neuron 81 (4): 930–942.

Cohen JD, Servan-Schreiber D (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. Psychol Rev 99: 45–77.

Cohen JD, Dunbar K, McClelland JL (1990). On the control of automatic processes: a parallel distributed processing model of the stroop effect. Psychol Rev 97 (3): 332–361.

Cohen JD, McClure SM, Yu AJ (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos Trans R Soc Lond Ser B Biol Sci 362 (1481): 933–942.

Collins AGE, Frank MJ (2013). Cognitive control over learning: creating, clustering, and generalizing task-set structure. Psychol Rev 120 (1): 190–229.

Daw ND, Dayan P (2014). The algorithmic anatomy of model-based evaluation. Philos Trans R Soc B Biol Sci 369 (1655), pii: 20130478.

Daw ND, Niv Y, Dayan P (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8 (12): 1704–1711.

Dayan P (1993). Improving generalization for temporal difference learning: the successor representation. Neural Comput 5 (4): 613–624.

Dayan P (2007). Bilinearity, rules, and prefrontal cortex. Front Comput Neurosci 1 (1): 1–14.

Dayan P (2008). Simple substrates for complex cognition. Front Neurosci 2 (2): 255.

Dell GS, Burger LK, Svec WR (1997). Language production and serial order: a functional analysis and a model. Psychol Rev 104 (1): 123–147.

Desimone R (1996). Neural mechanisms for visual memory and their role in attention. Proc Natl Acad Sci 93 (24): 13494–13499.

Diester I, Nieder A (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. PLoS Biol 5 (11): e294.

Doya K (2003). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw 12: 961–974.

Durstewitz D (2009). Implications of synaptic biophysics for recurrent network dynamics and active memory. Neural Netw 22: 1189–1200.

Durstewitz D, Seamans JK (2002). The computational role of dopamine D1 receptors in working memory. Neural Netw 15: 561–572.

Fitzgerald KD, Welsh RC, Gehring WJ et al. (2005). Error-related hyperactivity of the anterior cingulate cortex in obsessive-compulsive disorder. Biol Psychiatry 57 (3): 287–294.

Flagel SB, Robinson TE, Clark JJ et al. (2010). An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction. Neuropsychopharmacology 35 (2): 388–400.

Frank MJ (2005). When and when not to use your subthalamic nucleus: lessons from a computational model of the basal ganglia. Modelling natural action selection: proceedings of an international workshop, pp. 53–60.

Frank MJ (2006). Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw 19 (8): 1120–1136.

Frank MJ (2011). Computational models of motivated action selection in corticostriatal circuits. Curr Opin Neurobiol 21 (3): 381–386.

Frank MJ, Loughry B, O'Reilly RC (2001). Interactions between the frontal cortex and basal ganglia in working

memory: a computational model. Cogn Affect Behav Neurosci 1: 137–160.

Freedman DJ, Riesenhuber M, Poggio T et al. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci 23 (12): 5235–5246.

Fusi S, Miller EK, Rigotti M (2016). Why neurons mix: high dimensionality for higher cognition. Curr Opin Neurobiol 37: 66–74.

Fuster JM (1991). Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle. In: A Diamond (Ed.), The development and neural bases of higher cognitive functions, 608. New York Academy of Science Press, New York, pp. 318–336.

Fuster JM, Alexander GE (1971). Neuron activity related to short-term memory. Science 173: 652–654.

Gerfen CR (2001). Molecular effects of dopamine on striatal-projection pathways. Trends Neurosci 23: S64–S70.

Gerfen CR, Surmeier DJ (2011). Modulation of striatal projection systems by dopamine. Annu Rev Neurosci 34: 441–466.

Gobet F, Simon HA (1996). The roles of recognition processes and look-ahead search in time-constrained expert problems solving: evidence from grand-master-level chess. Psychol Sci 7: 52.

Goldman MS, Kaneko CRS, Major G et al. (2002). Linear regression of eye velocity on eye position and head velocity suggests a common oculomotor neural integrator. J Neurophysiol 88: 659–665.

Goldman-Rakic PS (1995). Architecture of the prefrontal cortex and the central executive. Ann N Y Acad Sci 769: 71–84.

Goodale MA, Milner AD (1992). Separate visual pathways for perception and action. Trends Neurosci 15 (1): 20–25.

Gurney K, Prescott TJ, Redgrave P (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biol Cybern 84 (6): 401–410.

Haber SN, Behrens TEJ (2014). The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. Neuron 83 (5): 1019–1039.

Haber SN, Fudge JL, McFarland NR (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci 20: 2369–2382.

Hampshire A, Chamberlain SR, Monti MM et al. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. NeuroImage 50: 1313–1319.

Hazy TE, Frank MJ, O'Reilly RC (2007). Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. Philos Trans R Soc Lond Ser B Biol Sci 362 (1485): 1601–1613.

Hazy TE, Frank MJ, O'Reilly RC (2010). Neural mechanisms of acquired phasic dopamine responses in learning. Neurosci Biobehav Rev 34 (5): 701–720.

Herd SA, Banich MT, O'Reilly RC (2006). Neural mechanisms of cognitive control: an integrative model of Stroop task performance and fMRI data. J Cogn Neurosci 18 (1): 22–32.

Herd SA, Krueger K, Nair A et al. (2019). Neural mechanisms of human decision-making. (submitted for publication).

Hochreiter S, Schmidhuber J (1997). Long short term memory. Neural Comput 9: 1735–1780.

Holroyd CB, Yeung N (2012). Motivation of extended behaviors by anterior cingulate cortex. Trends Cogn Sci 16 (2): 122–128.

Houk JC, Wise SP (1995). Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cereb Cortex 5 (2): 95–110.

Humphries MD, Stewart RD, Gurney KN (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. J Neurosci 26 (50): 12921–12942.

Hunt LT, Malalasekera WMN, de Berker AO et al. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. Nat Neurosci 21: 1471–1481.

Huys QJM, Maia TV, Frank MJ (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 19 (3): 404.

Itami S, Uno H (2002). Orbitofrontal cortex dysfunction in attention-deficit hyperactivity disorder revealed by reversal and extinction tasks. Neuroreport 13 (18): 2453.

Kennerley SW, Behrens TEJ, Wallis JD (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat Neurosci 14 (12): 1581–1589.

Kim MJ, Gee DG, Loucks RA et al. (2011). Anxiety dissociates dorsal and ventral medial prefrontal cortex functional connectivity with the amygdala at rest. Cereb Cortex 21 (7): 1667–1673.

Koechlin E, Jubault T (2006). Broca's area and the hierarchical organization of human behavior. Neuron 50 (6): 963–974.

Koechlin E, Ody C, Kouneiher F (2003). Neuroscience: the architecture of cognitive control in the human prefrontal cortex. Science 302: 1181–1184.

Koulakov AA, Raghavachari S, Kepecs A et al. (2002). Model for a robust neural integrator. Nat Neurosci 5: 775–782.

Kouneiher F, Charron S, Koechlin E (2009). Motivation and cognitive control in the human prefrontal cortex. Nat Neurosci 12 (9): 659–669.

Kriete T, Noelle DC (2011). Generalisation benefits of output gating in a model of prefrontal cortex. Connect Sci 23 (2): 119–129.

Kriete T, Noelle DC, Cohen JD et al. (2013). Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. Proc Natl Acad Sci U S A 110 (41): 16390–16395.

Kubota K, Niki H (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. J Neurophysiol 34 (3): 337–347.

Kulkarni TD, Narasimhan K, Saeedi A et al. (2016). Hierarchical deep reinforcement learning: integrating temporal abstraction and intrinsic motivation. In: DD Lee, M Sugiyama, UV Luxburg et al. (Eds.), Advances in neural information processing systems (NIPS). 29. MIT Press, Cambridge, MA.

Lhermitte F, Pillon B, Serdaru M (1986). Human autonomy and the frontal lobes. Part I: imitation and utilization

behavior: a neuropsychological study of 75 patients. Ann Neurol 19 (4): 326–334.

Lim S, Goldman MS (2014). Balanced cortical microcircuitry for spatial working memory based on corrective feedback control. J Neurosci 34 (20): 6790–6806.

Mante V, Sussillo D, Shenoy KV et al. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503 (7474): 78–84.

McClelland JL, McNaughton BL, O'Reilly RC (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102 (3): 419–457.

Miller R (1981). Meaning and purpose in the intact brain: a philosophical, psychological, and biological account of conscious processes, Oxford University Press.

Miller EK, Cohen JD (2001). An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24: 167–202.

Miller EK, Desimone R (1994). Parallel neuronal mechanisms for short-term memory. Science 263: 520–522.

Miller EK, Nieder A, Freedman DJ et al. (2003). Neural correlates of categories and concepts. Curr Opin Neurobiol 13: 198–203.

Mishkin M, Malamut B, Bachevalier J (1984). Memories and habits: two neural systems. In: G Lynch, JL McGaugh, NM Weinberger (Eds.), Neurobiology of learning and memory, Guilford, New York, pp. 65–77.

Mnih V, Kavukcuoglu K, Silver D et al. (2015). Human-level control through deep reinforcement learning. Nature 518 (7540): 529–533.

Momennejad I, Russek EM, Cheong JH et al. (2017). The successor representation in human reinforcement learning. Nat Hum Behav 1 (9): 680–692.

Montague PR, Dayan P, Sejnowski TJ (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16 (5): 1936–1947.

Montague PR, Dolan RJ, Friston KJ et al. (2012). Computational psychiatry. Trends Cogn Sci 16 (1): 72–80.

Munakata Y, Herd SA, Chatham CH et al. (2011). A unified framework for inhibitory control. Trends Cogn Sci 15 (10): 453–459.

Nieder A (2009). Prefrontal cortex and the evolution of symbolic reference. Curr Opin Neurobiol 19 (1): 99–108.

O'Reilly RC (2001). Generalization in interactive networks: the benefits of inhibitory competition and Hebbian learning. Neural Comput 13 (6): 1199–1242.

O'Reilly R (2006). Biologically based computational models of high-level cognition. Science 314 (5796): 91–94.

O'Reilly RC, Busby RS (2002). Generalizable relational binding from coarse-coded distributed representations. In: TG Dietterich, S Becker, Z Ghahramani (Eds.), Advances in neural information processing systems (NIPS) 14, MIT Press, Cambridge, MA.

O'Reilly RC, Frank MJ (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput 18 (2): 283–328.

O'Reilly RC, Braver TS, Cohen JD (1999). A biologically based computational model of working memory. In: A Miyake, P Shah (Eds.), Models of working memory: mechanisms of active maintenance and executive control, Cambridge University Press, New York, pp. 375–411.

O'Reilly RC, Munakata Y, Frank MJ et al. (2012). Computational cognitive neuroscience, first ed. Wiki Book. http://ccnbook.colorado.edu.

O'Reilly RC, Petrov AA, Cohen JD et al. (2014). How limited systematicity emerges: a computational cognitive neuroscience approach. In: IP Calvo, J Symons (Eds.), The architecture of cognition: rethinking fodor and pylyshyn's systematicity challenge, MIT Press, Cambridge, MA.

O'Reilly RC, Wyatte DR, Rohrlich J (2017). Deep predictive learning: a comprehensive model of three visual streams. arXiv: 1709.04654 [q-bio]. https://arxiv.org/abs/1709.04654.

Ongür D, Price JL (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb Cortex 10 (3): 206–219.

O'Reilly RC (2010). The what and how of prefrontal cortical organization. Trends Neurosci 33 (8): 355–361.

Packard MG, Knowlton BJ (2002). Learning and memory functions of the basal ganglia. Annu Rev Neurosci 25: 563–593.

Pan W-X, Schmidt R, Wickens JR et al. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. J Neurosci 25 (26): 6235–6242.

Pasupathy A, Miller EK (2005). Different time courses for learning-related activity in the prefrontal cortex and striatum. Nature 433: 873–876.

Pauli WM, Atallah HE, O'Reilly RC (2010). Integrating what & how/where with instrumental and Pavlovian learning: a biologically based computational model. In: PA Frensch, R Schwarzer (Eds.), Cognition and neuropsychology—international perspectives on psychological science, vol. 1. Psychology Press, East Sussex, UK, pp. 71–95.

Pauli WM, Clark AD, Guenther H et al. (2012a). Inhibiting PKMζ reveals dorsal lateral and dorsal medial striatum store the different memories needed to support adaptive behavior. Learn Mem 19: 307–314.

Pauli WM, Hazy TE, O'Reilly RC (2012b). Expectancy, ambiguity, and behavioral flexibility: separable and complementary roles of the orbital frontal cortex and amygdala in processing reward expectancies. J Cogn Neurosci 24 (2): 351–366.

Ranganath C, Hsieh L-T (2016). The hippocampus: a special place for time. Ann N Y Acad Sci 1369 (1): 93–110.

Redgrave P, Rodriguez M, Smith Y et al. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. Nat Rev Neurosci 11 (11): 760–772.

Reynolds J, Zacks J, Braver T (2007). A computational model of event segmentation from perceptual prediction. Cogn Sci 31 (4): 613–634.

Reynolds JR, O'Reilly RC, Cohen JD et al. (2012). The function and organization of lateral prefrontal cortex: a test of competing hypotheses. PLoS One 7: e30284.

Rich EL, Wallis JD (2016). Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19 (7): 973–980.

Rigotti M, Barak O, Warden MR et al. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature 497 (7451): 585–590.

Rougier NP, Noelle D, Braver TS et al. (2005). Prefrontal cortex and the flexibility of cognitive control: rules without symbols. Proc Natl Acad Sci 102 (20): 7338–7343.

Rumelhart DE, McClelland JL (1986). PDP models and general issues in cognitive science. In: DE Rumelhart, JL McClelland, P. R. Group (Eds.), Parallel distributed processing. Vol. 1: foundations, MIT Press, Cambridge, MA, pp. 110–146.

Rushworth MFS, Behrens TEJ, Rudebeck PH et al. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn Sci 11 (4): 168–176.

Samejima K, Ueda Y, Doya K et al. (2005). Representation of action-specific reward values in the striatum. Science (New York, NY) 310 (5752): 1337–1340.

Schacter DL, Addis DR, Buckner RL (2007). Remembering the past to imagine the future: the prospective brain. Nat Rev Neurosci 8 (9): 657.

Schapiro AC, Rogers TT, Cordova NI et al. (2013). Neural representations of events arise from temporal community structure. Nat Neurosci 16 (4): 486–492.

Schapiro AC, Turk-Browne NB, Botvinick MM et al. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. Philos Trans R Soc Lond B Biol Sci 372 (1711), pii: 20160049.

Semendeferi K, Armstrong E, Schleicher A et al. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. Am J Phys Anthropol 114: 224–241.

Shenhav A, Botvinick MM, Cohen JD (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79 (2): 217–240.

Shiffrin RM, Schneider W (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychol Rev 84: 127–190.

Silver D, Huang A, Maddison CJ et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature 529 (7587): 484.

Sommer MA, Wurtz RH (2000). Composition and topographic organization of signals sent from the frontal eye field to the superior colliculus. J Neurophysiol 83 (4): 1979–2001.

Squire LR (1992). Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. J Cogn Neurosci 4 (3): 232–243.

Stalnaker TA, Franz TM, Singh T et al. (2007). Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. Neuron 54: 51–58.

Stedron JM, Sahni SD, Munakata Y (2005). Common mechanisms for working memory and attention: the case of perseveration with visible solutions. J Cogn Neurosci 17: 623–631.

Stuss DT, Alexander MP (2007). Is there a dysexecutive syndrome? Philos Trans R Soc Lond 362 (1481): 901–915.

Sutton RS, Barto A (1981). Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88 (2): 135–170.

Sutton RS, Barto AG (1998). Reinforcement learning: an introduction, MIT Press, Cambridge, MA.

Sutton R, Precup D, Singh S (1999). Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning. Artif Intell 112 (1–2): 181–211.

Tesauro G (1995). Temporal difference learning and TD-gammon. Commun ACM 38 (3): 58–68.

Thiebaut de Schotten M, Dell'Acqua F, Valabregue R et al. (2012). Monkey to human comparative anatomy of the frontal lobe association tracts. Cortex 48 (1): 82–96.

Todd MT, Niv Y, Cohen JD (2008). Learning to use working memory in partially observable environments through dopaminergic reinforcement. In: D Koller (Ed.), Advances in neural information processing systems (NIPS) 21, vol. 21. Curran Associates.

Tolman E (1948). Cognitive maps in rats and men. Psychol Rev 55 (4): 189–208.

Tricomi E, Balleine B, O'Doherty J (2009). A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci 29: 2225–2232.

Ungerleider LG, Mishkin M (1982). Two cortical visual systems. In: DJ Ingle, MA Goodale, RJW Mansfield (Eds.), The analysis of visual behavior, MIT Press, Cambridge, MA, pp. 549–586.

Uylings H, Groenewegen H, Kolb B (2003). Do rats have a prefrontal cortex? Behav Brain Res 146 (1–2): 3–17.

Walton ME, Chau BK, Kennerley SW (2014). Prioritising the relevant information for learning and decision making within orbital and ventromedial prefrontal cortex. Curr Opin Behav Sci 1: 78–85.

Wang X-J, Krystal JH (2014). Computational psychiatry. Neuron 84 (3): 638–654.

Watkins CJCH, Dayan P (1992). Technical note: Q-learning. Mach Learn 8: 279.

Yin HH, Knowlton BJ (2006). The role of the basal ganglia in habit formation. Nat Rev Neurosci 7 (6): 464–476.

Zylberberg J, Strowbridge BW (2017). Mechanisms of persistent activity in cortical circuits: possible neural substrates for working memory. Annu Rev Neurosci 40 (1): 603–627.