

Hierarchical Control Over Effortful Behavior by Rodent Medial Frontal Cortex: A Computational Model

Clay B. Holroyd
University of Victoria

Samuel M. McClure
Stanford University

The anterior cingulate cortex (ACC) has been the focus of intense research interest in recent years. Although separate theories relate ACC function variously to conflict monitoring, reward processing, action selection, decision making, and more, damage to the ACC mostly spares performance on tasks that exercise these functions, indicating that they are not in fact unique to the ACC. Further, most theories do not address the most salient consequence of ACC damage: impoverished action generation in the presence of normal motor ability. In this study we develop a computational model of the rodent medial prefrontal cortex that accounts for the behavioral sequelae of ACC damage, unifies many of the cognitive functions attributed to it, and provides a solution to an outstanding question in cognitive control research—how the control system determines and motivates what tasks to perform. The theory derives from recent developments in the formal study of hierarchical control and learning that highlight computational efficiencies afforded when collections of actions are represented based on their conjoint goals. According to this position, the ACC utilizes reward information to select tasks that are then accomplished through top-down control over action selection by the striatum. Computational simulations capture animal lesion data that implicate the medial prefrontal cortex in regulating physical and cognitive effort. Overall, this theory provides a unifying theoretical framework for understanding the ACC in terms of the pivotal role it plays in the hierarchical organization of effortful behavior.

Keywords: anterior cingulate cortex, prefrontal cortex, hierarchical reinforcement learning, cognitive control, effort

The anterior cingulate cortex (ACC) has been implicated in a wide range of seemingly irreconcilable cognitive functions (e.g., Bush, 2009). We recently proposed that many of these functions can be captured by assuming that the ACC motivates extended behaviors to achieve larger task goals (Holroyd & Yeung, 2012). This hypothesis harmonizes well with recently documented verbal reports of a patient receiving intracranial ACC stimulation:

I started getting this feeling like . . . I was driving into a storm . . . Like you're headed toward a storm that's on the other side, maybe a couple of miles away, and you've got to get across the hill and all of a sudden you're sitting there going how am I going to get over that, through that? (Parvizi, Rangarajan, Shirer, Desai, & Greicius, 2013, p. 1361)

Couched within the formal theoretical framework of hierarchical reinforcement learning (HRL; Sutton, Precup, & Singh, 1999), our theory of ACC function holds that the ACC exploits computational efficiencies afforded by collections of actions that are represented together based on their conjoint goals, called “options” in the language of HRL, that allow for behavior to be selected on the basis of superordinate tasks that are manipulated at higher levels of abstraction (Botvinick, 2012; Botvinick, Niv, & Barto, 2009; Hengst, 2012). In this view, ACC stimulation encouraged the patient to execute an extended, difficult task (like traveling miles to pass through a storm) rather than any task-related action in particular (such as turning a sharp bend in the road). In contrast to alternative theories of ACC function, the HRL theory also crucially accounts for behavioral sequelae of ACC damage, namely, slowed responding and reduced motor activity that, in extreme cases, is observed as akinetic mutism, or the near absence of willed behavior despite normal motor capability (Holroyd & Yeung, 2012)—predicting that ACC damage would have attenuated this patient's motivation to drive through the storm. Yet although the theory is consistent with a wide range of data, it has still to be instantiated in a computational model that makes explicit its underlying assumptions, while demonstrating its internal consistency. In the present study, we develop such a model.

The model addresses two prominent, competing ideas about ACC function. First, numerous studies have indicated that the ACC is responsible for linking chosen actions with their associated reward outcomes in the service of supporting adaptive behavior (Rushworth, Buckley, Behrens, Walton, & Bannerman, 2007;

This article was published Online First December 1, 2014.

Clay B. Holroyd, Department of Psychology, University of Victoria; Samuel M. McClure, Department of Psychology, Stanford University.

Clay B. Holroyd is supported, in part, by funding from the Canada Research Chairs program, a Michael Smith Foundation for Health Research Scholar Award, and a Natural Sciences and Engineering Research Council of Canada Discovery Grant and Discovery Accelerator Supplement (312409-05). Samuel M. McClure is supported by National Institutes of Health Grant MH091068. Simulation materials can be downloaded from http://web.uvic.ca/~lccl/sites/default/files/Holroyd_McClure_files.zip

Correspondence concerning this article should be addressed to Clay B. Holroyd, P. O. Box 1700, University of Victoria, Victoria, British Columbia, Canada, V8W 2Y2. E-mail: holroyd@uvic.ca

Walton, Croxson, Behrens, Kennerley, & Rushworth, 2007). When formalized according to principles of reinforcement learning (RL)—a body of algorithms used to train autonomous agents to navigate uncertain environments to obtain rewards and avoid punishments (Sutton & Barto, 1998)—this hypothesis suggests that the ACC serves as an “actor” mechanism that adaptively regulates behavior according to feedback received from the environment (Holroyd & Coles, 2002). By contrast, most existing computationally specific theories of ACC function propose a role for this brain area in performance monitoring (e.g., W. H. Alexander & Brown, 2011; Botvinick, 2007; Botvinick, Braver, Barch, Carter, & Cohen, 2001; J. W. Brown & Braver, 2005; Shenhav, Botvinick, & Cohen, 2013; Silvetti, Alexander, Verguts, & Brown, 2014; Silvetti, Seurinck, & Verguts, 2011), a function that is more closely related to a “critic” mechanism in the language of RL (Holroyd & Coles, 2008). Further, other neural systems, such as the striatum, are also believed to implement an actor-related RL function, which raises the question of whether the ACC brings added value to solving RL problems (Holroyd & Yeung, 2011, 2012). Thus, despite a rich theoretical framework from which to draw (Sutton & Barto, 1998), a specific actor-related function for the ACC has yet to be precisely specified (but see Khamassi, Lallée, Enel, Procyk, & Dominey, 2011, for a recently proposed alternative). Our modeling efforts provide such an account.

Second, several prominent theories implicate the ACC in the deployment of cognitive control (e.g., E. K. Miller & Cohen, 2001). Although we have suggested that trial-to-trial changes in ACC activation related to control processes such as conflict monitoring may be incidental to its core function (Holroyd & Yeung, 2012), the HRL theory also holds that sustained activation of the ACC ensures ongoing control over task performance. Accordingly, we propose an actor-related mechanism for the ACC that specifically delineates its role in supporting cognitive control. The model also presents a possible answer to an unresolved question in control research: Previous theories have posited that the frontal cortex represents task-related, contextual information that biases information processing in other neural systems, especially when surface features of the task elicit overlearned or automatic responses that are inappropriate for the given task context (e.g., J. D. Cohen, Dunbar, & McClelland, 1990; Kimberg & Farah, 1993). This control mechanism has provided a foundation for numerous studies examining trial-to-trial changes in response times (RTs; e.g., Botvinick et al., 2001), but has left unspecified how the system determines what task to execute in the first place, and how vigorously and carefully the task should be carried out (Umemoto & Holroyd, 2014). Accordingly, our computational model links the cognitive control and RL literatures to demonstrate how the control system utilizes reward-related information to select and motivate particular tasks for execution. In so doing, the model demonstrates that an actor mechanism described within the HRL framework can encompass many of the cognitive control-related characteristics attributed to the ACC.

In sum, we present a computational model of ACC function that addresses these challenges while demonstrating that the proposed function is fragile to neural insults. This work develops an earlier suggestion that the ACC guides action selection by deciding between which of several possible high-level action plans to follow (Holroyd & Coles, 2002) in the context of our more recent HRL theory of ACC function. Because nonhuman animal studies have

provided a wealth of fertile data that are not readily obtainable in humans, and because rodent behavior studied in the laboratory is simple enough for model simulation, we focused our initial efforts on understanding the rodent medial frontal cortex. The model proposes a hierarchical relationship between the prelimbic cortex, the ACC, and the striatum, each of which control an effortful process on the basis of a common set of principles carried out at differing levels of abstraction. This formal instantiation of the HRL–ACC theory provides a means for explicitly delineating the theory’s underpinning assumptions while also demonstrating its internal consistency.

Computational Theory

The model is based on three principles related to hierarchy, control, and reinforcement (Holroyd & Yeung, 2011, 2012). These principles are implemented in an abstract RL framework that captures the higher level computational goals of the system, as opposed to a detailed neural model that is more biophysically realistic (M. X. Cohen & Frank, 2009). Although the model simulates rat behavior, its structure is informed by data from both rat and primate (including human) studies, as described in what follows.

Hierarchy

For problems characterized by hierarchical structure, representing sequences of behavior as temporal abstractions can enhance computational efficiency by reducing the size of the problem space. Consider that it is usually easier to plan a road trip as a sequence of high-level goals (e.g., drive to the nearest town, stop for lunch, get gasoline, and so on) rather than as a computationally exhaustive sequence of actions that compose those goals (e.g., open the car door, sit in the driver’s seat, and so on). In the language of HRL, the high-level actions constitute options that represent options-specific action policies, where each policy is described by a rule or algorithm that indicates what “primitive” action to elicit for each environmental state encountered. The policy can consist as a single set of stimulus–response mappings that is applied over and over again, such as “turn left at each rock and turn right at each pine tree,” or as a sequence of actions, such as “follow the trail until it reaches the lake”—the unifying principle being that it is usually more tractable to learn about the policy as a whole rather than about the individual actions that comprise the policy. In principle, representing tasks at higher levels of temporal abstraction can facilitate problem solving by manipulating behavior at the task level rather than at the level of primitive actions (Botvinick, Niv, et al., 2009). Our simulations assume that the medial frontal cortex exploits such hierarchy.

By definition, hierarchical theories describe systems composed of at least two levels. In the case of HRL, these consist of a higher level that selects options and a lower level that selects primitive actions specific to the option in play—hereafter simply called “actions”—so as to maximize value relative to the overarching option (Sutton & Barto, 1998). Following RL convention, we assume that subjects navigate through a “gridworld,” wherein each element of the grid represents a particular location in the environment. The simulated animal transitions from cell to cell each time the action level selects among five possible actions: moving north,

south, east, or west, or not moving (“sitting”). In turn, the higher option level selects and maintains an option that guides the lower action level—which is to say that it selects and maintains the task or behavioral strategy. Thus, for example, if the option level were to select cleaning dishes over doing laundry, then the action level would be biased to execute actions consonant with the goal of cleaning dishes because only those actions maximize option-specific value.

Of course, many real-world problems are sufficiently complicated for a hierarchical system to benefit from more than two levels of abstraction. For instance, in a typical task-switching experiment, human participants are asked to switch between different tasks from trial to trial. As we have described it here, the individual stimulus–response mappings associated with each task would be carried out by the action level, whereas the task itself would be selected at the options level. But these particular options only make sense in the higher level context of performing the experiment itself, which affords the opportunity for selecting between the tasks. Alternatively, the subjects might have chosen to participate in other competing

high-level behaviors, such as going to the beach, which would afford a different set of options (e.g., swimming, sunbathing, playing volleyball, and so on), each of which would come with their own set of options-specific actions. We term this third level of abstraction as a “metaoption” level that specifies a set of metaoptions, in which each metaoption comprises a set of meta-option-specific options. Our model thus contains three levels (see Figure 1). In principle, more levels could be possible, but three are sufficient to demonstrate the core principles of the HRL theory. Furthermore, because rodent subjects are normally compelled to participate in their experiments, the metaoption level consists of only a single metaoption corresponding to the experiment itself.

To be specific, the task environment for each experiment is described by an 11×11 gridworld, in which each cell constitutes a unique environmental state (s). Experiments are separated into trials (T) divided into discrete time steps (t). At each time t , five potential actions (a) are allowable from any state s : moving one cell north, south, east, or west, or remaining in the same cell

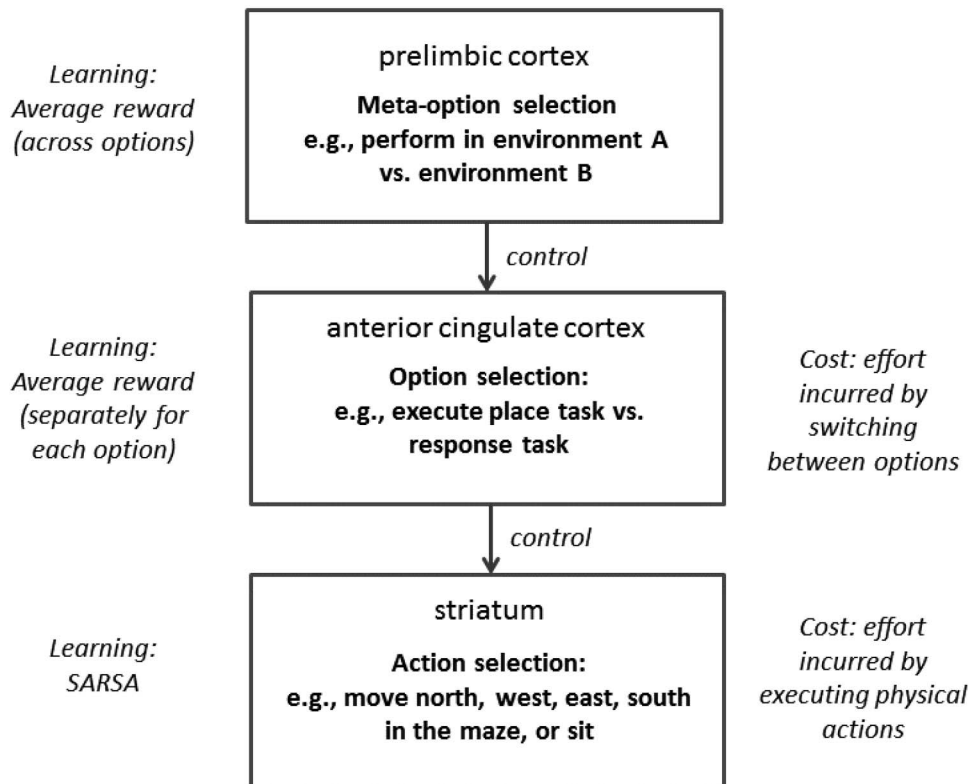


Figure 1. Schematic of frontal midline function. Bottom: The striatum and other brain areas select low-level actions on the basis of learned values of future reward relative to the energetic costs of executing those actions. Action selection is determined probabilistically over choices to move north, south, east, or west, or to sit. Energetic costs incurred by action selection are attenuated by a top-down control signal supplied to the striatum by the anterior cingulate cortex (ACC). Middle: The ACC selects and implements high-level options, such as whether to execute a task based on allocentric (place task) or egocentric (response task) frames of reference, as determined by integrating across trials the reward received during option execution and relative to a cost associated with switching between options. The ACC regulates the degree of control over the striatum according to whether rewards received are better or worse than the average for the option in play. Top: The prefrontal cortex implements the metaoption to participate in the experiment or not and applies control over ACC to regulate the cost of switching between options. The degree of prefrontal control is determined according to whether rewards are better or worse than average across the experiment.

(“sitting”). Each trial terminates when a primary reward (r) is encountered or when $t = 500$, whichever occurs first. Action selection is governed by a three-level hierarchy, with levels (L) from 1 to 3 ($L = 1-3$) that select actions, options, and metaoptions, respectively, according to the softmax equation (Sutton & Barto, 1998):

$$P_L(o_{L+1}, o_L, s) = \frac{e^{\tau_L^{-1} Q_L(o_{L+1}, o_L, s)}}{\sum_{o_M=1}^{n_L} e^{\tau_L^{-1} Q_L(o_{L+1}, o_M, s)}} \quad (1)$$

where $P_L(o_{L+1}, o_L, s)$ is the probability that level L selects from state s option o_L specific to the option o_{L+1} selected by level $L + 1$. Options o_L correspond to actions, options, and metaoptions for $L = 1-3$, respectively. n_L indicates the number of options available at level L : $n_1 = 5$ (one for each action), $n_2 = 1$ or 2 (as determined by the experiment), and $n_3 = 1$ (which enforces the selection of a single metaoption—the experiment). τ_L is a level-specific temperature constant that regulates the relative probabilities of option selection ($\tau_L > 0$ for all L) (Sutton & Barto, 1998). Note that trial notation T and time notation t are dropped for convenience; metaoption and option selection occur at the start of each trial T (i.e., whenever $t = 1$), and action selection occurs on each time t for each trial T . Finally, $Q_L(o_{L+1}, o_L, s)$ denotes option value.

In keeping with established neural models of RL, we assume that the action-selection mechanism is subserved by the basal ganglia (e.g., Botvinick, Niv, et al., 2009; Niv, 2009), a heuristic that we adopt because of long-standing precedent (e.g., Barto, 1995; Houk, Adams, & Barto, 1995). We view the striatum’s role in the model as summarizing the behavior of a complex network of neural systems that contribute to the production of actions, including the parietal cortex, the cerebellum, and the hippocampus. Further, in line with our previous proposals that the primate ACC (in what has been termed, more precisely, the anterior midcingulate cortex; Vogt, 2009) uses RL signals to select between action policies (Holroyd & Coles, 2002), we assume that option selection and maintenance is carried out by the ACC (Holroyd & Yeung, 2012). This proposal is partly inspired by the work of Botvinick, Niv, et al. (2009), who, considering possible neural mechanisms underlying HRL in humans, suggested that the dorsolateral prefrontal cortex (DLPFC) is responsible for option selection and maintenance. Our idea is isomorphic with their proposal, but instead holds that option selection and maintenance is mediated by the ACC rather than by the DLPFC, and aligns with the observation that the rodent ACC codes for contextual, high-order rules that specify individual task sets or strategies (Haddon & Killcross, 2006; Ma, Hyman, Lindsay, Phillips, & Seamans, 2014). The theory thus proposes that the rodent ACC is concerned with the selection and maintenance of each option—that is, with the task itself—rather than with the details of task execution: By learning option values according to principles of RL, and by selecting options based on those learned values, the ACC decides what task to perform and directs the basal ganglia to implement the task.

Where in the brain is the metaoption level located? To elucidate this question, we draw insight from the human neuroimaging literature. Although previous theories of the human ACC have posited that the rostral and caudal ACC subserve affective and cognitive processing, respectively (Bush, Luu, & Posner, 2000),

recent observations have cast doubt on this suggestion (Shackman et al., 2011). Rather, a growing body of evidence in humans indicates that increasingly rostral parts of the frontal cortex implement increasingly higher levels of hierarchical abstraction (Badre & D’Esposito, 2007, 2009; Badre & Frank, 2012; Collins, Cavanagh, & Frank, 2014; Dixon, Fox, & Christoff, 2014; Jeon & Friederici, 2013; Koehlin & Summerfield, 2007; Kounieher, Charron, & Koehlin, 2009; O’Reilly, 2010; Orr & Banich, 2014; but see Crittenden & Duncan, 2014; Farooqui, Mitchell, Thompson, & Duncan, 2012; Reynolds, O’Reilly, Cohen, & Braver, 2012). We thus assume that the metaoption level occurs in an area of the rat medial frontal cortex rostral to the ACC. In rodents, a candidate structure is the prelimbic cortex, a brain area said to be anatomically related to the rostral ACC in primates (Seamans, Lapish, & Durstewitz, 2008). These two suggestions—that the frontal cortex is organized along a rostral–caudal hierarchy, and that the medial frontal cortex is similarly hierarchically structured in rats and humans—are controversial and therefore constitute strong assumptions of the model. We return to these issues in the Discussion section.

Note that each option represents a temporal abstraction of the sequences of smaller actions that move the simulated rats within each trial from various starting locations to distant end locations. Such temporal abstraction across collections of extended behaviors constitutes a fundamental aspect of HRL (Botvinick, Niv, et al., 2009). A related computational efficiency afforded by HRL is the ability to sequence the options themselves for a greater purpose. However, the experiments described here entail rats engaging either in a single option repeatedly or in switching between two options, but never in executing a succession of different options. Our emphasis on examining options in isolation is consistent with Botvinick, Niv, et al.’s (2009) identification of options with task sets that, as noted earlier, serve to represent sets of stimulus–response mappings. For this reason, task sets provide a convenient entry to the empirical study of option selection and maintenance. In this way, the simulations—despite explicitly accounting for rat behavior—speak to numerous primate studies that have examined ACC responses to task selection (Forstmann, Brass, Koch, & von Cramon, 2006; Haynes et al., 2007), task initiation (Murtha, Chertkow, Beauregard, Dixon, & Evans, 1996), task switching (Boorman, Rushworth, & Behrens, 2013; Hayden, Pearson, & Platt, 2011; Hikosaka & Isoda, 2010; Hyafil, Summerfield, & Koehlin, 2009; Johnston, Levin, Koval, & Everling, 2007; Parris, Thai, Benattayallah, Summers, & Hodgson, 2007; Rushworth, Hadland, Paus, & Sipila, 2002) and task maintenance and coordination (Aarts, Roelofs, & van Turenout, 2008; M. X. Cohen, 2011; Dosenbach et al., 2006, 2007; Dosenbach, Fair, Cohen, Schlaggar, & Petersen, 2008; Gevins, Smith, McEvoy, & Yu, 1997; Torta, Costa, Duca, Fox, & Cauda, 2013; Womelsdorf, Johnston, Vinck, & Everling, 2010; Woodward, Ruff, & Ngan, 2006).

Control

Cognitive control refers to a collection of neurocognitive processes concerned with the orchestration of complex behaviors, especially when the appropriate course of action is underconstrained by information from the external environment, or when

the environment elicits automatic behaviors that are inconsistent with task goals. The DLPFC, the ACC, and other regions of the frontal cortex are thought to implement a diverse array of control functions, including goal setting, planning, and performance monitoring (E. K. Miller & Cohen, 2001; Stuss & Knight, 2002). One important function is the ability to override habitual or impulsive behaviors that are inappropriate for the task at hand. For example, a driver must inhibit the impulse to accelerate at a green traffic signal if the way is blocked by a road crew; and a dieter can struggle against an overwhelming desire to finish a nearby package of cookies. Such inner duels are a hallmark of human experience and the subject of numerous two-system theories of behavioral choice (Epstein, 1994; Fudenberg & Levine, 2006; Heatherton & Wagner, 2011; Hofmann, Friese, & Strack, 2009; Kahneman, 2003; Sloman, 1996).

At the computational level, these observations suggest that the mechanisms for action selection and option selection can be decoupled. Yet current HRL-inspired theories of neural function do not include mechanisms for cognitive control. Rather, models of choice based on RL have proposed either “flat” mechanisms that continuously select between goal-driven and habit-driven neural systems for action selection (Daw, Niv, & Dayan, 2005; Keramati, Dezfouli, & Piray, 2011; Shah & Barto, 2009), or hierarchical mechanisms in which the task specified by the higher level is invariably executed faithfully by the lower level (Caluwaerts et al., 2012; Dezfouli & Balleine, 2012, 2013). Other approaches provide separate accounts of HRL (Ribas-Fernandes et al., 2011) and effortful control (Shenhav et al., 2013). By contrast, our proposal spans these literatures by presenting a hierarchical mechanism that permits conflict between levels.

We propose that the ACC exerts control over the striatum to prevent decoupling between the lower-level mechanism for action selection and the higher level mechanism that maintains task context as specified by the currently activated option. Our account holds that the ACC not only chooses the option but also determines the level of control to apply toward executing the option once selected, while maintaining the option in working memory until the option reaches its termination state. This proposal is supported by evidence from both rat and primate studies (Baeg et al., 2003; Dosenbach et al., 2008; Petit, Courtney, Ungerleider, & Haxby, 1998). When the selected option is associated with a high value and control is necessary for successful task completion, then the ACC exerts vigorous top-down control over the basal ganglia (Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013), consistent with an “energizing” function of the ACC inferred from rat studies (Warden et al., 2012) and human clinical and neuroimaging studies (M. P. Alexander, Stuss, Shallice, Picton, & Gillingham, 2005; Kounieher et al., 2009; Picton, Stuss, Alexander, et al., 2007; Picton, Stuss, Shallice, Alexander, & Gillingham, 2006; Stuss et al., 2005). Conversely, when control is required but the level of activation provided by the ACC to the basal ganglia is weak, then subjects fail to maintain attention to the task at hand (Passetti, Chudasama, & Robbins, 2002; Rushworth, Hadland, Gaffan, & Passingham, 2003) and task performance is unsustainable (Amiez, Joseph, & Procyk, 2006; Camille, Tsuchida, & Fellows, 2011; Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006). This proposal is also supported by observations that ACC activity is often inversely correlated with RTs (Mulert, Gallinat, Dorn, Herrmann, & Winterer, 2003; Naito et al., 2000; Winterer,

Adams, Jones, & Knutson, 2002), and that ACC damage typically results in longer and more variable RTs (M. P. Alexander et al., 2005; Løvstad et al., 2012; Modirrousta & Fellows, 2008; Naccache et al., 2005; Picton, Stuss, Alexander, et al., 2007; Picton, Stuss, Shallice, et al., 2006; Stuss et al., 2005; Vendrell et al., 1995) and higher false alarm rates (Tsuchida & Fellows, 2009). We further propose that the influence of the prelimbic cortex over option selection by the ACC comports to the same arrangement. These option and metaoption control signals are implemented in the model in the form of variables that minimize cost terms related to action selection and option selection, respectively.

This architecture addresses the issue of action–option decoupling but invites an equally puzzling question: Why should the system *ever* relinquish maximum control? In humans, control famously waxes and wanes such that a diet can be obeyed one day and not the next, yet we would never design a computer that “tried hard” on some days but not on others. There would appear to be no good reason for implementing a variable control signal at the algorithmic level. As might be expected, then, the underlying cause of control variability is highly controversial. A dominant theory holds that the exertion of control utilizes a resource that depletes with use, such that control is withdrawn when the resource on which it depends is depleted or is being husbanded for future exploitation (Ackerman, 2011; Baumeister & Heatherton, 1996; Hagger, Wood, Stiff, & Chatzisarantis, 2010; van der Linden, Frese, & Meijman, 2003). Consistent with this possibility, application of top-down control is associated with an effortful cost (Kool, McGuire, Wang, & Botvinick, 2013) that the system minimizes to a level just sufficient to sustain adequate task performance (Kool, McGuire, Rosen, & Botvinick, 2010; Westbrook, Kester, & Braver, 2013; Yeung & Monsell, 2003). But the intuitive idea that control varies over time *because* it utilizes a limited resource has been challenged by competing theories that instead point to the opportunity costs of staying versus switching tasks (Kurzban, Duckworth, Kable, & Myers, 2013) or to evolving task priorities (Inzlicht & Schmeichel, 2012; Inzlicht, Schmeichel, & Macrae, 2014). Although we favor resource-based accounts (Holroyd, 2013, 2014), our model is consistent with any theory that proposes a relaxation of control for easy or automatic tasks (e.g., J. D. Cohen et al., 1990).

A third strong assumption of the model—in addition to the assumptions that the frontal cortex is organized along a rostral–caudal hierarchy, and that the medial frontal cortex is similarly hierarchically structured in rats and humans—is that each of the higher levels of the hierarchy control a cost term associated with the selection process in their immediately lower levels. Thus, the metaoption level controls a cost associated with option selection by the option layer, and the option layer controls a cost associated with action selection by the action layer. These costs are assumed to derive from the effort entailed in regulatory control over behavioral and cognitive processing (Kool et al., 2013). In the case of actions, we propose that action execution incurs a cost associated with the effortful expenditure of physical energy, as suggested by observations that people prefer to do nothing over something, even for actions that exact only minimal energetic costs (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Brockner, Shaw, & Rubin, 1979). Further, though not explicitly incorporated in the model, these costs would also relate to the effort entailed in

overcoming physical pain. This idea was implemented in the action selection algorithm as a bias against actions with high energetic (or painful) costs that, in turn, could be attenuated by an option-dependent control signal. The proposal aligns with substantial evidence that the ACC in rats (e.g., Cowen, Davis, & Nitz, 2012; Hauber & Sommer, 2009; Hillman & Bilkey, 2010, 2012; Walton, Bannerman, & Rushworth, 2002) and primates (Botvinick, Huffstetler, & McGuire, 2009; Croxson, Walton, O'Reilly, Behrens, & Rushworth, 2009; Davis, Hutchison, Lozano, Tasker, & Dostrovsky, 2000; Kurniawan, Guitart-Masip, Dayan, & Dolan, 2013; Mulert et al., 2008; Paus, Koski, Caramanos, & Westbury, 1998; Prévost, Pessiglione, Metereau, Cléry-Melin, & Dreher, 2010) is differentially activated by effortful costs and in response to painful stimuli (Dum, Levinthal, & Strick, 2009; Johansen & Fields, 2004; Vogt & Sikes, 2009). The idea is also consistent with evidence from rat studies that ACC neurons provide a neural substrate for the storage and expression of mental schemas that, like options, map context and events onto appropriate actions (Euston, Gruber, & McNaughton, 2012; S. H. Wang, Tse, & Morris, 2012), providing long-term memories for control (Jung, Baeg, Kim, Kim, & Kim, 2008).

What effortful cost is incurred by *option* selection? Suggestively in this regard, Botvinick, Niv, et al. (2009) observed that the concepts of task sets in cognitive psychology and options in HRL share important commonalities, namely, that both ideas

postulate a unitary representation that (1) can be selected or activated; (2) remains active for some period of time following its initial selection; (3) leads to the imposition of a specific stimulus-response mapping or policy; and (4) can participate in hierarchical relations with other representations of the same kind.

These considerations indicate that task sets constitute a particular type of high-level option. Further, it is commonly observed in humans that switching between two different tasks results in longer RTs and higher error rates compared with repeating the same task (Kiesel et al., 2010; Monsell, 2003). Given the proposed correspondence between task sets and options, it follows that switching between options exacts a switch cost. Importantly, in humans, overcoming this switch cost is effortful (Kool et al., 2013) and aversive (Dreisbach & Fischer, 2012), resulting in a bias for sticking with the same option in order to minimize the cost (Botvinick, 2007; Kool et al., 2010). Further, switch costs are normally larger for easier (faster and more accurate) options than for harder (slower and less accurate) options (Allport, Styles, & Hsieh, 1994), especially when the options share characteristics that invite cross-talk (Yeung & Monsell, 2003). Remarkably, this asymmetrical switch cost is large enough to inhibit participants from switching away from the harder option to the easier option, despite the higher demands of the harder task (Millington, Poljac, & Yeung, 2013; Yeung, 2010). In the model, we assume that rat behavior is similarly constrained by a switch cost, as observed in humans, which is implemented as a cost in the option-selection algorithm that penalizes switches and that is modulated by a control signal applied from the metaoption level.

We implemented these ideas as follows. $Q_L(o_{L+1}, o_L, s)$ denotes the value of state-option pair (o, s) at level L specific to the

option selected at level $L + 1$. This value can be separated into two terms:

$$Q_L(o_{L+1}, o_L, s) = V_L(o_{L+1}, o_L, s) - \frac{C_L(o_L, s)}{1 + \epsilon_{L+1}} \quad (2)$$

where $V_L(o_{L+1}, o_L, s)$ is the value of state-option pair (o, s) at level L specific to the option selected at level $L + 1$, as determined by the learning algorithm for that level, C_L is the level-dependent cost for executing option o_L in state s ($C_L \geq 0$), and ϵ_{L+1} is a “control” variable determined by the immediately higher level $L + 1$. Note that for $L = 1$ and a given option o_{L+1} , V constitutes a $5 \times 11 \times 11$ array corresponding to the values of the five actions available at each cell within the 11×11 gridworld. For $L = 2$, V constitutes a two-element array corresponding to the value of two available behavioral tasks (e.g., response vs. place; see Cross-Maze Task simulation); and for $L = 3$, V constitutes a scalar variable denoting the value of the experiment itself. For $L = 3$, τ_L , ϵ_{L+1} , and o_{L+1} are not defined, as the model allows for only one metaoption and the probability of selecting this metaoption is always 1.0. This equation indicates that high levels of control from the metaoption and option levels attenuate the costs of selecting options or actions in the levels that they control, which vanish in the limit. Conversely, the absence of control results in maximum exposure to the cost of option and action selection, which biases the system against selecting these possibilities.

Finally, in addition to the action- and option-selection costs described thus far, we assume that maintaining an option (or metaoption) in working memory over the course of a trial exerts an effortful cost that, all other things being equal, results in reduced control over time. The decay of control is implemented in an update rule (discussed later).

Reinforcement

The HRL theory of the ACC integrates concepts of hierarchy and control with reinforcement (Holroyd & Yeung, 2012). The involvement of the ACC in RL and reward processing is well known in primate studies and discussed elsewhere (e.g., Amiez, Joseph, & Procyk, 2005; Cai & Padoa-Schioppa, 2012; Hayden & Platt, 2010; Holroyd & Coles, 2002; Holroyd & Yeung, 2011; Kennerley, Behrens, & Wallis, 2011; Luk & Wallis, 2009; Matsumoto, Matsumoto, Abe, & Tanaka, 2007; Sallet et al., 2007; Shidara & Richmond, 2002; Shima & Tanji, 1998; Walsh & Anderson, 2012). In the model, reinforcement exerts two primary effects. First, the reinforcement determines values for action and option selection according to learning algorithms that are specific to each level of the hierarchy. At the lower level, action selection operates according to standard principles of RL such that the animal learns the value of actions available in different states of the world (Sutton & Barto, 1998; see Figure 1). Values are learned progressively from reward prediction error signals that indicate when events are better or worse than expected. Current thinking holds that the reward prediction errors are encoded as phasic increases and decreases in the firing rate of midbrain dopamine neurons (Schultz, 2013), which are utilized by the striatum for the purpose of adaptive decision making (M. X. Cohen & Frank, 2009; Montague, Hyman, & Cohen, 2004), although it is likely that value learning is subserved by the striatum in conjunction with other

neural systems to which the striatum interconnects (Haber & Knutson, 2010).

To be specific, learning of action values (V) follows principles of flat RL. For $L = 1$, V is learned according to the SARSA (State-Action-Reward-State-Action) update rule (Sutton & Barto, 1998)

$$V_{L=1,t-1}(o_{L=2}, a, s) \leftarrow V_{L=1,t-1}(o_{L=2}, a, s) + \alpha \delta'_t \quad (3)$$

where (a, s) is the state-action pair occurring at time step $t-1$; $o_{L=2}$ specifies the $11 \times 11 \times 5$ array of state-action values specific to the option selected at level $L = 2$; α is the learning rate ($0 < \alpha < 1$); and the subscript t is provided for clarity. Further, δ'_t is the temporal difference error at time t , as given by

$$\delta_t = r_t + \gamma V_{L=1,t}(o_{L=2}, a', s') - V_{L=1,t-1}(o_{L=2}, a, s) \quad (4)$$

and

$$\delta'_t = \begin{cases} \delta_t, & \text{if } \delta_t \geq 0 \\ \delta_t \xi, & \text{if } \delta_t < 0 \end{cases} \quad (5)$$

where (a', s') is the state-action pair occurring at time step t ; r_t is the “primary” reward encountered at time t ($r \geq 0$); γ is the discount parameter ($0 \leq \gamma \leq 1$) (Sutton & Barto, 1998); and ξ is a constant ($0 \leq \xi \leq 1$). $V_{L=1,t}$ is defined as zero for terminal states. Note that Equation 5 introduces an asymmetry between the strength of negative and positive temporal difference errors that effectively reduces the impact of negative temporal difference errors relative to positive temporal difference errors because $\xi < 1$ (see, e.g., Frank, Moustafa, Haughey, Curran, & Hutchison, 2007), resulting in slowed reversal learning.

At the next level of the motor control hierarchy, the ACC is hypothesized to learn the value of option execution and to apply top-down control over the striatum in line with the value of the current task. Likewise, the prelimbic cortex is hypothesized to learn the value of the metaoption and to apply top-down control over the ACC in line with the experiment itself. This proposal holds that the ACC (or the prelimbic cortex) associates each option (or metaoption) with the average reward received across trials during the execution of that option (or metaoption). Such learning may be mediated by tonic dopamine levels, the impact of which on medial frontal cortex function has been well explored in rats. A wealth of evidence indicates that tonic dopamine levels strongly modulate the working memory functions of the medial frontal cortex (Floresco & Magyar, 2006; Seamans & Yang, 2004). Tonic dopamine levels have been proposed to code for average reward rate (Niv, 2007; see also Beierholm et al., 2013), which appears to facilitate the decision-making functions of the medial frontal cortex (Floresco, 2013). Consistent with this idea, reward-related dopamine release in the medial frontal cortex is critical to the acquisition of novel goal-directed behavioral strategies (i.e., options; Stark, Rothe, Wagner, & Scheich, 2004; Stefani & Moghaddam, 2006), and depends on the relative amount of work required to get the reward (Richardson & Gratton, 1998). Both D1 and D2 receptors in the prefrontal cortex also contribute to the formation of long-term memories (Blond, Crépel, & Otani, 2002; Puig & Miller, 2012; Sheynikhovich, Otani, & Arleo, 2011; Xu et al., 2009; Young & Yang, 2005), as well as to rapid updates of network connectivity (Arnsten, Paspalas, Gamo, Yang, & Wang,

2010), perhaps providing the means for the ACC to learn option values.

Accordingly, option values V for levels $L = 2, 3$ are determined by the average reward received across trials while that option is selected, updated at the end of each trial according to

$$V_{L=2,3}(o_L) \leftarrow \nu R + (1 - \nu) V_{L=2,3}(o_L) \quad (6)$$

where R is the total reward received on that trial and ν is a constant that determines the average reward learning rate ($0 \leq \nu \leq 1$). State notation is removed for convenience, as the update rule is applied at trial completion independently of the specific state s . Note that o_L corresponds to one of two available options for level $L = 2$ (e.g., response vs. place in Cross-Maze simulation) and to only one metaoption for level $L = 3$ (i.e., engage in task or not). Effectively, Level 2 of the hierarchy maintains separate average reward values for each option, whereas Level 3 of the hierarchy maintains the average reward value across options.

This proposal holds that the ACC (or prelimbic cortex) associates each option (or metaoption) with the average reward received across trials during the execution of that option (or metaoption) to learn about task (or experiment) value rather than about action (or option) values. The idea dovetails with evidence in rats (Sul, Kim, Huh, Lee, & Jung, 2010), as well as in primates (Amiez et al., 2006; Holroyd & Coles, 2008; Kennerley et al., 2006; Seo & Lee, 2007), that the ACC is concerned with integrating reward values across multiple trials rather than in instigating trial-to-trial adjustments in behavior.

The second effect of reinforcement is to adjust the level of control applied by each level over its immediately lower level. In the model, the ACC and the prelimbic cortex rely on a reward-dependent feedback loop to determine the minimal amount of control needed to maintain high average reward: Top-down control increases when received rewards are worse than expected and decreases otherwise (as observed in a study with human participants; Venables & Fairclough, 2009). In this way, the ACC and the prelimbic cortex relax regulatory control when events unfold smoothly and boost control when they do not, with a stable equilibrium point at minimum control required to maximize reward.

To be specific, for levels $L = 2, 3$, the control term ϵ_L in Equation 2 is initialized at the start of each block of trials with value ϵ_{max} (> 0) and updated at the end of each trial according to the prediction error Δ_L as

$$\begin{aligned} \Delta_L &= R - V_{L=2,3}(o_L) \\ \epsilon_L &\leftarrow \begin{cases} [\epsilon_L - \beta, 0]_{MAX}, & \text{if } \Delta_L \geq 0 \\ [\epsilon_L + \beta', \epsilon_{max}]_{MIN}, & \text{if } \Delta_L < 0 \end{cases} \end{aligned} \quad (7)$$

where β and β' are constants ($\beta' > \beta > 0$). This formulation ensures that the control for level L decreases linearly to zero when the total reward received on a given trial is equal to or exceeds the average reward associated with the option in play at that level, and increases linearly to a maximum of ϵ_{max} otherwise. For $L = 1$, ϵ_L is undefined (as the bottom level cannot apply control over a still lower level, although it may be plausible that the striatum may differentially regulate motor cortex activity based on expected reward).

Maximal control thus begins afresh at the start of each block of trials. For a novel task, this refresh enables the system to

explore the problem space and determine the maximum average amount of reward that it can expect to receive. If the organism continues to achieve the expected reward, then control is gradually disengaged. However, if received rewards are less than expected, then control is reinstated. In this way, effortful control over goal-directed behavior is recruited only when successful task performance is necessary for high payoff. These features are supported by observations that neuronal activity in the rat ACC begins each experimental session high (Hillman & Bilkey, 2012), and then decays to the minimal threshold necessary to sustain adequate task performance (Hillman & Bilkey, 2010), suggesting that the ACC is responsible for “staying the course” toward achieving long-term goals (Hillman & Bilkey, 2013; Williams, Mohler, & Givens, 1999). In humans, this formulation likewise resonates with observations that ACC activity is responsible for task persistence in the face of adversity (Kurniawan et al., 2010; Parvizi et al., 2013), especially when ongoing performance is predictive of high payoff (Dixon & Christoff, 2012; Shenhav et al., 2013).

Simulation Details

We conducted seven sets of simulations associated with four tasks, conducted across 13 experiments in five studies. For each experiment, simulated performance data were averaged across groups of 100 subjects separately for the sham and lesion conditions. Each experiment consisted of an initial habituation phase during which option selection was equiprobable (produced by setting $\tau_L = 1,000$), followed by a subsequent training phase and the experiment proper. The experiments were characterized by separate blocks of trials, the number of which depended on the specifics of each experiment. For each subject, values $V_L(o_{L+1}, o_L, s)$ for option-specific actions ($L = 1$), and $V_L(o_L)$ for options ($L = 2$) and for the metaoption ($L = 3$), were initialized at zero at the start of each experiment. Further, the control signals ϵ_L for levels $L = 2, 3$ were reinitialized at the start of each block of trials as $\epsilon_L = \epsilon_{max}$.

Table 1 lists the parameter values across the seven sets of simulations. The parameter values were determined using evolutionary optimization techniques (Ashlock, 2010) that minimized the sum of squared errors between simulated and empirical behavior. Because of the high computational demands of the simulations, these parameters were optimized, at most, four at a time. This approach necessitated fixing some parameters at reason-

able—though, by necessity, somewhat arbitrary—values when optimizing over other parameters. In particular, values for ξ , ϵ_{max} , β , and β' were all set by fiat (β' was fixed across studies as a multiple of β), whereas the optimization procedure centered on obtaining values for α , γ , τ_1 , τ_2 , and v . To the extent possible, the same parameter values were utilized across all of the experiments. It was not possible to find a single set of values that accounted for the two cross-maze experiments as well as the two barrier maze and their related experiments (delay maze and water maze). Thus, α , ξ , τ_1 , and v differed across these two sets of simulations. Further, β (and β') was increased for the tonic dopamine-related simulations. Note that the task demands of the cross-maze and the barrier maze tasks are quite different from one another—which is why the tasks were, in fact, selected for simulation—so it is unsurprising that some simulation parameters differed between experiments.

Results

Cross-Maze Task

Our first goal was to illustrate the essential role of the rat medial frontal cortex over hierarchical behavior. To do so, we simulated performance of rodents on a cross-maze task in which two superordinate task goals—appropriate for the “place” and “response” conditions of the experiment—were alternately required for successful performance (Ragozzino, Detrick, & Kesner, 1999). On each trial in the place condition, animals were started from either of two arm locations of the cross-maze and were required to find a reward located in a third arm that remained fixed across trials, that is, always in the north arm, irrespective of whether the rat started from the south or west arms. By contrast, on each trial of a response condition of the task, the animals were required to turn in a particular direction from the start arms with respect to their egocentric frame of reference, that is, to turn right, irrespective of whether they started from the west and east arms (see Figure 2). Rats were first trained on one task condition and were then retrained on the other task condition. Temporary inactivation of the prelimbic cortex was accomplished by infusion of a calcium channel blocker. When this lesion was imposed as rats learned the second task, it was observed to retard learning in treatment animals relative to sham controls, regardless of which task was learned first (Figure 3, top left, A vs. B). However, if the drug was administered as animals learned the first task condition, then it had no effect on

Table 1
Parameter Values for Each Experiment

	α	γ	ξ	τ_1	τ_2	v	ϵ_{max}	β	β'
Barrier	0.8	0.92	1.0	0.7	0.18	.67	6.0	.5	1.0
BarRev	0.8	0.92	1.0	0.7	0.18	.67	6.0	.5	1.0
Delay	0.8	0.92	1.0	0.7	0.18	.67	6.0	.5	1.0
BarDop	0.8	0.92	1.0	0.7	0.18	.67	6.0	.9	1.8
Water	0.8	0.92	1.0	0.7	0.18	.67	15.0	.5	1.0
Cross	0.2	0.92	.5	0.2	0.18	.96	6.0	.5	1.0
CrossDop	0.2	0.92	.5	0.2	0.18	.96	6.0	1.2	2.4

Note. Rows = experiments; columns = parameters. Barrier = barrier maze task; BarRev = barrier maze task with reversal; Delay = delay maze task; BarDop = barrier maze with ACC dopamine disruption; Water = water maze task; Cross = cross-maze task; CrossDop = cross-maze task with prelimbic dopamine disruption.

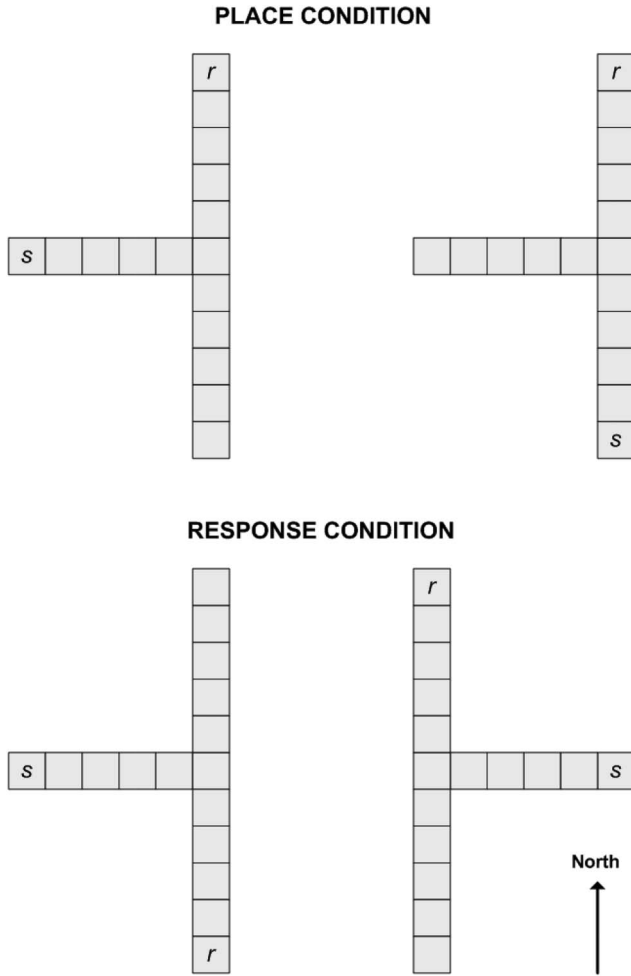


Figure 2. Example maze configurations for the place and response conditions of the cross-maze task. In the place condition (top), the agent starts each trial at location s in either the west or south arms and the reward occurs at location r in the north arm; irrespective of starting location, the east arm is blocked. In the response condition (bottom), the agent starts each trial at location s in either the west or east arms and the reward occurs at locations r , that is, at the end of the right arm relative to the starting location; irrespective of starting location, the forward arm is blocked. Arrow indicates north for all trial types across both place and response conditions.

performance when learning the second task (Figure 3, top left, A vs. C; Ragozzino et al., 1999). The impaired ability to shift to the new task strategy resulted from a tendency to perseverate on behaviors learned during the first task condition (Figure 3, middle left), and was selective for shifts to a new task. In particular, the temporary prelimbic lesions had no impact on reversal learning within a given task (Figure 3, bottom left): Lesioned rats were capable of relearning, but were selectively impaired in switching task goals (i.e., switching between actions determined by place or response).

Figure 1 illustrates the model design. On each trial, the ACC selected either an option for executing the task according to a place strategy, or an option for executing the task according to a response strategy (see also Caluwaerts et al., 2012, for a

similar model). The striatum learned to navigate individual steps of the maze on the basis of separate sets of option-specific state-action values (one set for each of the place and response options), supported by top-down control from the ACC. The ACC selected between the options by weighting their learned average reward values against an assumed cost of switching from one task to the other, which biased the ACC to favor continued selection of the same option. Meanwhile, whereas the ACC learned the average reward value of each option separately, the prelimbic cortex learned the average reward of the experiment itself, across options. The prelimbic cortex then applied a top-down control signal over ACC activity to overcome the switch cost to facilitate switching between options. The degree of top-down control was minimized by a feedback control loop that increased control when the received rewards were worse than average and decreased control otherwise.

We simulated performance on Experiments 1 through 4 of Ragozzino and colleagues (1999); trial sequences and maze specifics (e.g., reward magnitude for each trial) are described in the methods section of that article. The cross-maze was represented as a central cell from which four alleys of five cells radiated outward in each of the cardinal directions (Figure 4a). Possible start and end locations are indicated in Figure 4a by Xs, which varied from trial to trial depending on the experiment and condition, as detailed in Ragozzino and colleagues (1999). Only two possible alley choices were available to the agent on a given trial (Ragozzino et al., 1999); Figure 2 provides example maze configurations for the different conditions. Optimal performance across conditions depended on use of both place and response strategies, so $n_{L=2} = 2$ (Equation 1), yielding two separate sets of $11 \times 11 \times 5$ option-specific actions values. On trials in which the agent selected the “response” option, the gridworld coordinates were rotated into an egocentric reference frame (simplified by fixing the start location in the south alley irrespective of the actual start location of the agent). A switch cost was incurred when the agent evaluated switching to the different option ($C_{L=2} = 1.0$, Equation 2), except for the first trial in each block of trials when there was no immediately prior trial from which to switch ($C_{L=2} = 0$). Reward values were $r = 1.5$ for every half-pellet of food found at the reward locations. The effect of prelimbic lesions on performance was simulated by setting $\epsilon_{L=3} = 0$.

This instantiation of the HRL theory accounts for the impaired switching observed following prelimbic lesions, as follows (Figure 3, right column). The prelimbic cortex averages the reward received across the entire experiment and applies increased top-down control to the ACC whenever this value decreases, as occurs following the task shift when error rates increase. In turn, this additional prelimbic control reduces the impact of the switch cost in the ACC option selection, which would otherwise lead to response perseveration associated with the first option strategy. Lesions to the prelimbic cortex remove this source of top-down control, resulting in impaired switching during the shift phase (Figure 3, top right) because of increased perseverations (Figure 3, middle right). By contrast, simulated prelimbic lesions spare reversal learning because the new response mappings are consistent with the superordinate goal associated with the previous option selected and maintained by the ACC (Figure 3, bottom right). In this case, choosing to go

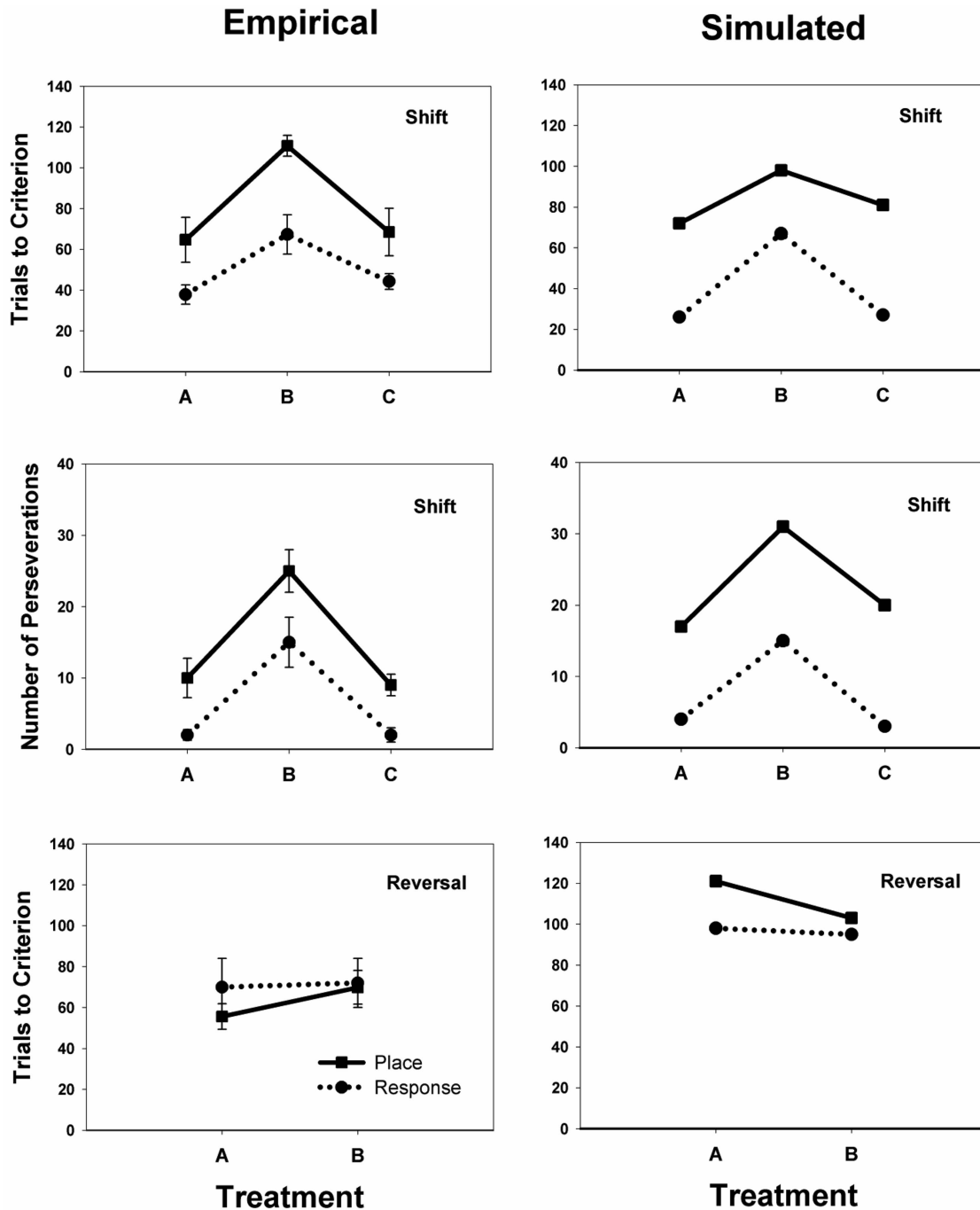


Figure 3. Empirical and simulated data of rats engaged in the cross-maze task. Rats were trained initially on either the place task (solid lines) or the response task (dotted lines). Empirical data are shown in the left column and simulated data are shown in the right column. Top panels: Number of trials to reach learning criterion when shifting to the new task. Middle panels: Number of perseverations on the previous task when shifting to the new task. Bottom panels: Number of trials to reach criterion when learning the reverse stimulus–response mappings associated with the first task condition. The abscissa indicates treatment condition. A = sham group; B = treatment group in which prelimbic cortex is inactivated during the second task condition; C = treatment group in which prelimbic cortex is inactivated during the first task condition. Error bars indicate standard error of the mean. Empirical data from [Ragozzino and colleagues \(1999\)](#); all data correspond to the second task condition of each experiment, that is, to the shift condition in the shift experiments and to the reversal condition in the reversal experiments. Prelimbic cortex lesions are simulated by setting the control value for the metaoption to zero.

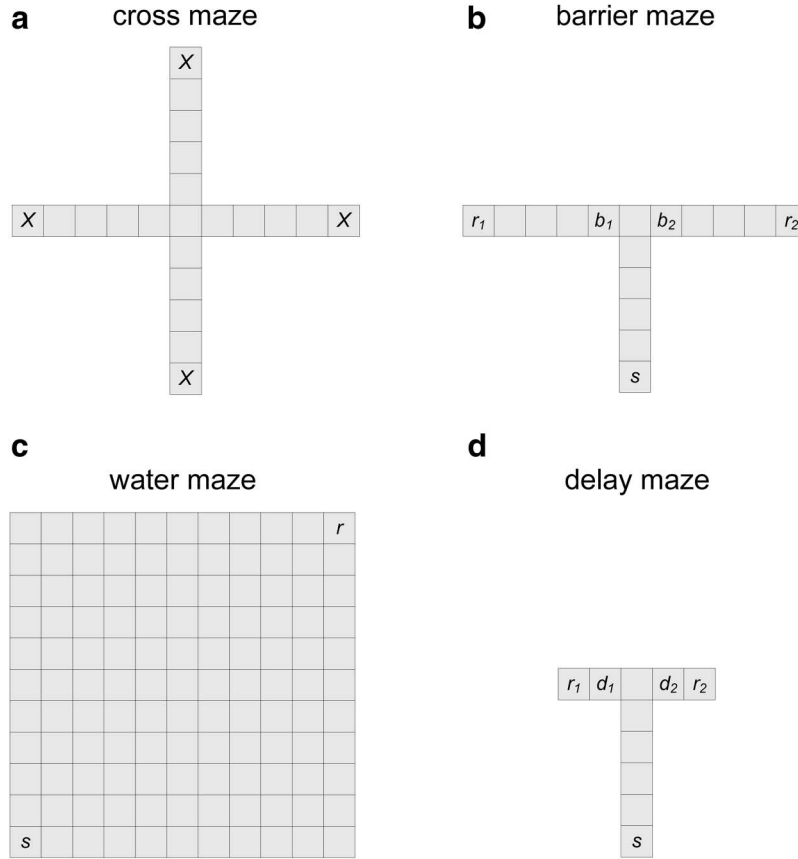


Figure 4. Simulation gridworlds. (a) Cross-maze (Ragozzino et al., 1999): X indicates possible start and reward locations that varied from trial to trial, as detailed in Ragozzino and colleagues (1999); only two alleys choices were available to the agent on a given trial. (b) Barrier maze (Walton et al., 2002): r_1 and r_2 denote reward amounts and associated locations for the left and right alleys, respectively; b_1 and b_2 denote barrier heights and associated locations for the left and right alleys, respectively; s indicates the start location at the stem of the maze. (c) Water maze: s and r indicate start and reward locations, respectively; moving in any direction in the maze incurred a small energetic cost. (d) Delay maze (Rudebeck et al., 2006): r_1 and r_2 denote reward amounts and associated locations for the left and right alleys, respectively; d_1 and d_2 denote delay lengths for the left and right alleys, respectively; s indicates the start location at the stem of the maze.

left instead of right, or north instead of east, requires learning new state-action values for the same “response” or “place” option, respectively. Unlike the switch condition, in which an increase in perseverative errors occurs following the switch from the previous (incorrect) task, in the reversal condition, perseverative errors are avoided because new state-action values are learned for the previous (correct) task.

The theory thus proposes a division of labor between prelimbic and ACC mechanisms for behavioral regulation. At one level of control, the ACC selects options for execution that have high expected value relative to the cost in switching between options. At a higher level, the prelimbic cortex learns the value of the metaoption as a whole and applies top-down control over the ACC that facilitates option selection. This organization ensures successful completion of the high-level system goal by enabling exploration over tasks: When the ACC selects an option that fails to meet the objective of its parent metaoption in the prelimbic cortex, then the prelimbic cortex commands the

ACC to try out alternative options that might better achieve its objective.

Finally, we explored the consequences of disruption of tonic dopamine levels in the prelimbic cortex. Because we assumed that this dopamine activity is encoded as the average reward value of the metaoption, we set this variable to zero (the decay rate β was also increased for both the sham and lesion simulations). Figure 5 (left column) illustrates the results of an empirical study in which rats in the cross-maze experiment were given a dopamine D1 antagonist (Ragozzino, 2002). Simulated performance on this task when the average reward value of the metaoption is set to zero is also shown (Figure 5, right column). As can be seen by inspection, capping this variable at a low value qualitatively reproduces the consequences of dopamine disruption in the prelimbic cortex, which, in turn, mirrors the effects of direct lesions to the prelimbic cortex itself (see Figure 3). In the model, artificially reducing the value of the metaoption causes the received rewards to be continually better than

Cross Maze: prelimbic cortex dopamine disruption

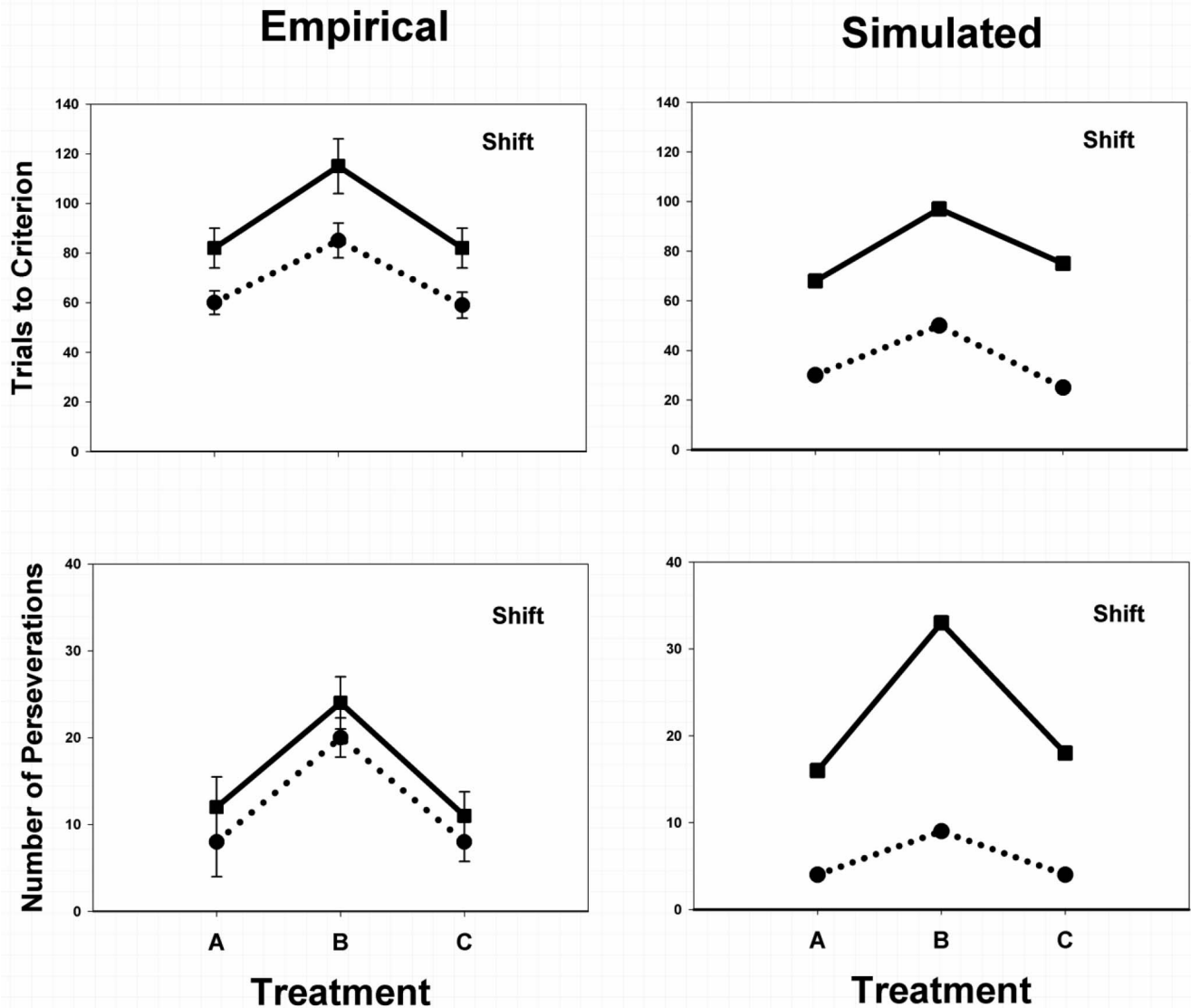


Figure 5. Empirical (left column) and simulated (right column) effects of dopamine disruption in the prelimbic cortex on cross-maze task performance. Data correspond to the second (shift) condition of each experiment; rats were trained on the place task and shifted to the response task (solid lines), or were trained on the response task and shifted to the place task (dotted lines). Empirical data associated with dopamine system disruption are shown in the left column and simulated data are shown in the right column. Top panels: number of trials to reach learning criterion when shifting to the new task. Bottom panels: number of perseverations on the previous task when shifting to the new task. The abscissa indicates treatment condition. A = sham group; B = treatment group in which the prelimbic cortex was infused with a dopamine D1 antagonist, SCH23390, during the second task condition; C = Treatment group in which the prelimbic cortex was infused with SCH23390 during the first task condition. Error bars indicate standard error of the mean. Empirical data from [Ragozzino \(2002\)](#). Note that the tasks in [Ragozzino \(2002\)](#) and [Ragozzino and colleagues \(1999\)](#) are characterized by minor methodological differences, and task performance was simulated using the maze design described in [Ragozzino and colleagues \(1999\)](#). Reduced dopamine levels in the prelimbic cortex were simulated by setting the average reward value for the metaoption to zero.

the (incorrectly low) average, which allows control over task selection to dissipate even when control is needed (Equation 7).

Barrier Maze Task

Our second goal was to demonstrate that the ACC implements the same hierarchical principles as the prelimbic cortex, with the only difference being the nature of the controlled signal: Whereas the prelimbic cortex controls the switch costs associated with option selection by the ACC, the ACC controls the effort-related costs associated with action selection by the striatum. To do so, we simulated the results of an influential study on rodents that demonstrated the essential role of the ACC in motivating effortful behavior (Walton et al., 2002). Rats in the experiment were required to navigate a “barrier maze” for reward. On each trial of the experiment, the rats were presented with a decision to either climb a barrier to receive a high reward in one arm of a T maze—incurring a large energetic cost—or to obtain an unobstructed but low reward in the other maze arm. As revealed in Figure 6 (solid lines), in Condition A (white background), the rats initially pre-

ferred to climb over the large barrier to receive the high reward. However, following surgery after Block 6, rats with ACC lesions eschewed the high-reward arm in favor of the low-reward arm (thick solid line). When an additional barrier was placed in the low-reward arm (Condition B, gray background), these rats again preferred the high-reward arm, demonstrating continued sensitivity to reward values when energetic costs were equated across choices. Subsequently, in Condition C (dotted background), the barrier in the high-reward arm was reduced in height, attracting the rats to that arm with increased probability. Finally, when the difference in the reward value between the two arms was increased from 4:2 pellets (Conditions A to C) to 5:1 pellets (Condition D, hatched background), preference for the high-reward arm also increased. By contrast, the sham-lesioned rats (thin solid line) chose the high-reward arm throughout all of the experimental conditions.

We simulated performance on Experiments 1 through 4 of the study by Walton and colleagues (2002). Trial sequences and maze specifics (e.g., reward magnitude and barrier height for each condition) are described in the methods section of that article. The T maze was represented within an 11×11 gridworld as a stem consisting of six cells with two alleys, five cells each, radiating perpendicularly from the stem end (Figure 4b). Rewards (r) were delivered at the end of each alley, in which $r = 1.5$ for each pellet of food. Barrier placement in each alley was associated with the unit immediately adjacent to the maze junction (Figure 4b), represented as a penalty $C_{L=1} = 4.5$ for 30-cm barriers when the agent evaluated entering that cell (Equation 2). Values of r and $C_{L=1}$ were scaled linearly for different reward amounts and barrier heights, respectively (e.g., $r = 3$ for two pellets of food). The effect of ACC lesions on performance were simulated by setting $\epsilon_{L=2} = 0$ —that is, fixing the ACC control signal over striatal action selection to zero. Note that there is no equivalent to the option-level choice in the T maze as in the cross-maze task previously (i.e., place vs. response options). We therefore set $n_{L=2} = 1$ (Equation 1) and $C_{L=2} = 0$ (Equation 2), yielding a single $11 \times 11 \times 5$ set of option-specific primitive action values.

The model accounts for the effects of ACC damage in terms of the relative values associated with task execution at different levels of abstraction. Action selection in the model is driven by the striatum, which weights the immediate costs of executing actions against their predicted benefits. By itself, we assume that this cost–benefit comparison produces a strong preference for low rewards that are easily obtained over high rewards that demand a large energetic investment (Condition A, white background). ACC lesions expose these preferences (simulated in the model by setting the control term to zero; thick dotted line following ACC ablation). Of course, when the costs of the two choices are equated (Condition B, gray background), the size of the barrier is reduced (Condition C, dotted background), or the reward differential between arms is increased (Condition D, hatched background), the cost–benefit analysis performed by the striatum tilts in favor of the high-reward arm. By contrast, when functioning normally, the ACC learns the average reward value associated with task execution as a whole and applies top-down control over the striatum as necessary to maintain the higher reward rate associated with the larger reward. We model this by assuming that the ACC begins each session with sufficient top-down control to the striatum to overcome the energetic cost of climbing over the barrier, propel-

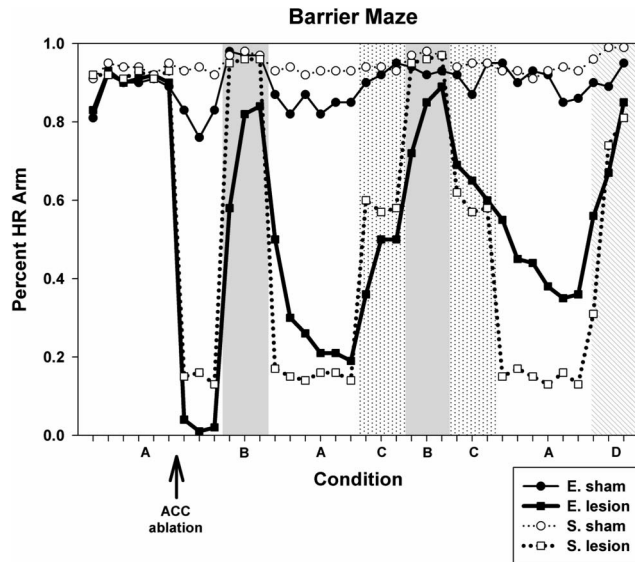


Figure 6. Empirical (solid lines) and simulated (dotted lines) data of rats engaged in the barrier maze task (Walton et al., 2002). Across all conditions, rats choose between one arm containing a high reward (HR) and a second arm containing a low reward (LR). In Condition A (white background), the HR arm is obstructed by a large barrier and the LR arm is unobstructed. In Condition B (gray background), both arms are obstructed by large barriers. In Condition C (dotted background), the HR arm is obstructed by a medium-size barrier and the LR arm is unobstructed. In Condition D (hatched background), the differential value between the HR and LR arms is increased. Ordinate indicates percent of trials in which rats select the HR arm. Abscissa indicates condition, where performance data for each condition are averaged according to blocks of 10 trials each. E. sham = empirical results for sham-lesioned animals; E. lesion = empirical results for animals with anterior cingulate cortex (ACC) lesions; S. sham = simulated results for sham lesioned animals; S. lesion = simulated results for animals with ACC lesions. Note that ACC lesions were induced during Condition A following the first six blocks of trials as indicated by the arrow, which was simulated in the model by setting the ACC control term to zero. Empirical data from Walton et al. (2002).

ling the system to obtain the high reward (thin dotted line). The minimum level of top-down control necessary to maximize overall reward is then automatically determined by a feedback control loop: Control decays while average reward remains high, enabling the striatum to explore the low-reward arm, but the resulting decrease in the average reward precipitates increased control by the ACC, which returns the system to a state of effortful but more rewarding activity.

In parallel to our analysis of the cross-maze experiment, we also considered the effect of simulated disruption to tonic dopamine levels in the ACC. To do so, we set the average reward value of the selected option to zero (the decay rate β was also increased for both the sham and lesion simulations). As can be seen in Figure 7, physical disruption of dopamine levels in the ACC by local application of 6-hydroxydopamine impairs performance on the task in a manner similar to direct ablation of the ACC (Schweimer, Saft & Hauber, 2005). Further, artificially reducing the value of the selected option in the simulation produces a comparable effect. This result obtains because all of the received rewards are then better than the (incorrectly low) model-estimated average reward, permitting control levels to plummet (Equation 7).

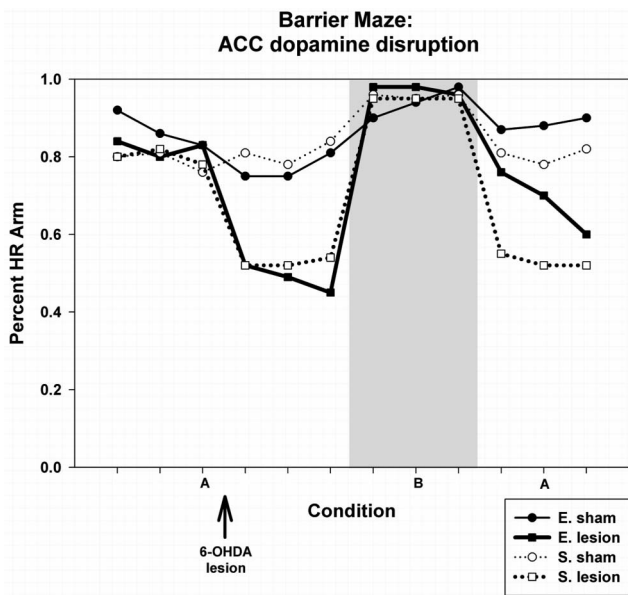


Figure 7. Empirical (solid lines) and simulated (dotted lines) effects on barrier maze performance of reduced dopamine levels in the anterior cingulate cortex (ACC) as reported in Schweimer and colleagues (2005). Across all blocks, one arm of the T maze contained a high reward (HR) and the other arm contained a low reward. Ordinate indicates percent of trials that subjects chose the HR arm; abscissa indicates block. Block A (white background) corresponds to trials in which the HR arm was obstructed with a high barrier, and Block B corresponds to trials in which both arms were obstructed with high barriers. Dopamine processing was disrupted locally in the ACC by application of 6-hydroxydopamine (6-OHDA) neurotoxin at the time indicated by the arrow. This effect was simulated by setting the value for the selected option to zero. E. sham = empirical results for sham-lesioned animals; E. lesion = empirical results for animals with 6-OHDA ACC lesions; S. sham = simulated results for sham-lesioned animals; S. lesion = simulated results for animals with 6-OHDA ACC lesions.

The barrier maze results are commonly interpreted to mean that the ACC is responsible for enabling the animal to work for reward (Walton et al., 2002). Our results instead point toward a more nuanced explanation (see also Walton, Rudebeck, Bannerman, & Rushworth, 2007): The ACC encodes a course of action at the level of task representation and applies sufficient control to ensure its successful completion. Without the ACC's input based on the higher level value of performing the task, action selection is dominated by simple actions mediated by the basal ganglia that are reflexive, effort-averse, and characterized by immediate gratification.

Critically, the ACC facilitates exploration of actions within a task in parallel to how the prelimbic cortex facilitates the exploration of task strategies within the experiment. To illustrate this point, we simulated a novel, reversal learning condition of the barrier maze task. The simulation details and parameter values were identical to Condition B of the barrier maze task, in which both arms were blocked by high barriers (as described earlier; Figure 6), except that the high-reward arm contained seven pellets and the low-reward arm contained zero pellets (instead of 4:2). Following an initial exploration phase with no barriers, the rats engaged in 20 trials of the barrier maze task (high reward in Arm 1 and no reward in Arm 2, both arms blocked), after which the reward location was suddenly switched to the other arm (no reward in Arm 1 and high reward in Arm 2, both arms blocked). As before, ACC lesions were simulated by setting the control term for the option to zero.

Figure 8 illustrates the model predictions before and after the reversal. Prior to the reversal (top panel), the model reproduces the results seen previously in Condition B of the barrier maze task (see Figure 6): When barriers are placed in front of both arms, both the simulated animals with sham lesions and the animals with ACC lesions prefer the high-reward arm (Arm 1) almost exclusively. As before, the lesion group chooses the high-reward arm despite the cost incurred by the barrier for two reasons: first, because the costs between the two choices are equated (Figure 6, Condition B), and second, because the reward values between the arms are skewed to relatively extreme values (Figure 6, Condition D), rendering the choice for the high-reward arm particularly obvious. By contrast, the two groups exhibit qualitatively different behaviors after the reversal (Figure 8, middle panel). When the high reward is suddenly moved to Arm 2, the simulated control rats shift their preference to that arm. However, instead of choosing one arm over the other, the rats with ACC lesions hesitate and choose neither—even though they are entirely capable of climbing over the barriers, which they readily surmounted before the switch.

Insight into this dissociation is provided by the simulated control levels (Figure 8, bottom panel). Before the reversal, the animals with ACC lesions climb the barrier in Arm 1 to get the high reward, which is achieved on the basis of the striatal efforts alone. Following the reversal, they quickly learn that Arm 1 no longer contains the reward, but without the support from the ACC are unmotivated to explore Arm 2, which they remember as being empty. So the rats with ACC lesions choose neither alternative. By contrast, the sham-lesioned animals also climb the barrier in Arm 1 to get the reward before the reversal, and because the striatum can satisfactorily perform this task on its own, the control level in the sham group gradually decays to zero. But following the reversal, the sudden absence of rewards in Arm 1 precipitates a spike in

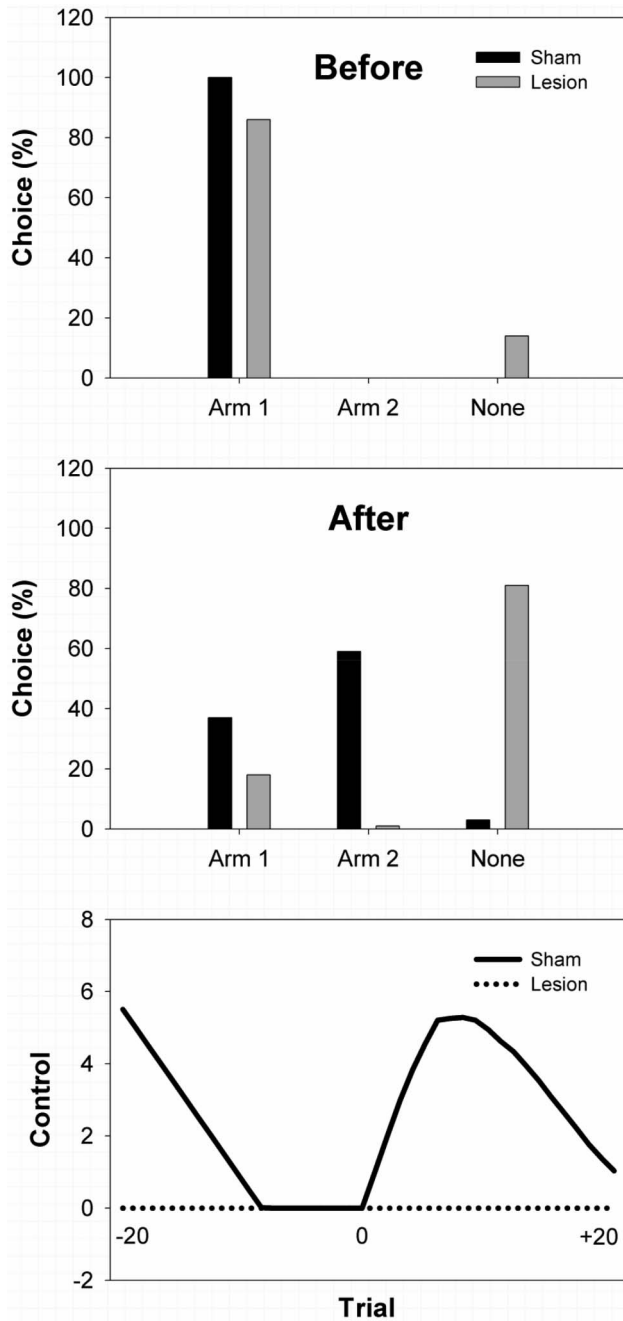


Figure 8. Simulated performance on the barrier maze reversal task. Top panel: Percent choices from Arm 1, Arm 2, and no choice (None) during 20 trials before reversal. Middle panel: Percent choices from Arm 1, Arm 2, and no choice (None) during 20 trials after reversal. Bottom panel: Control levels for 20 trials before and after the reversal. Zero on abscissa indicates time of the reversal. Units of control are arbitrary. Sham = simulated control rats; Lesion = Simulated rats with ACC lesions.

control that motivates the striatum to explore Arm 2. Once the striatum finds the new reward location and learns the new set of action–reward contingencies, the ACC again gradually withdraws control. The sham group succeeds where the lesioned group fails

because the ACC has come to associate the task itself with a high reward value. When that high expectation is not met, the ACC, in effect, assumes that the problem must have an alternative solution—and is willing to put the work in to find it. By contrast, without knowledge of the higher context in which the rewards are delivered, the striatum assumes the truth of what it has learned: that neither of the arms contains any reward.

Water Maze Task

The pseudoparalysis of the rats with ACC lesions in the barrier maze simulation points to the source of akinetic mutism, which we highlight here by simulating rat performance on an additional, imagined test. The simulated rats were required to move to a small, distant reward in a “water maze” that was implemented as an 11×11 gridworld, as illustrated in Figure 4c. The agent started each trial at location s and was required to find a reward ($r = 6$) located cater-cornered in the maze. The maze exacted an effort penalty, $C_{L=1} = .5$, for all actions except sitting (i.e., move north, south, east, or west), and $C_{L=1} = 0$ for sitting. The experiment consisted of a three-block habituation phase ($\tau_1, \tau_2 = 1000$), followed by 12 blocks of 10 trials each for the experiment proper. The ACC was lesioned by setting the option-level control signal to zero ($\epsilon_{L=2} = 0$). Results are illustrated in Figure 9. Under the assumption that all movements through the maze exact small energetic costs, simulated ACC lesions reduced the number of trials that the rat successfully obtained the reward because of an increased propensity to sit.

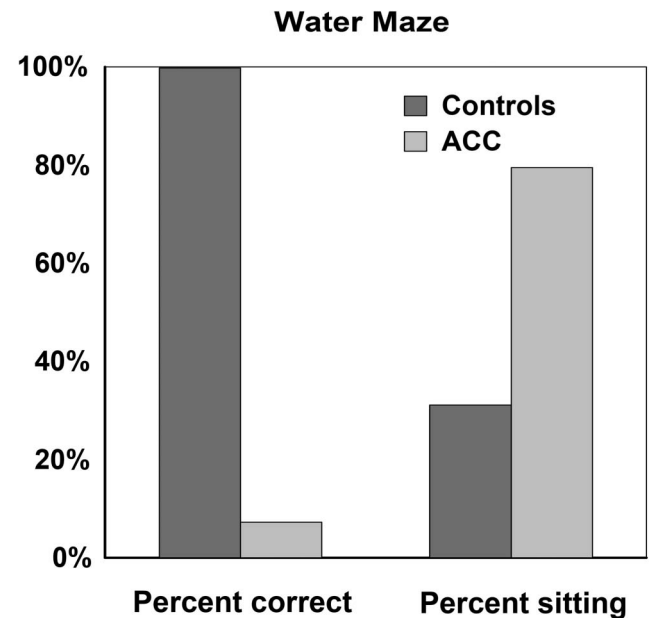


Figure 9. Simulated effects of anterior cingulate cortex (ACC) lesions on performance of a hypothetical water maze task. Left: percent trials that the animal reaches the reward location relative to the total number of trials. Right: percent of actions that the animal spends sitting relative to the total number of actions taken. Dark: sham control simulations. Light: ACC lesion simulations. ACC lesions cause a significant decrease in rewarded activity because of increased overall sitting.

The theory thus proposes a division of labor between striatal and ACC mechanisms for behavioral regulation. At one level of control, the striatum selects actions for execution that have high expected value relative to their energetic costs. At a higher level, the ACC learns the value of tasks as a whole and applies top-down control over the striatum to ensure successful task completion. This organization provides a scaffolding that supports extended behaviors, particularly when energetic costs would otherwise interfere with performance on the basis of striatal value signals alone. This account answers why ACC damage can result in akinetic mutism: Given that all actions require at least some physical effort, the absence of an organizing goal for action selection, together with weak distal rewards, can lead to a pseudoparalysis in which the organism fails to move despite being able to do so.

Delay Maze Task

A final simulation rules out a possible confound, which is that impaired performance on the barrier maze task following ACC lesions results not from disruption to a hierarchical representation of the task, but rather to increased temporal discounting of future rewards. In principle, rats with ACC damage might eschew large rewards, requiring high effort if the values of the rewards were highly discounted. To demonstrate that this is not the case, Rudebeck, Walton, Smyth, Bannerman, and Rushworth (2006) revealed a double dissociation between the functions of the orbitofrontal cortex (OFC) and the ACC on delay discounting versus barrier versions of the T-maze task. On each trial of the delay discounting task, rats were again required to choose between two arms of a T maze containing high and low rewards; however, the rats had to first submit to a several-second delay to obtain the high reward. Whereas ACC lesions dramatically impaired performance on the barrier maze but not on the delay maze, OFC lesions impaired performance on the delay maze but not on the barrier maze (see Figure 10). The authors concluded that decision making in the rat is influenced by two independent neural systems that separately process delay- and effort-related costs.

To address this issue, we simulated rat performance on the delay maze and on the barrier maze in Experiments 1 and 2 of Rudebeck and colleagues (2006). Although our model does not explicitly represent the OFC, we accounted for the effect of increased temporal discounting related to OFC damage by decreasing the time discount parameter (γ , Equation 4), thereby reducing the effective strength of the reinforcer at times preceding reward delivery. Trial sequences and maze specifics (e.g., reward magnitude, barrier height and delay length for each trial) are described in the methods section of Rudebeck and colleagues. The barrier maze (Experiment 2 in Rudebeck et al.) was unchanged from the previous simulation (Figure 4b). The delay maze was represented as a six-cell-long stem joined perpendicularly at one end to two alleys, each two cells long (Figure 4d). Subjects were detained between 0 and 20 s when passing through units (d), represented abstractly as additional cells that the subject automatically traversed to reach the reward (seven additional cells per each 5-s increment in delay). For both the barrier and delay mazes, rewards (r) were delivered at the end of each arm, with $r = 1.5$ for 1 pellet of food. Values of r and $C_{L=1}$ (for the barrier maze) were scaled linearly for different reward amounts and barrier heights, respectively. The effect of

ACC lesions on performance were simulated by setting $\epsilon_{L=2} = 0$, as previously. The effect of OFC lesions was simulated by multiplying the discount parameter γ by .9.

Reducing the control term in the simulation spared delay discounting but profoundly disrupted effortful behavior, whereas reducing the reward discounting parameter produced the converse result (see Figure 10). This result indicates that reducing the discount parameter can impair performance on a delayed reward task while sparing performance on an effort-related task, and further, that reducing the control parameter can impair performance on the effort-related task while sparing performance on the delay task. Our conclusion therefore agrees with that of Rudebeck and colleagues (2006): The organizing mechanism implemented by the ACC can be distinguished from the discounting of future rewards, which determines the impact of predicted rewards on present behavior irrespective of the contextual milieu.

Discussion

Holroyd and Yeung (2012) recently proposed the broad thesis that the ACC functions to motivate extended, goal-directed behaviors according to principles of HRL. Here, we formalized these ideas in a series of simulations that makes explicit the theory's underlying assumptions while demonstrating its internal consistency. Although the model simulated rat behavior specifically, its structure was informed by findings from both rat and primate literature. This effort advances current thinking on the frontal medial cortex and ACC function along multiple avenues. First, going beyond a few previous models that tentatively suggested that the ACC carries out an actor-related RL role over behavioral

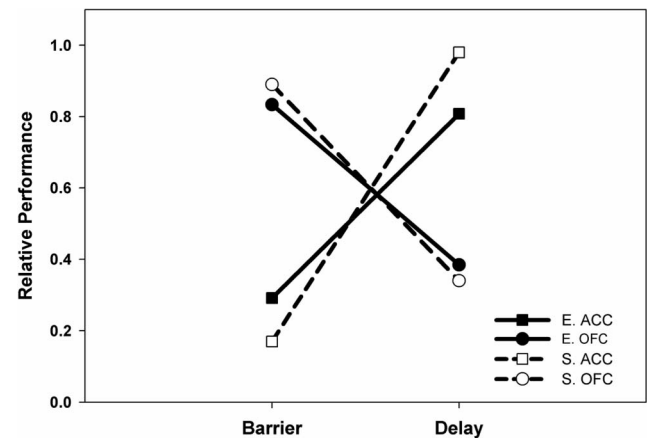


Figure 10. Empirical (solid lines) and simulated (dashed lines) data of rats engaged in barrier- and delay- versions of the T maze task. The ordinate indicates the proportion of high reward arm choices for anterior cingulate cortex (ACC)/orbitofrontal cortex (OFC) lesioned animals to high reward arm choices for sham-lesioned animals. E. ACC = empirical data for rats with anterior cingulate cortex lesions; E. OFC = empirical data for rats with orbitofrontal cortex lesions; S. ACC = simulated data for rats with anterior cingulate cortex lesions; S. OFC = simulated data for rats with orbitofrontal cortex lesions. The effects of ACC and OFC lesions were simulated by reducing the control and discount parameters, respectively. Note the double dissociation between the effects of ACC and OFC damage on the barrier and delay versions of the T maze, for both the real and simulated subjects. Empirical data from Rudebeck et al. (2006).

choice (e.g., Holroyd & Coles, 2002, 2008; Holroyd, Yeung, Coles, & Cohen, 2005), the model provides a computationally specific account of the idea that the ACC is responsible for guiding decision making by associating actions with their outcomes (Rushworth et al., 2007; Walton, Rudebeck, et al., 2007). By contrast, existing computational theories have proposed that the ACC is more concerned with performance monitoring than with action generation (e.g., W. H. Alexander & Brown, 2011; Botvinick, 2007; Botvinick et al., 2001; J. W. Brown & Braver, 2005; Shenhav et al., 2013; Silvetti et al., 2011, in press; but for recent actor-related alternatives, see Khamassi et al., 2011; Zendeirouh, Gharibzadeh, & Towhidkhan, 2013, 2014), or have discussed hierarchical function without reference to the ACC (Dezfouli & Balleine, 2012; Frank & Badre, 2012; Keramati & Gutkin, 2013). Second, the model demonstrates how the ACC can assist with complex RL problems that other neural systems (such as the striatum) might otherwise solve on their own (Holroyd & Yeung, 2012). Third, the model accounts for the behavioral consequences of ACC damage—namely, that damage sometimes results in slowed responding and reduced behavioral output despite otherwise normal motor ability—whereas functions proposed by other theories are mostly spared by such damage (Holroyd & Yeung, 2012). And fourth, the simulations indicate a mechanism for associating abstract task representations underlying cognitive control (e.g., J. D. Cohen et al., 1990; Kimberg & Farah, 1993) with reward signals that specify which task to execute as well as how vigorously the task should be carried out (Umemoto & Holroyd, 2014). In so doing, the model links the ACC literatures on cognitive control (Botvinick et al., 2001) and RL (Holroyd & Coles, 2002; Rushworth et al., 2007; Walton, Rudebeck, et al., 2007) by providing what is, to our knowledge, the first mechanistic demonstration of how principles of RL can regulate the deployment of control. We hope that these efforts provide a foundation for further development and a benchmark for comparison with other competing models.

Hierarchy and Control

The model develops an earlier hierarchical proposal that the ACC uses dopamine-related reward signals to select between action policies (Holroyd & Coles, 2002) in the context of recent explorations by Botvinick and colleagues into whether biological systems implement principles of HRL to solve difficult control problems (Botvinick, 2012; Botvinick, Niv, et al., 2009; Diuk, Schapiro, et al., 2013; Diuk, Tsai, Wallis, Botvinick, & Niv, 2013; Ribas-Fernandes et al., 2011). Based on apparent commonalities between the concepts of options in HRL and task representations mediated by the DLPFC, Botvinick and colleagues suggested that the DLPFC serves as a mechanism for option selection and maintenance. Our proposal is similar to their account, but suggests that the ACC, rather than the DLPFC, is responsible for option selection and maintenance (Holroyd & Yeung, 2011, 2012). Notably, Botvinick and colleagues observed HRL signals in the ACC in a human neuroimaging study but did not discuss this result in the context of ACC function (Ribas-Fernandes et al., 2011), perhaps because the ACC does not play a specific role in their brain-based theory of HRL (Botvinick, Niv, et al., 2009).

In a separate line of research, Shenhav et al. (2013) recently proposed an elegant theory that the ACC encodes a quantity called

the “expected value of control” that is utilized by the DLPFC to regulate the application of top-down control over task execution. The expected value of control and HRL theories share some salient commonalities—for example, both theories propose that effortful control over goal-directed behavior is recruited by the ACC only when successful task performance is predictive of high payoff (Dixon & Christoff, 2012)—but only the HRL theory in its present form explicitly links the principles of control with hierarchy. Further, whereas the expected value of control theory incorporates effort-related costs into the control calculations (Shenhav et al., 2013), the HRL theory of ACC function excludes such costs when calculating option values. Whether such discrepancies reflect fundamental differences between the two theories or are merely incidental is difficult to ascertain. We suggest that the comparison could be facilitated if the expected value of control framework were also instantiated in a computational model.

That effort-related costs are not directly factored into the model’s action update equation (Equation 4) constitutes an especially strong model assumption, as it renders the system incapable of predicting future cost expenditures. Yet because these costs are evaluated in the action selection algorithm (Equations 1 and 2), the system tends to avoid selecting actions with high energetic costs when suitably valuable alternatives are available. In this way, the system does not predict upcoming effortful behaviors but avoids choosing such behaviors at the time of action selection. Note that were the effort-related costs actually learnable, then the striatal and ACC mechanisms for selection would never conflict: Costs that were high enough to deter the striatal mechanism from obtaining a large reward would also devalue the average reward associated with the superordinate option, resulting in reduced control and a bias toward less-rewarding outcomes—a possibility that is precluded by relegating the impact of the costs to the moment of selection. Despite this assumption, effort-related costs still affect behavior over the long term: When the gradual reduction of control by the ACC frees the striatum to avoid actions with high-effort-related costs, the resulting decrease in reward propels the ACC to increase control. This feedback–control loop can give the appearance that the ACC learns and applies effort-related costs prospectively when in fact it does not.

In any event, we assume that effort-related costs are unpredictable only as a heuristic; our formal position is that such costs are relatively difficult, if not necessarily impossible, to predict. This assumption could be relaxed by including a cost term in the action update equation with an especially small discount parameter (Equation 4), and a cost term in the option–metaoption reward learning equation with a relatively small learning rate (Equation 6), which would render the costs somewhat learnable.

The theory thus proposes that the evaluation of effort-related costs is qualitatively different from that of unconditioned stimuli (such as rewards and punishments) that are inherently learnable and predictable (Boksem & Tops, 2008). It has been observed that cost estimates of tasks associated with extended physical exertion or tolerating pain may be difficult to predict and might fluctuate (McGuire & Kable, 2013; see also Kahneman, 2000). Note that effort-related prediction entails not predicting the cost itself per se (e.g., the amount of pain), but rather the degree of control that the system can expect to muster in meeting the cost, as it is the latter rather than the former that determines whether or not the offending obstacle is actually overcome. Hence, in contrast to

predicting reward and punishment signals that arise from sources external to the agent, prediction of effort-related costs entails predicting the agent's own internally generated control signal in relation to an external quantity. These predictions are especially ambiguous precisely when the required control levels are more or less equal to the task at hand, just as an elite runner may not be able to predict whether or not they will win a particular marathon. We suggest that ACC function is especially necessary for sustaining a selected option through to its completion when the control demands over option execution are greater than anticipated. At minimum, the ongoing debate about the putative resource dependence of effortful control underscores the unique nature of this psychological construct (Inzlicht & Schmeichel, 2012; Inzlicht et al., 2014; Kurzban et al., 2013).

At the neural level, empirical support for this assumption is mixed. Similar to our suggestion, Phillips and colleagues have proposed that the acute effects of dopamine on behavior are determined based on a comparison between the predictive values of future reward (as coded by extrasynaptic dopamine concentrations) with the costs required to elicit the rewards, such that approach behavior is elicited only when the predicted reward value exceeds a cost-defined threshold (Phillips, Walton, & Jhou, 2007). Further, they argue that this comparison unfolds in the nucleus accumbens, where dopamine release biases behavior toward predicted rewards in the context of cost-related information conveyed from the ACC (Phillips et al., 2007; Walton, Kennerley, Bannerman, Phillips, & Rushworth, 2006). Substantial evidence that the nucleus accumbens is sensitive to such cost–benefit comparisons is partially supportive of this hypothesis (e.g., Botvinick, Huffstetler, et al., 2009; Cousins, Atherton, Turner, & Salamone, 1996; Cousins & Salamone, 1994; Kurniawan, Guitart-Masip, & Dolan, 2011; Ostlund, Wassum, Murphy, Balleine, & Maidment, 2011; Salamone, Correa, Farrar, Nunes, & Pardo, 2009; but see Walton et al., 2009).

More troubling with respect to the model—given that Equation 3 indicates that effort-related costs should not factor into state-value learning—are recent observations that phasic dopamine release to cues predictive of forthcoming rewards reflect discounting of the effort-related costs needed to obtain the rewards (Day, Jones, & Carelli, 2011; Day, Jones, Wightman, & Carelli, 2010; Kurniawan et al., 2013; Pasquereau & Turner, 2013). However, these effort-sensitive dopamine neurons constitute only a small proportion of the collective midbrain dopamine system, raising the question as to whether these signals are powerful enough to drive learning of extended control policies as opposed to simple cue–outcome associations (Pasquereau & Turner, 2013). Further, other studies have failed to find effort-related discounting of dopamine release in the rat nucleus accumbens (Gan, Walton, & Phillips, 2010; Wanat, Kuhn, & Phillips, 2010) and of BOLD activation of the human nucleus accumbens (Schmidt, Lebreton, Cléry-Melin, Daunizeau, & Pessiglione, 2012). Note that the model's action selection mechanism (Equation 1) factors effort-related costs into the decision *before* the action has actually been carried out (Equation 2), necessitating a mechanism that predicts effort-related costs over very short time intervals. In this context, observations that reward-predictive cues elicit phasic dopamine signals that are sensitive to effort-related costs (Day et al., 2010, 2011; Pasquereau & Turner, 2013) may actually reflect the acute energizing effects of dopamine on behavior, rather than a longer term

learning process that operates over extended sequences of actions (Pasquereau & Turner, 2013; Phillips et al., 2007).

We note that, in its current form, the model does not store previous context-specific control levels as a starting point for future sessions, nor does it provide a mechanism for recruiting control in the absence of negative prediction errors. Nevertheless, we argue that the simple control mechanism proposed here can successfully address many everyday challenges: So long as control begins in a novel context with a high value, then the system will gradually relax into a state that simultaneously maximizes average received reward while minimizing the level of control needed to achieve it (Hillman & Bilkey, 2010, 2012, 2013; Williams et al., 1999). Control in rewarding situations then need only be recruited when performance drops below the average value. This process could be facilitated by retrieving option-related reward values from long-term memory storage in the ACC (Euston et al., 2012; Jung et al., 2008; S. H. Wang et al., 2012). In this case, poor initial performance would boost the degree of control to a level that achieves a reward rate commensurate to that retrieved from long-term store.

Finally, we emphasize that the hierarchical control mechanism that we have proposed provides a means for exploring solutions to problems at multiple levels of abstraction simultaneously. To make this point concrete, imagine the following scenario: That you are traveling for the first time from your home to a new job located somewhere across the city. Assume that your objective is to arrive at work in a timely fashion, and that reward is earned as a linear function of the distance traveled toward your goal per unit time. In this case, the metaoption to arrive quickly at work would afford several specific options: driving, taking the bus, walking, hitchhiking, and so on. If you were to select and execute the option for driving, the average reward received might steadily increase, resulting in reduced control at both the option and metaoption levels—especially if you are a good driver and can handle the car relatively automatically. But if the vehicle were suddenly to break down, then the consequent decrease in reward rate would precipitate increased control at both the metaoption and option levels. Greater control over driving would be ineffectual, but the increased control by the metaoption would facilitate a shift to a different option, perhaps to taking the bus, and if that strategy failed, then to still another option such as walking. Once the next selected option is successful, the rising reward rate would stabilize the option as the metaoption withdraws control.

Moreover, in parallel to how the metaoption would facilitate exploration over options, the selected option would facilitate exploration over actions. To continue with the story, imagine that the initial leg of your walk from the bus stop begins by going downhill, such that you make rapid progress toward your goal and reward quickly accumulates. Because little effort is required to walk downhill, the option level would rapidly relinquish control over behavior. But at the bottom of the hill, the terrain levels out into a valley that stretches in the wrong direction. In the absence of control you might find yourself walking the long way around to your destination. Instead, the sudden decrease in reward rate would cause a spike in control from the walking option, encouraging you to explore alternative routes that might be more effortful—for example, by cutting across a hill with tall grass—but more direct. In this way, effortful control at both the option and metaoption

levels ensures thorough exploration of decision space, especially for naturalistic environments that afford multiple potential actions.

Functional–Neuroanatomical Homologies

Although we developed our model to account for studies of the rodent medial frontal cortex, we believe that the simulations inform understanding of the human medial frontal cortex. Comparison across species is complicated by the fact that putative homologies between rat and human medial frontal structures have been perennially controversial (Balleine & O'Doherty, 2010; V. J. Brown & Bowman, 2002; Laubach, 2011; Preuss, 1995; Seamans et al., 2008; Uylings, Groenewegen, & Kolb, 2003; Vogt, 2009; Vogt et al., 2013; Vogt & Paxinos, 2014; Wallis, 2012). An ongoing debate has centered on whether the rat medial prefrontal cortex is homologous to the human medial prefrontal cortex or to the human DLPFC (Laubach, 2011). In fact, the rat medial frontal cortex appears to combine aspects of both structures at a rudimentary level (Seamans et al., 2008). An old scheme defines the rat frontal midline as containing the ACC, which consists of areas Cg1, Cg2, and the dorsal ACC, as well as the prelimbic cortex and infralimbic cortex (Rushworth, Walton, Kennerley, & Bannerman, 2004). The prelimbic cortex, infralimbic cortex, and ACC are observed to cooperate for effecting goal-directed behavior, yet their functions are dissociable (e.g., Chudasama et al., 2003; Dalley, Cardinal, & Robbins, 2004; Dias & Aggleton, 2000; Killcross & Coutureau, 2003; Ragozzino, 2007). In primates, motor areas in the cingulate sulcus in the caudal ACC project directly to the spinal cord (Dum & Strick, 1996), basal ganglia (Takada et al., 2001), and other motor areas (Morecraft & Tanji, 2009). However, the rat frontal midline does not contain a cingulate sulcus, and so contains no cingulate motor areas per se (Vogt, 2009; Vogt & Paxinos, 2014; Vogt et al., 2013). Rather, the densest projections to the spinal cord (M. W. Miller, 1987) and basal ganglia (Gabbott, Warner, Jays, Salway, & Busby, 2005; Laubach, 2011) emanate from a pregenual area in the rat ACC corresponding to parts of Area 32 and Area 24. It has been suggested that this region, which corresponds approximately to the dorsal ACC and receives projections from the prelimbic cortex (Jones, Groenewegen, & Witter, 2005), resembles the primate motor areas in the ACC sulcus (Rushworth et al., 2004; Vogt et al., 2013). Although not explicitly modeled here, the proposed pathway from the prelimbic cortex to the striatum via the ACC may parallel a direct prelimbic–striatal pathway that facilitates task shifts (Goto & Grace, 2005).

Our assertion that the function of the rodent medial frontal cortex is organized along a rostral–caudal hierarchy of control is supported by evidence of such a hierarchy in the human frontal cortex (Badre & D'Esposito, 2007, 2009; Badre & Frank, 2012; Collins et al., 2014; Dixon et al., 2014; Jeon & Friederici, 2013; Koechlin & Summerfield, 2007; Kounieher et al., 2009; O'Reilly, 2010; Orr & Banich, 2014), but several studies have recently challenged this view (Crittenden & Duncan, 2014; Farooqui et al., 2012; Reynolds et al., 2012). Further, there is reason to question whether hierarchies of arbitrary depth can be mapped to different brain areas. Clearly the rigid association that we have proposed between three hierarchical levels and three neuroanatomical structures would be inflexible for many complex, real-world problems. Partly for this reason, we previously speculated that multilevel problems could be flexibly translated into series of two-level

problems that are recursively loaded into working memory (Holroyd & Yeung, 2011), but evidence now seems more consistent with the existence of a rostral–caudal hierarchy along the midline. Alternatively, the frontal midline might encode *emergent* solutions that are intermediate to these two extremes (Dixon et al., 2014): Hierarchical problems could be flexibly decomposed into their components and subcomponents along the rostral–caudal extent of the ACC to the degree of coarseness necessary to accommodate the entire depth of the hierarchy. The upper limit on the number of levels could be constrained by the availability of *stripes*—groups of interconnected neurons in the frontal cortex that have been estimated to number about 20,000 (Frank, Loughry, & O'Reilly, 2001).

Despite this debate, there is good reason to believe that regions of the medial frontal cortex share functional correspondences between rodents, nonhuman primates, and humans. As noted the ACC in rats (Cowen et al., 2012; Hauber & Sommer, 2009; Hillman & Bilkey, 2010, 2012; Walton et al., 2002) and the caudal ACC in primates (Botvinick, Huffstetler, et al., 2009; Croxson et al., 2009; Davis et al., 2000; Kurniawan et al., 2013; Mulert et al., 2008; Paus et al., 1998; Prévost et al., 2010) are both plainly sensitive to effortful costs, suggesting a functional homology between these areas across species. And as we have argued, lesions of this area in humans result in reduced motor output despite normal motor ability (Holroyd & Yeung, 2012), in apparent correspondence with the effect of ACC lesions on rat behavior (Walton et al., 2002). Moreover, detailed examination of ACC activity from rats (Balaguer-Ballester, Lapish, Seamans, & Durstewitz, 2011; Cowen et al., 2012; Cowen & McNaughton, 2007; Durstewitz, Vitoz, Floresco, & Seamans, 2010; Fujisawa, Amarasingham, Harrison, & Buzsáki, 2008; Hyman, Ma, Balaguer-Ballester, Durstewitz, & Seamans, 2012; Lapish, Durstewitz, Chandler, & Seamans, 2008; Ma et al., 2014; Mulder, Nordquist, Örgüt, & Pennartz, 2003; see also Ostlund, Winterbauer, & Balleine, 2009) to monkeys (Hayden et al., 2011; Procyk & Joseph, 2001; Shidara, Mizuhiki, & Richmond, 2005; Shidara & Richmond, 2002) and humans (Badre & Wagner, 2004; Duncan, 2010; Woodward et al., 2006) provides converging evidence across species that the region is responsible for managing transitions between successive task stages, consonant with application of top-down control by the ACC over each action within an option sequence, or recurrent information fed back to the ACC upon each action's completion (Holroyd & Yeung, 2012). These brain areas also produce electrophysiological signals of error and reward processing observed in both rats (Narayanan, Cavanagh, Frank, & Laubach, 2013; Warren, Hyman, Seamans, & Holroyd, 2014) and humans (M. X. Cohen, 2011; Walsh & Anderson, 2012), supporting the functional convergence between the rat ACC and human caudal ACC proposed here.

The human rostral ACC also appears to be functionally homologous with the rat prelimbic cortex. As described, direct recordings from neurons in the prelimbic cortex (Rich & Shapiro, 2009) and disruption of prelimbic cortex function (Birrell & Brown, 2000; Floresco, Block, & Tse, 2008; Ragozzino et al., 1999; Stefani & Moghaddam, 2005) strongly implicate this region in facilitating shifts between tasks. Further, direct connections from the prelimbic cortex to the dorsal ACC (Jones et al., 2005) facilitate neuronal synchronization between these regions (Totah, Jackson, & Moghaddam, 2013). Consistent with these observations and

with the hierarchical assumption of the present simulations, functional neuroimaging evidence in humans indicates that rostral ACC activity biases task processing in the caudal ACC (Nakao et al., 2010), such that the rostral ACC appears to regulate task selection and the caudal ACC implements task execution (Dixon & Christoff, 2012; Haynes et al., 2007; Venkatraman & Huettel, 2012; Venkatraman, Rosati, Taren, & Huettel, 2009; see also Haruno & Kawato, 2006). Although the foci of these rostral activations span a range of ACC areas from the pregenual ACC ($y = 34$, $z = 35$ in Talairach coordinates; Venkatraman & Huettel, 2012) to rostral portions of the anterior midcingulate cortex ($y = 56$, $z = 16$; Nakao et al., 2010), in all cases the activations occur anterior to a caudal focus of ACC activity concerned with relatively subordinate aspects of task execution. Rostral ACC activation is also negatively correlated with shifting performance in task switching (Wager, Jonides, Smith, & Nichols, 2005; see also Pollmann, Weidner, Müller, & von Cramon, 2000), and consistent with the assumed function of the prefrontal cortex, lesions of the rostral ACC near the genu of the corpus callosum impair task switching in humans, as indicated by lesion-symptom mapping of a large neuropsychological test battery data set obtained from 344 individuals with focal brain lesions (Gläscher et al., 2012). By contrast, as would be predicted by the hypothesis that the caudal ACC is responsible for energizing behavior, lesions of the caudal ACC in this same study were associated with impaired performance on a measure of word generation.

In contrast to our proposal that the rodent ACC controls the striatum directly, we have previously suggested that the primate ACC controls striatal action selection indirectly via its influence over DLPFC activity (Holroyd & Yeung, 2012; Umemoto & Holroyd, 2014; cf. Shenhav et al., 2013). Consistent with this idea, connections from Area 32 in the rostral ACC in primates target neurons in the DLPFC that are responsible for top-down control over behavior (Medalla & Barbas, 2009, 2010), and damage to the DLPFC in humans induces strategic impairments in task switching comparable with those observed following damage to the rat prefrontal cortex (Munakata, Morton, & Stedron, 2003; O'Reilly, Noelle, Braver, & Cohen, 2002). We propose that this apparent separation of function in primates between brain areas responsible for task selection (ACC) and task implementation (DLPFC) may afford computational advantages comparable with that achieved by dissociating the action and evaluative components of actor-critic models of behavior (Sutton & Barto, 1998). This view suggests that previous observations of greater ACC-DLPFC coactivation with increasing task difficulty, which occurs independently of the details of motor execution (Koski & Paus, 2000), may reflect transmission of a bias signal from the ACC to the DLPFC for task selection and motivation (Holroyd & Yeung, 2012; Umemoto & Holroyd, 2014). The HRL theory also suggests that the withdrawal of support by the ACC from a control-demanding task will result in a gradual decay in performance as memory for the task declines over time (Altmann, 2002), leading to goal neglect and perseverative activity (Duncan, Emslie, Williams, Johnson, & Freer, 1996; Morton & Munakata, 2002; Munakata et al., 2003).

The Role of Dopamine

The model also provides insight into a wealth of evidence that has associated tonic dopamine levels with the working memory

functions of the medial frontal cortex (Floresco & Magyar, 2006; Seamans & Yang, 2004). In particular, at a network-wide level, tonic dopamine appears to regulate the gain of cortical neurons in order to stabilize information held in working memory against distracters or neural noise (Durstewitz & Seamans, 2008; Servan-Schreiber, Printz, & Cohen, 1990; Thurley, Senn, & Lüscher, 2008). Stimulation of the midbrain dopamine system also elicits sustained neural firing of bistable neurons in the ACC (Onn & Wang, 2005; Peters, Barnhardt, & O'Donnell, 2004), which evidently functions to update the contents of working memory (Baeg et al., 2003; D'Ardenne et al., 2012; Petit et al., 1998). Reward-related changes of cortical dopamine levels also partly depend on the relative amount of work required to achieve the reward (Richardson & Gratton, 1998).

We found that disruption of tonic dopamine levels in the medial frontal cortex, which we simulated by setting the average reward values of the selected metaoption (for the prefrontal cortex) and option (for the ACC) to zero, adversely affected task performance in a manner comparable with direct lesions of dopaminergic nuclei. In both cases, this manipulation ensured that the rewards received were evaluated by the affected systems as being better than expected, resulting in reduced control over the lower levels even as performance deteriorated. As well, according to HRL theory, a selected option must be actively sustained until its completion (Botvinick, Niv, et al., 2009). Given the fundamental concern of the frontal cortex with maintaining task goals in working memory (E. K. Miller & Cohen, 2001), we have assumed that the rat medial frontal cortex maintains the selected options and metaoptions in working memory. These considerations suggest that tonic dopamine levels code for a threshold that determines the degree to which an option should be maintained in working memory. Outcomes that miss this target result in increased activation of the option in working memory, providing additional top-down support over task performance in order to boost the reward rate. Conversely, outcomes that surpass the target allow for the option to decay from working memory, because, in this case, the target reward rate can be maintained in the absence of control. Disruption to tonic dopamine levels thus sets this threshold to an artificially low level, which allows the contents of working memory to be emptied and control to dissipate freely (see also Floresco, Magyar, Ghods-Sharifi, Vexelman, & Tse, 2006; Nagano-Saito et al., 2008; Tunbridge, Bannerman, Sharp, & Harrison, 2004).

Model Fits and Future Directions

The model simulations are not perfect. For example, Figure 6 indicates that, following changes in task contingencies, the model immediately adopts a stable rate of performance, whereas the empirical data suggest a slower rate of adaptation. Although it would be possible to account for such differences by equipping the model with additional degrees of freedom (e.g., by introducing a perseverative bias that slows between-arm switches), we elected not to make any such changes in order to focus the discussion on the core principles of the theory: hierarchy, control, and reinforcement. Likewise, the model performs worse than actual rats do on the reversal learning task (Figure 3, bottom panels). This discrepancy likely arises from the fact that the model does not include a module for OFC function, which is widely believed to facilitate reversal learning (Boulougouris, Dalley, & Robbins, 2007; Burke,

Takahashi, Correll, Brown, & Schoenbaum, 2009; Hornak et al., 2004; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009; see also Wheeler & Fellows, 2008). We decided not to include such a module in order to highlight the functions of the ACC and pre-limbic cortex, but doing so would be an important step for incorporating ACC function within a complete neural model of reward processing and action selection.

Although it has previously been suggested that ACC damage shifts the relative weight that the system accords to rewards versus effort-related costs in a one-level cost–benefit analysis (e.g., Walton et al., 2006), the HRL theory provides a mechanism whereby such control can be targeted to specific tasks. Consistent with this assertion, optogenetic stimulation of medial frontal cortex neurons that selectively project to the dorsal raphe nucleus (DRN), a serotonergic nucleus, induces task-dependent behavioral activation: Whereas direct stimulation of the DRN increases mobility overall, indirect stimulation of the DRN, by way of the medial frontal cortex, increases mobility only in a challenging task (Warden et al., 2012; see also Welberg, 2013). By contrast, the anterior portion of the dorsal medial striatum appears to motivate or energize behavior overall rather than in a task-specific way (A. Y. Wang, Miura, & Uchida, 2013; see also Levy & Dubois, 2006; Schmidt et al., 2008). This aspect of the theory could be further tested in an experiment that exercised all three levels of the hierarchy simultaneously, for example, with a task-switching design in which multiple tasks demanded varying degrees of effortful control. A straightforward prediction of the model would be that rats allocate effort specific to the demands of each task, rather than according to a flat cost–benefit analysis, and that this allocation depends on the integrity of the ACC.

How options are established in the first place is an open question for future research. We have speculated elsewhere that a repertoire of options may be learned during the course of early development (Lukie, Montazer-Hojat, & Holroyd, 2014). In this view, new options would be generated in a trial-and-error learning process sustained by an intrinsic motivation to explore (Singh, Barto, & Chentanez, 2005), which could later be recombined in combinatorial fashion to address more complex problems (Elman, Bates, Johnson, & Karmiloff-Smith, 1996). Options could also be formed around the learning of latent high-level structure in the environment (Gershman & Niv, 2010, 2012), and/or by identifying sub-goals through a structural analysis of the task domain (Diuk, Schapiro, et al., 2013). But it is also the case that many options must be fashioned on the fly, for example, according to the task instructions given to subjects in a psychology experiment. In general, this is an outstanding issue in both psychology and artificial intelligence, and, as such, exceeds the scope of the present investigation.

Our account was motivated by the observation that previous research on ACC function has focused on detailed examinations of moment-to-moment changes in control, thereby overlooking the higher level structure associated with more extended, everyday activities (Holroyd & Yeung, 2012). Foraging provides an instructive example of such a naturalistic behavior. This activity is shared between birds in the wild and humans at the mall (Stephens, Brown, & Ydenberg, 2007), and, like so many other behaviors, is associated with ACC activity in rats (Hillman & Bilkey, 2012), monkeys (Blanchard & Hayden, 2014; Hayden et al., 2011), and

humans (Kolling, Behrens, Mars, & Rushworth, 2012). However, unlike most other behaviors, foraging is also specifically impaired by ACC damage in rats (Li et al., 2012; Seamans, Floresco, & Phillips, 1995). Current explanations of the role of the ACC in foraging hold that the ACC codes the average value of the foraging environment and the cost of foraging (Kolling et al., 2012), or is responsible for overriding the pursuit of immediate rewards associated with the fruit at hand in favor of more delayed rewards to be found at a distal source (Shenhav et al., 2013). ACC activity is also observed to motivate changes in behavioral strategies rather than to affect current decision making (Blanchard & Hayden, 2014). Our model can be thought of as a computational approach that formalizes these presumed functions. It does so by couching the contribution of the ACC to foraging within a hierarchical motor control system: Foraging is characterized by collections of relatively primitive actions during resource exploitation (such as fishing from a pond), punctuated by high-level strategic shifts during resource exploration (such as deciding to leave a pond in search of a new pond). If the principles of the model indeed generalize to humans, then the model predicts that the caudal ACC implements individual tasks (such as fishing vs. walking) and the rostral ACC (which is sensitive to the overall value of the environment as opposed to the specific values of either fishing or walking) supports the shift between tasks. This prediction can be tested by recording the fMRI BOLD signal from humans immersed in a virtual foraging task.

Conclusion

Our critique of current theories of ACC function was predicated on the observation that ACC lesions tend to spare performance on experimental tasks that depend on those proposed functions (Holroyd & Yeung, 2012). Yet apart from rare accounts of akinetic mutism, the effect of ACC damage on humans is often subtle and nonapparent, even in everyday life. The earliest clinical observations noted that a person with ACC damage “shows nothing obviously unusual in his behavior when one sees him, meets him or talks to him: to all intents and purposes he is normal” (Tow & Whitty, 1953, p. 192). Closer examination revealed that such patients exhibit a subtle array of deficits, including reduced drive and less interest in engaging in creative activities (Tow & Whitty, 1953). Conversely, stimulation of the ACC (as discussed in the introduction) motivates effortful behaviors in humans much as it motivates rats to climb large barriers, especially to execute extended sequences of actions directed toward larger task goals like traveling miles to pass through a storm. The HRL theory accounts for such observations by proposing that the ACC applies the motivational control necessary to engage in specific activities as a task-related whole (such as learning to play the piano) rather than the effort needed to initiate any actions in particular (such as pressing the individual keys). This discrepancy between the ACC’s apparently paramount role in cognitive control and decision making, on the one hand, and the seemingly modest consequences of ACC damage on these processes, on the other, underscores the fact that the current tests of ACC function are relatively insensitive to the longer term aspects of control regulation. On the contrary, the HRL theory suggests that the effects of ACC damage will manifest most strongly on extended, naturalistic tasks rather than on the

tests of short-term control that characterize conventional laboratory investigations of ACC function.

References

- Aarts, E., Roelofs, A., & van Turenout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *The Journal of Neuroscience*, 28, 4671–4678. <http://dx.doi.org/10.1523/JNEUROSCI.4400-07.2008>
- Ackerman, P. L. (2011). 100 years without resting. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 11–43). Washington, DC: American Psychological Association.
- Alexander, M. P., Stuss, D. T., Shallice, T., Picton, T. W., & Gillingham, S. (2005). Impaired concentration due to frontal lobe damage from two distinct lesion sites. *Neurology*, 65, 572–579. <http://dx.doi.org/10.1212/01.wnl.0000172912.07640.92>
- Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, 14, 1338–1344. <http://dx.doi.org/10.1038/nn.2921>
- Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention & performance XV* (pp. 421–452). Cambridge, MA: MIT Press.
- Altmann, E. M. (2002). Functional decay of memory for tasks. *Psychological Research*, 66, 287–297. <http://dx.doi.org/10.1007/s00426-002-0102-9>
- Amiez, C., Joseph, J.-P., & Procyk, E. (2005). Anterior cingulate error-related activity is modulated by predicted reward. *European Journal of Neuroscience*, 21, 3447–3452. <http://dx.doi.org/10.1111/j.1460-9568.2005.04170.x>
- Amiez, C., Joseph, J.-P., & Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex*, 16, 1040–1055. <http://dx.doi.org/10.1093/cercor/bhj046>
- Arntsen, A. F. T., Paspalas, C. D., Gamo, N. J., Yang, Y., & Wang, M. (2010). Dynamic network connectivity: A new form of neuroplasticity. *Trends in Cognitive Sciences*, 14, 365–375. <http://dx.doi.org/10.1016/j.tics.2010.05.003>
- Ashlock, D. (2010). *Evolutionary computation for modeling and optimization*. New York, NY: Springer.
- Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 19, 2082–2099. <http://dx.doi.org/10.1162/jocn.2007.19.12.2082>
- Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, 10, 659–669. <http://dx.doi.org/10.1038/nrn2667>
- Badre, D., & Frank, M. J. (2012). Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: Evidence from fMRI. *Cerebral Cortex*, 22, 527–536. <http://dx.doi.org/10.1093/cercor/bhr117>
- Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, 41, 473–487. [http://dx.doi.org/10.1016/S0896-6273\(03\)00851-1](http://dx.doi.org/10.1016/S0896-6273(03)00851-1)
- Baeg, E. H., Kim, Y. B., Huh, K., Mook-Jung, I., Kim, H. T., & Jung, M. W. (2003). Dynamics of population code for working memory in the prefrontal cortex. *Neuron*, 40, 177–188. [http://dx.doi.org/10.1016/S0896-6273\(03\)00597-X](http://dx.doi.org/10.1016/S0896-6273(03)00597-X)
- Balaguer-Ballester, E., Lapish, C. C., Seamans, J. K., & Durstewitz, D. (2011). Attracting dynamics of frontal cortex ensembles during memory-guided decision-making. *PLoS Computational Biology*, 7, e1002057. <http://dx.doi.org/10.1371/journal.pcbi.1002057>
- Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35, 48–69. <http://dx.doi.org/10.1038/npp.2009.131>
- Barto, A. G. (1995). Adaptive critics and the basal ganglia. In J. Houk, J. Davis, & D. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 215–232). Cambridge, MA: MIT Press.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265. <http://dx.doi.org/10.1037/0022-3514.74.5.1252>
- Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, 7, 1–15. http://dx.doi.org/10.1207/s15327965pli0701_1
- Beierholm, U., Guitart-Masip, M., Economides, M., Chowdhury, R., Düzel, E., Dolan, R., & Dayan, P. (2013). Dopamine modulates reward-related vigor. *Neuropsychopharmacology*, 38, 1495–1503. <http://dx.doi.org/10.1038/npp.2013.48>
- Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *The Journal of Neuroscience*, 20, 4320–4324.
- Blanchard, T. C., & Hayden, B. Y. (2014). Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *The Journal of Neuroscience*, 34, 646–655. <http://dx.doi.org/10.1523/JNEUROSCI.3151-13.2014>
- Blond, O., Crépel, F., & Otani, S. (2002). Long-term potentiation in rat prefrontal slices facilitated by phased application of dopamine. *European Journal of Pharmacology*, 438, 115–116. [http://dx.doi.org/10.1016/S0014-2999\(02\)01291-8](http://dx.doi.org/10.1016/S0014-2999(02)01291-8)
- Boksem, M. A. S., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, 59, 125–139. <http://dx.doi.org/10.1016/j.brainresrev.2008.07.001>
- Boorman, E. D., Rushworth, M. F., & Behrens, T. E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *The Journal of Neuroscience*, 33, 2242–2253. <http://dx.doi.org/10.1523/JNEUROSCI.3022-12.2013>
- Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, 7, 356–366. <http://dx.doi.org/10.3758/CABN.7.4.356>
- Botvinick, M. M. (2012). Hierarchical reinforcement learning and decision making. *Current Opinion in Neurobiology*, 22, 956–962. <http://dx.doi.org/10.1016/j.conb.2012.05.008>
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624–652. <http://dx.doi.org/10.1037/0033-295X.108.3.624>
- Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective & Behavioral Neuroscience*, 9, 16–27. <http://dx.doi.org/10.3758/CABN.9.1.16>
- Botvinick, M. M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113, 262–280. <http://dx.doi.org/10.1016/j.cognition.2008.08.011>
- Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research*, 179, 219–228. <http://dx.doi.org/10.1016/j.bbr.2007.02.005>
- Brockner, J., Shaw, M. C., & Rubin, J. Z. (1979). Factors affecting withdrawal from an escalating conflict: Quitting before it's too late. *Journal of Experimental Social Psychology*, 15, 492–503. [http://dx.doi.org/10.1016/0022-1031\(79\)90011-8](http://dx.doi.org/10.1016/0022-1031(79)90011-8)
- Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, 307, 1118–1121. <http://dx.doi.org/10.1126/science.1105783>

- Brown, V. J., & Bowman, E. M. (2002). Rodent models of prefrontal cortical function. *Trends in Neurosciences*, 25, 340–343. [http://dx.doi.org/10.1016/S0166-2236\(02\)02164-1](http://dx.doi.org/10.1016/S0166-2236(02)02164-1)
- Burke, K. A., Takahashi, Y. K., Correll, J., Brown, P. L., & Schoenbaum, G. (2009). Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with “disinhibition” of responding for previously unrewarded cues. *European Journal of Neuroscience*, 30, 1941–1946. <http://dx.doi.org/10.1111/j.1460-9568.2009.06992.x>
- Bush, G. (2009). Dorsal anterior midcingulate cortex: Roles in normal cognition and disruption in attention-deficit/hyperactivity disorder. In B. A. Vogt (Ed.), *Cingulate neurobiology and disease* (pp. 246–274). Oxford, UK: Oxford University Press.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4, 215–222. [http://dx.doi.org/10.1016/S1364-6613\(00\)01483-2](http://dx.doi.org/10.1016/S1364-6613(00)01483-2)
- Cai, X., & Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *The Journal of Neuroscience*, 32, 3791–3808. <http://dx.doi.org/10.1523/JNEUROSCI.3864-11.2012>
- Caluwaerts, K., Staffa, M., N’Guyen, S., Grand, C., Dollé, L., Favre-Félix, A., . . . Khamassi, M. (2012). A biologically inspired meta-control navigation system for the Psikharpx rat robot. *Bioinspiration & Biomimetics*, 7, 025009. <http://dx.doi.org/10.1088/1748-3182/7/2/025009>
- Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *The Journal of Neuroscience*, 31, 15048–15052. <http://dx.doi.org/10.1523/JNEUROSCI.3164-11.2011>
- Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *The Journal of Neuroscience*, 33, 8541–8548. <http://dx.doi.org/10.1523/JNEUROSCI.5754-12.2013>
- Chudasama, Y., Passetti, F., Rhodes, S. E. V., Lopian, D., Desai, A., & Robbins, T. W. (2003). Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: Differential effects on selectivity, impulsivity and compulsivity. *Behavioural Brain Research*, 146, 105–119. <http://dx.doi.org/10.1016/j.bbr.2003.09.020>
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, 332–361. <http://dx.doi.org/10.1037/0033-295X.97.3.332>
- Cohen, M. X. (2011). Error-related medial frontal theta activity predicts cingulate-related structural connectivity. *NeuroImage*, 55, 1373–1383. <http://dx.doi.org/10.1016/j.neuroimage.2010.12.072>
- Cohen, M. X., & Frank, M. J. (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behavioural Brain Research*, 199, 141–156. <http://dx.doi.org/10.1016/j.bbr.2008.09.029>
- Collins, A. G., Cavanagh, J. F., & Frank, M. J. (2014). Human EEG uncovers latent generalizable rule structure during learning. *The Journal of Neuroscience*, 34, 4677–4685. <http://dx.doi.org/10.1523/JNEUROSCI.3900-13.2014>
- Cousins, M. S., Atherton, A., Turner, L., & Salamone, J. D. (1996). Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. *Behavioural Brain Research*, 74, 189–197. [http://dx.doi.org/10.1016/0166-4328\(95\)00151-4](http://dx.doi.org/10.1016/0166-4328(95)00151-4)
- Cousins, M. S., & Salamone, J. D. (1994). Nucleus accumbens dopamine depletions in rats affect relative response allocation in a novel cost/benefit procedure. *Pharmacology, Biochemistry and Behavior*, 49, 85–91. [http://dx.doi.org/10.1016/0091-3057\(94\)90460-X](http://dx.doi.org/10.1016/0091-3057(94)90460-X)
- Cowen, S. L., Davis, G. A., & Nitz, D. A. (2012). Anterior cingulate neurons in the rat map anticipated effort and reward to their associated action sequences. *Journal of Neurophysiology*, 107, 2393–2407. <http://dx.doi.org/10.1152/jn.01012.2011>
- Cowen, S. L., & McNaughton, B. L. (2007). Selective delay activity in the medial prefrontal cortex of the rat: Contribution of sensorimotor information and contingency. *Journal of Neurophysiology*, 98, 303–316. <http://dx.doi.org/10.1152/jn.00150.2007>
- Crittenden, B. M., & Duncan, J. (2014). Task difficulty manipulation reveals multiple demand activity but no frontal lobe hierarchy. *Cerebral Cortex*, 24, 532–540. <http://dx.doi.org/10.1093/cercor/bhs333>
- Croxxon, P. L., Walton, M. E., O’Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-based cost-benefit valuation and the human brain. *The Journal of Neuroscience*, 29, 4531–4541. <http://dx.doi.org/10.1523/JNEUROSCI.4515-08.2009>
- Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: Neural and neurochemical substrates. *Neuroscience and Biobehavioral Reviews*, 28, 771–784. <http://dx.doi.org/10.1016/j.neubiorev.2004.09.006>
- D’Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L. E., & Cohen, J. D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 19900–19909. <http://dx.doi.org/10.1073/pnas.1116727109>
- Davis, K. D., Hutchison, W. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (2000). Human anterior cingulate cortex neurons modulated by attention-demanding tasks. *Journal of Neurophysiology*, 83, 3575–3577. <http://dx.doi.org/10.1523/JNEUROSCI.2315-05.2005>
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8, 1704–1711. <http://dx.doi.org/10.1038/nn1560>
- Day, J. J., Jones, J. L., & Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *European Journal of Neuroscience*, 33, 308–321. <http://dx.doi.org/10.1111/j.1460-9568.2010.07531.x>
- Day, J. J., Jones, J. L., Wightman, R. M., & Carelli, R. M. (2010). Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. *Biological Psychiatry*, 68, 306–309. <http://dx.doi.org/10.1016/j.biopsych.2010.03.026>
- Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience*, 35, 1036–1051. <http://dx.doi.org/10.1111/j.1460-9568.2012.08050.x>
- Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Computational Biology*, 9, e1003364. <http://dx.doi.org/10.1371/journal.pcbi.1003364>
- Dias, R., & Aggleton, J. P. (2000). Effects of selective excitotoxic prefrontal lesions on acquisition of nonmatching- and matching-to-place in the T-maze in the rat: Differential involvement of the prelimbic-infralimbic and anterior cingulate cortices in providing behavioural flexibility. *European Journal of Neuroscience*, 12, 4457–4466. <http://dx.doi.org/10.1046/j.0953-816X.2000.01323.x>
- Diuk, C., Schapiro, A., Cordova, N., Ribas-Fernandes, J., Niv, Y., & Botvinick, M. (2013). Divide and conquer: Hierarchical reinforcement learning and task decomposition in humans. In G. Baldassarre & M. Mirolli (Eds.), *Computational and robotic models of the hierarchical organization of behavior* (pp. 271–291). New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-39875-9_12
- Diuk, C., Tsai, K., Wallis, J., Botvinick, M., & Niv, Y. (2013). Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *The Journal of Neuroscience*, 33, 5797–5805. <http://dx.doi.org/10.1523/JNEUROSCI.5445-12.2013>
- Dixon, M. L., & Christoff, K. (2012). The decision to engage cognitive control is driven by expected reward-value: Neural and behavioral evidence. *PLoS ONE*, 7, e51637. <http://dx.doi.org/10.1371/journal.pone.0051637>

- Dixon, M. L., Fox, K. C. R., & Christoff, K. (2014). Evidence for rostral-caudal functional organization in multiple brain areas related to goal-directed behavior. *Brain Research*, 1572, 26–39. <http://dx.doi.org/10.1016/j.brainres.2014.05.012>
- Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12, 99–105. <http://dx.doi.org/10.1016/j.tics.2008.01.001>
- Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., . . . Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 11073–11078. <http://dx.doi.org/10.1073/pnas.0704320104>
- Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., . . . Petersen, S. E. (2006). A core system for the implementation of task sets. *Neuron*, 50, 799–812. <http://dx.doi.org/10.1016/j.neuron.2006.04.031>
- Dreisbach, G., & Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, 78, 94–98. <http://dx.doi.org/10.1016/j.bandc.2011.12.003>
- Dum, R. P., Levinthal, D. J., & Strick, P. L. (2009). The spinothalamic system targets motor and sensory areas in the cerebral cortex of monkeys. *The Journal of Neuroscience*, 29, 14223–14235. <http://dx.doi.org/10.1523/JNEUROSCI.3398-09.2009>
- Dum, R. P., & Strick, P. L. (1996). Spinal cord terminations of the medial wall motor areas in macaque monkeys. *The Journal of Neuroscience*, 16, 6513–6525.
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14, 172–179. <http://dx.doi.org/10.1016/j.tics.2010.01.004>
- Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, 30, 257–303. <http://dx.doi.org/10.1006/cogp.1996.0008>
- Durstewitz, D., & Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biological Psychiatry*, 64, 739–749. <http://dx.doi.org/10.1016/j.biopsych.2008.05.015>
- Durstewitz, D., Vitoz, N. M., Floresco, S. B., & Seamans, J. K. (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron*, 66, 438–448. <http://dx.doi.org/10.1016/j.neuron.2010.03.029>
- Elman, J. L., Bates, E. A., Johnson, M. H., & Karmiloff-Smith, A. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, 49, 709–724. <http://dx.doi.org/10.1037/0003-066X.49.8.709>
- Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, 76, 1057–1070. <http://dx.doi.org/10.1016/j.neuron.2012.12.002>
- Farooqui, A. A., Mitchell, D., Thompson, R., & Duncan, J. (2012). Hierarchical organization of cognition reflected in distributed frontoparietal activity. *The Journal of Neuroscience*, 32, 17373–17381. <http://dx.doi.org/10.1523/JNEUROSCI.0598-12.2012>
- Floresco, S. B. (2013). Prefrontal dopamine and behavioral flexibility: Shifting from an “inverted-U” toward a family of functions. *Frontiers in Neuroscience*, 7, 62. <http://dx.doi.org/10.3389/fnins.2013.00062>
- Floresco, S. B., Block, A. E., & Tse, M. T. (2008). Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behavioural Brain Research*, 190, 85–96. <http://dx.doi.org/10.1016/j.bbr.2008.02.008>
- Floresco, S. B., & Magyar, O. (2006). Mesocortical dopamine modulation of executive functions: Beyond working memory. *Psychopharmacology*, 188, 567–585. <http://dx.doi.org/10.1007/s00213-006-0404-5>
- Floresco, S. B., Magyar, O., Ghods-Sharifi, S., Vexelman, C., & Tse, M. T. L. (2006). Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology*, 31, 297–309. <http://dx.doi.org/10.1038/sj.npp.1300825>
- Forstmann, B. U., Brass, M., Koch, I., & von Cramon, D. Y. (2006). Voluntary selection of task sets revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 18, 388–398. <http://dx.doi.org/10.1162/jocn.2006.18.3.388>
- Frank, M. J., & Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Computational analysis. *Cerebral Cortex*, 22, 509–526. <http://dx.doi.org/10.1093/cercor/bhr114>
- Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective & Behavioral Neuroscience*, 1, 137–160. <http://dx.doi.org/10.3758/CABN.1.2.137>
- Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 16311–16316. <http://dx.doi.org/10.1073/pnas.0706111104>
- Fudenberg, D., & Levine, D. K. (2006). A dual-self model of impulse control. *The American Economic Review*, 96, 1449–1476. <http://dx.doi.org/10.1257/aer.96.5.1449>
- Fujisawa, S., Amarasingham, A., Harrison, M. T., & Buzsáki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nature Neuroscience*, 11, 823–833. <http://dx.doi.org/10.1038/nn.2134>
- Gabbott, P. L. A., Warner, T. A., Jays, P. R. I., Salway, P., & Busby, S. J. (2005). Prefrontal cortex in the rat: Projections to subcortical autonomic, motor, and limbic centers. *The Journal of Comparative Neurology*, 492, 145–177. <http://dx.doi.org/10.1002/cne.20738>
- Gan, J. O., Walton, M. E., & Phillips, P. E. M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature Neuroscience*, 13, 25–27. <http://dx.doi.org/10.1038/nn.2460>
- Gershman, S. J., & Niv, Y. (2010). Learning latent structure: Carving nature at its joints. *Current Opinion in Neurobiology*, 20, 251–256. <http://dx.doi.org/10.1016/j.conb.2010.02.008>
- Gershman, S. J., & Niv, Y. (2012). Exploring a latent cause theory of classical conditioning. *Learning & Behavior*, 40, 255–268. <http://dx.doi.org/10.3758/s13420-012-0080-8>
- Gevens, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, 7, 374–385. <http://dx.doi.org/10.1093/cercor/7.4.374>
- Gläscher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., . . . Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 14681–14686. <http://dx.doi.org/10.1073/pnas.1206608109>
- Goto, Y., & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, 8, 805–812. <http://dx.doi.org/10.1038/nn1471>
- Haber, S. N., & Knutson, B. (2010). The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35, 4–26. <http://dx.doi.org/10.1038/npp.2009.129>
- Haddon, J. E., & Killcross, S. (2006). Prefrontal cortex lesions disrupt the contextual control of response conflict. *The Journal of Neuroscience*, 26, 2933–2940. <http://dx.doi.org/10.1523/JNEUROSCI.3243-05.2006>
- Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, 136, 495–525. <http://dx.doi.org/10.1037/a0019486>

- Haruno, M., & Kawato, M. (2006). Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: FMRI examination in stimulus-action-reward association learning. *Neural Networks*, 19, 1242–1254. <http://dx.doi.org/10.1016/j.neunet.2006.06.007>
- Hauber, W., & Sommer, S. (2009). Prefrontostriatal circuitry regulates effort-related decision making. *Cerebral Cortex*, 19, 2240–2247. <http://dx.doi.org/10.1093/cercor/bhn241>
- Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, 14, 933–939. <http://dx.doi.org/10.1038/nn.2856>
- Hayden, B. Y., & Platt, M. L. (2010). Neurons in anterior cingulate cortex multiplex information about reward and action. *The Journal of Neuroscience*, 30, 3339–3346. <http://dx.doi.org/10.1523/JNEUROSCI.4874-09.2010>
- Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Current Biology*, 17, 323–328. <http://dx.doi.org/10.1016/j.cub.2006.11.072>
- Heatherington, T. F., & Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in Cognitive Sciences*, 15, 132–139. <http://dx.doi.org/10.1016/j.tics.2010.12.005>
- Hengst, B. (2012). Hierarchical approaches. In M. Wiering & M. Otterlo (Eds.), *Adapting, learning and optimization, Vol. 12: Reinforcement learning* (pp. 293–323). Berlin, Germany: Springer-Verlag.
- Hikosaka, O., & Isoda, M. (2010). Switching from automatic to controlled behavior: Cortico-basal ganglia mechanisms. *Trends in Cognitive Sciences*, 14, 154–161. <http://dx.doi.org/10.1016/j.tics.2010.01.006>
- Hillman, K. L., & Bilkey, D. K. (2010). Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. *The Journal of Neuroscience*, 30, 7705–7713. <http://dx.doi.org/10.1523/JNEUROSCI.1273-10.2010>
- Hillman, K. L., & Bilkey, D. K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. *Nature Neuroscience*, 15, 1290–1297. <http://dx.doi.org/10.1038/nn.3187>
- Hillman, K. L., & Bilkey, D. K. (2013). Persisting through subjective effort: A key role for the anterior cingulate cortex? *Behavioral and Brain Sciences*, 36, 691–692. <http://dx.doi.org/10.1017/S0140525X13001039>
- Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, 4, 162–176. <http://dx.doi.org/10.1111/j.1745-6924.2009.01116.x>
- Holroyd, C. B. (2013). Theories of anterior cingulate cortex function: Opportunity cost. *Behavioral and Brain Sciences*, 36, 693–694. <http://dx.doi.org/10.1017/S0140525X13001052>
- Holroyd, C. B. (2014). *The waste disposal problem of effortful control*. Manuscript submitted for publication.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–709. <http://dx.doi.org/10.1037/0033-295X.109.4.679>
- Holroyd, C. B., & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex*, 44, 548–559. <http://dx.doi.org/10.1016/j.cortex.2007.08.013>
- Holroyd, C. B., & Yeung, N. (2011). An integrative theory of anterior cingulate cortex function: Option selection in hierarchical reinforcement learning. In R. B. Mars, J. Sallet, M. F. S. Rushworth, & N. Yeung (Eds.), *Neural basis of motivational and cognitive control* (pp. 332–349). Cambridge, MA: MIT Press. <http://dx.doi.org/10.7551/mitpress/9780262016438.003.0018>
- Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, 16, 122–128. <http://dx.doi.org/10.1016/j.tics.2011.12.008>
- Holroyd, C. B., Yeung, N., Coles, M. G. H., & Cohen, J. D. (2005). A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology: General*, 134, 163–191. <http://dx.doi.org/10.1037/0096-3445.134.2.163>
- Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., & Polkey, C. E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, 16, 463–478. <http://dx.doi.org/10.1162/089892904322926791>
- Houk, J. C., Adams, J. L., & Barto, A. G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. Houk, J. Davis, & D. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 249–270). Cambridge, MA: MIT Press.
- Hyafil, A., Summerfield, C., & Koehlin, E. (2009). Two mechanisms for task switching in the prefrontal cortex. *The Journal of Neuroscience*, 29, 5135–5142. <http://dx.doi.org/10.1523/JNEUROSCI.2828-08.2009>
- Hyman, J. M., Ma, L., Balaguer-Ballester, E., Durstewitz, D., & Seamans, J. K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 5086–5091. <http://dx.doi.org/10.1073/pnas.1114415109>
- Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, 7, 450–463. <http://dx.doi.org/10.1177/1745691612454134>
- Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, 18, 127–133. <http://dx.doi.org/10.1016/j.tics.2013.12.009>
- Jeon, H. A., & Friederici, A. D. (2013). Two principles of organization in the prefrontal cortex are cognitive hierarchy and degree of automaticity. *Nature Communications*, 4, 2041. <http://dx.doi.org/10.1038/ncomms3041>
- Johansen, J. P., & Fields, H. L. (2004). Glutamatergic activation of anterior cingulate cortex produces an aversive teaching signal. *Nature Neuroscience*, 7, 398–403. <http://dx.doi.org/10.1038/nn1207>
- Johnston, K., Levin, H. M., Koval, M. J., & Everling, S. (2007). Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron*, 53, 453–462. <http://dx.doi.org/10.1016/j.neuron.2006.12.023>
- Jones, B. F., Groenewegen, H. J., & Witter, M. P. (2005). Intrinsic connections of the cingulate cortex in the rat suggest the existence of multiple functionally segregated networks. *Neuroscience*, 133, 193–207. <http://dx.doi.org/10.1016/j.neuroscience.2005.01.063>
- Jung, M. W., Baeg, E. H., Kim, M. J., Kim, Y. B., & Kim, J. J. (2008). Plasticity and memory in the prefrontal cortex. *Reviews in the Neurosciences*, 19, 29–46. <http://dx.doi.org/10.1515/REVNEURO.2008.19.1.29>
- Kahneman, D. (2000). Evaluation by moments: Past and future. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 699–708). Cambridge, UK: Cambridge University Press.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720. <http://dx.doi.org/10.1037/0003-066X.58.9.697>
- Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, 14, 1581–1589. <http://dx.doi.org/10.1038/nn.2961>
- Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9, 940–947. <http://dx.doi.org/10.1038/nn1724>
- Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, 7, e1002055. <http://dx.doi.org/10.1371/journal.pcbi.1002055>
- Keramati, M., & Gutkin, B. (2013). Imbalanced decision hierarchy in addicts emerging from drug-hijacked dopamine spiraling circuit. *PLoS ONE*, 8, e61489. <http://dx.doi.org/10.1371/journal.pone.0061489>

- Khamassi, M., Lallée, S., Enel, P., Procyk, E., & Dominey, P. F. (2011). Robot cognitive control with a neurophysiologically inspired reinforcement learning model. *Frontiers in Neurorobotics*, 5, 1. <http://dx.doi.org/10.3389/fnbot.2011.00001>
- Kiesel, A., Steinhäuser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, 136, 849–874. <http://dx.doi.org/10.1037/a0019842>
- Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, 13, 400–408. <http://dx.doi.org/10.1093/cercor/13.4.400>
- Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General*, 122, 411–428. <http://dx.doi.org/10.1037/0096-3445.122.4.411>
- Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11, 229–235. <http://dx.doi.org/10.1016/j.tics.2007.04.005>
- Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science*, 336, 95–98. <http://dx.doi.org/10.1126/science.1216930>
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, 139, 665–682. <http://dx.doi.org/10.1037/a0020198>
- Kool, W., McGuire, J. T., Wang, G. J., & Botvinick, M. M. (2013). Neural and behavioral evidence for an intrinsic cost of self-control. *PLoS ONE*, 8, e72626. <http://dx.doi.org/10.1371/journal.pone.0072626>
- Koski, L., & Paus, T. (2000). Functional connectivity of the anterior cingulate cortex within the human frontal lobe: A brain-mapping meta-analysis. *Experimental Brain Research*, 133, 55–65. <http://dx.doi.org/10.1007/s002210000400>
- Kouneiher, F., Charron, S., & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, 12, 939–945. <http://dx.doi.org/10.1038/nn.2321>
- Kurniawan, I. T., Guitart-Masip, M., Dayan, P., & Dolan, R. J. (2013). Effort and valuation in the brain: The effects of anticipation and execution. *The Journal of Neuroscience*, 33, 6160–6169. <http://dx.doi.org/10.1523/JNEUROSCI.4777-12.2013>
- Kurniawan, I. T., Guitart-Masip, M., & Dolan, R. J. (2011). Dopamine and effort-based decision making. *Frontiers in Neuroscience*, 5, 81. <http://dx.doi.org/10.3389/fnins.2011.00081>
- Kurniawan, I. T., Seymour, B., Talmi, D., Yoshida, W., Chater, N., & Dolan, R. J. (2010). Choosing to make an effort: The role of striatum in signaling physical effort of a chosen action. *Journal of Neurophysiology*, 104, 313–321. <http://dx.doi.org/10.1152/jn.00027.2010>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36, 661–679. <http://dx.doi.org/10.1017/S0140525X12003196>
- Lapish, C. C., Durstewitz, D., Chandler, L. J., & Seamans, J. K. (2008). Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 11963–11968. <http://dx.doi.org/10.1073/pnas.0804045105>
- Laubach, M. (2011). A comparative perspective on executive and motivational control in medial prefrontal cortex. In R. B. Mars, J. Sallet, M. F. S. Rushworth, & N. Yeung (Eds.), *Neural basis of motivational and cognitive control* (pp. 95–109). Cambridge, MA: MIT Press. <http://dx.doi.org/10.7551/mitpress/9780262016438.003.0006>
- Levy, R., & Dubois, B. (2006). Apathy and the functional anatomy of the prefrontal cortex-basal ganglia circuits. *Cerebral Cortex*, 16, 916–928. <http://dx.doi.org/10.1093/cercor/bhj043>
- Li, F., Li, M., Cao, W., Xu, Y., Luo, Y., Zhong, X., . . . Li, C. (2012). Anterior cingulate cortical lesion attenuates food foraging in rats. *Brain Research Bulletin*, 88, 602–608. <http://dx.doi.org/10.1016/j.brainresbull.2012.05.015>
- Løvstad, M., Funderud, I., Meling, T., Krämer, U. M., Voytek, B., Due-Tønnessen, P., . . . Solbakk, A. K. (2012). Anterior cingulate cortex and cognitive control: Neuropsychological and electrophysiological findings in two patients with lesions to dorsomedial prefrontal cortex. *Brain and Cognition*, 80, 237–249. <http://dx.doi.org/10.1016/j.bandc.2012.07.008>
- Luk, C. H., & Wallis, J. D. (2009). Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *The Journal of Neuroscience*, 29, 7526–7539. <http://dx.doi.org/10.1523/JNEUROSCI.0386-09.2009>
- Lukie, C. N., Montazer-Hojat, S., & Holroyd, C. B. (2014). Developmental changes in the reward positivity: An electrophysiological trajectory of reward processing. *Developmental Cognitive Neuroscience*, 9, 191–199. <http://dx.doi.org/10.1016/j.dcn.2014.04.003>
- Ma, L., Hyman, J. M., Lindsay, A. J., Phillips, A. G., & Seamans, J. K. (2014). Differences in the emergent coding properties of cortical and striatal ensembles. *Nature Neuroscience*, 17, 1100–1106. <http://dx.doi.org/10.1038/nn.3753>
- Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, 10, 647–656. <http://dx.doi.org/10.1038/nn1890>
- McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review*, 120, 395–410. <http://dx.doi.org/10.1037/a0031910>
- Medalla, M., & Barbas, H. (2009). Synapses with inhibitory neurons differentiate anterior cingulate from dorsolateral prefrontal pathways associated with cognitive control. *Neuron*, 61, 609–620. <http://dx.doi.org/10.1016/j.neuron.2009.01.006>
- Medalla, M., & Barbas, H. (2010). Anterior cingulate synapses in prefrontal areas 10 and 46 suggest differential influence in cognitive control. *The Journal of Neuroscience*, 30, 16068–16081. <http://dx.doi.org/10.1523/JNEUROSCI.1773-10.2010>
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202. <http://dx.doi.org/10.1146/annurev.neuro.24.1.167>
- Miller, M. W. (1987). The origin of corticospinal projection neurons in rat. *Experimental Brain Research*, 67, 339–351. <http://dx.doi.org/10.1007/BF00248554>
- Millington, R. S., Poljac, E., & Yeung, N. (2013). Between-task competition for intentions and actions. *Quarterly Journal of Experimental Psychology* (2006), 66, 1504–1516. <http://dx.doi.org/10.1080/17470218.2012.746381>
- Modirrousta, M., & Fellows, L. K. (2008). Dorsal medial prefrontal cortex plays a necessary role in rapid error prediction in humans. *The Journal of Neuroscience*, 28, 14000–14005. <http://dx.doi.org/10.1523/JNEUROSCI.4450-08.2008>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7, 134–140. [http://dx.doi.org/10.1016/S1364-6613\(03\)00028-7](http://dx.doi.org/10.1016/S1364-6613(03)00028-7)
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, 431, 760–767. <http://dx.doi.org/10.1038/nature03015>
- Morecraft, R. J., & Tanji, J. (2009). Cingulofrontal interactions and the cingulate motor areas. In B. A. Vogt (Ed.), *Cingulate neurobiology and disease* (pp. 113–144). Oxford, UK: Oxford University Press.
- Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, 40, 255–265. <http://dx.doi.org/10.1002/dev.10033>
- Mulder, A. B., Nordquist, R. E., Orgüt, O., & Pennartz, C. M. A. (2003). Learning-related changes in response patterns of prefrontal neurons

- during instrumental conditioning. *Behavioural Brain Research*, 146, 77–88. <http://dx.doi.org/10.1016/j.bbr.2003.09.016>
- Mulert, C., Gallinat, J., Dorn, H., Herrmann, W. M., & Winterer, G. (2003). The relationship between reaction time, error rate and anterior cingulate cortex activity. *International Journal of Psychophysiology*, 47, 175–183. [http://dx.doi.org/10.1016/S0167-8760\(02\)00125-3](http://dx.doi.org/10.1016/S0167-8760(02)00125-3)
- Mulert, C., Seifert, C., Leicht, G., Kirsch, V., Ertl, M., Karch, S., . . . Jäger, L. (2008). Single-trial coupling of EEG and fMRI reveals the involvement of early anterior cingulate cortex activation in effortful decision making. *NeuroImage*, 42, 158–168. <http://dx.doi.org/10.1016/j.neuroimage.2008.04.236>
- Munakata, Y., Morton, J. B., & Stedron, J. M. (2003). The role of prefrontal cortex in perseveration: Developmental and computational explorations. In P. Quinlan (Ed.), *Connectionist models of development* (pp. 83–114). East Sussex, UK: Psychology Press.
- Murtha, S., Chertkow, H., Beauregard, M., Dixon, R., & Evans, A. (1996). Anticipation causes increased blood flow to the anterior cingulate cortex. *Human Brain Mapping*, 4, 103–112. [http://dx.doi.org/10.1002/\(SICI\)1097-0193\(1996\)4:2<103::AID-HBM2>3.0.CO;2-7](http://dx.doi.org/10.1002/(SICI)1097-0193(1996)4:2<103::AID-HBM2>3.0.CO;2-7)
- Naccache, L., Dehaene, S., Cohen, L., Habert, M. O., Guichart-Gomez, E., Galanaud, D., & Willer, J. C. (2005). Effortless control: Executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia*, 43, 1318–1328. <http://dx.doi.org/10.1016/j.neuropsychologia.2004.11.024>
- Nagano-Saito, A., Leyton, M., Monchi, O., Goldberg, Y. K., He, Y., & Dagher, A. (2008). Dopamine depletion impairs frontostriatal functional connectivity during a set-shifting task. *The Journal of Neuroscience*, 28, 3697–3706. <http://dx.doi.org/10.1523/JNEUROSCI.3921-07.2008>
- Naito, E., Kinomura, S., Geyer, S., Kawashima, R., Roland, P. E., & Zilles, K. (2000). Fast reaction to different sensory modalities activates common fields in the motor areas, but the anterior cingulate cortex is involved in the speed of reaction. *Journal of Neurophysiology*, 83, 1701–1709.
- Nakao, T., Osumi, T., Ohira, H., Kasuya, Y., Shinoda, J., Yamada, J., & Northoff, G. (2010). Medial prefrontal cortex-dorsal anterior cingulate cortex connectivity during behavior selection without an objective correct answer. *Neuroscience Letters*, 482, 220–224. <http://dx.doi.org/10.1016/j.neulet.2010.07.041>
- Narayanan, N. S., Cavanagh, J. F., Frank, M. J., & Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nature Neuroscience*, 16, 1888–1895. <http://dx.doi.org/10.1038/nn.3549>
- Niv, Y. (2007). Cost, benefit, tonic, phasic: What do response rates tell us about dopamine and motivation? *Annals of the New York Academy of Sciences*, 1104, 357–376. <http://dx.doi.org/10.1196/annals.1390.018>
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53, 139–154. <http://dx.doi.org/10.1016/j.jmp.2008.12.005>
- Onn, S. P., & Wang, X. B. (2005). Differential modulation of anterior cingulate cortical activity by afferents from ventral tegmental area and mediodorsal thalamus. *The European Journal of Neuroscience*, 21, 2975–2992. <http://dx.doi.org/10.1111/j.1460-9568.2005.04122.x>
- O'Reilly, R. C. (2010). The what and how of prefrontal cortical organization. *Trends in Neurosciences*, 33, 355–361. <http://dx.doi.org/10.1016/j.tins.2010.05.002>
- O'Reilly, R. C., Noelle, D. C., Braver, T. S., & Cohen, J. D. (2002). Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. *Cerebral Cortex*, 12, 246–257. <http://dx.doi.org/10.1093/cercor/12.3.246>
- Orr, J. M., & Banich, M. T. (2014). The neural mechanisms underlying internally and externally guided task selection. *NeuroImage*, 84, 191–205. <http://dx.doi.org/10.1016/j.neuroimage.2013.08.047>
- Ostlund, S. B., Wassum, K. M., Murphy, N. P., Balleine, B. W., & Maidment, N. T. (2011). Extracellular dopamine levels in striatal subregions track shifts in motivation and response cost during instrumental conditioning. *The Journal of Neuroscience*, 31, 200–207. <http://dx.doi.org/10.1523/JNEUROSCI.4759-10.2011>
- Ostlund, S. B., Winterbauer, N. E., & Balleine, B. W. (2009). Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *The Journal of Neuroscience*, 29, 8280–8287. <http://dx.doi.org/10.1523/JNEUROSCI.1176-09.2009>
- Parris, B. A., Thai, N. J., Benattayallah, A., Summers, I. R., & Hodgson, T. L. (2007). The role of the lateral prefrontal cortex and anterior cingulate in stimulus-response association reversals. *Journal of Cognitive Neuroscience*, 19, 13–24. <http://dx.doi.org/10.1162/jocn.2007.19.1.13>
- Parvizi, J., Rangarajan, V., Shirer, W. R., Desai, N., & Greicius, M. D. (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. *Neuron*, 80, 1359–1367. <http://dx.doi.org/10.1016/j.neuron.2013.10.057>
- Pasquereau, B., & Turner, R. S. (2013). Limited encoding of effort by dopamine neurons in a cost-benefit trade-off task. *The Journal of Neuroscience*, 33, 8288–8300. <http://dx.doi.org/10.1523/JNEUROSCI.4619-12.2013>
- Passetti, F., Chudasama, Y., & Robbins, T. W. (2002). The frontal cortex of the rat and visual attentional performance: Dissociable functions of distinct medial prefrontal subregions. *Cerebral Cortex*, 12, 1254–1268. <http://dx.doi.org/10.1093/cercor/12.12.1254>
- Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: A review of 107 PET activation studies. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*, 9, R37–R47. <http://dx.doi.org/10.1097/00001756-199806220-00001>
- Peters, Y., Barnhardt, N. E., & O'Donnell, P. (2004). Prefrontal cortical up states are synchronized with ventral tegmental area activity. *Synapse (New York, N.Y.)*, 52, 143–152. <http://dx.doi.org/10.1002/syn.20015>
- Petit, L., Courtney, S. M., Ungerleider, L. G., & Haxby, J. V. (1998). Sustained activity in the medial wall during working memory delays. *The Journal of Neuroscience*, 18, 9429–9437.
- Phillips, P. E., Walton, M. E., & Jhou, T. C. (2007). Calculating utility: Preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology*, 191, 483–495. <http://dx.doi.org/10.1007/s00213-006-0626-6>
- Picton, T. W., Stuss, D. T., Alexander, M. P., Shallice, T., Binns, M. A., & Gillingham, S. (2007). Effects of focal frontal lesions on response inhibition. *Cerebral Cortex*, 17, 826–838. <http://dx.doi.org/10.1093/cercor/bhk031>
- Picton, T. W., Stuss, D. T., Shallice, T., Alexander, M. P., & Gillingham, S. (2006). Keeping time: Effects of focal frontal lesions. *Neuropsychologia*, 44, 1195–1209. <http://dx.doi.org/10.1016/j.neuropsychologia.2005.10.002>
- Pollmann, S., Weidner, R., Müller, H. J., & von Cramon, D. Y. (2000). A fronto-posterior network involved in visual dimension changes. *Journal of Cognitive Neuroscience*, 12, 480–494. <http://dx.doi.org/10.1162/089892900562156>
- Preuss, T. M. (1995). Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *Journal of Cognitive Neuroscience*, 7, 1–24. <http://dx.doi.org/10.1162/jocn.1995.7.1.1>
- Prévost, C., Pessiglione, M., Méteureau, E., Cléry-Melin, M. L., & Dreher, J. C. (2010). Separate valuation subsystems for delay and effort decision costs. *The Journal of Neuroscience*, 30, 14080–14090. <http://dx.doi.org/10.1523/JNEUROSCI.2752-10.2010>
- Procyk, E., & Joseph, J. P. (2001). Characterization of serial order encoding in the monkey anterior cingulate sulcus. *European Journal of Neuroscience*, 14, 1041–1046. <http://dx.doi.org/10.1046/j.0953-816x.2001.01738.x>

- Puig, M. V., & Miller, E. K. (2012). The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron*, 74, 874–886. <http://dx.doi.org/10.1016/j.neuron.2012.04.018>
- Ragozzino, M. E. (2002). The effects of dopamine D(1) receptor blockade in the prelimbic-infralimbic areas on behavioral flexibility. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 9, 18–28. <http://dx.doi.org/10.1101/lm.45802>
- Ragozzino, M. E. (2007). The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Annals of the New York Academy of Sciences*, 1121, 355–375. <http://dx.doi.org/10.1196/annals.1401.013>
- Ragozzino, M. E., Detrick, S., & Kesner, R. P. (1999). Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *The Journal of Neuroscience*, 19, 4585–4594.
- Reynolds, J. R., O'Reilly, R. C., Cohen, J. D., & Braver, T. S. (2012). The function and organization of lateral prefrontal cortex: A test of competing hypotheses. *PLoS ONE*, 7, e30284. <http://dx.doi.org/10.1371/journal.pone.0030284>
- Ribas-Fernandes, J. J., Solway, A., Diuk, C., McGuire, J. T., Barto, A. G., Niv, Y., & Botvinick, M. M. (2011). A neural signature of hierarchical reinforcement learning. *Neuron*, 71, 370–379. <http://dx.doi.org/10.1016/j.neuron.2011.05.042>
- Rich, E. L., & Shapiro, M. (2009). Rat prefrontal cortical neurons selectively code strategy switches. *The Journal of Neuroscience*, 29, 7208–7219. <http://dx.doi.org/10.1523/JNEUROSCI.6068-08.2009>
- Richardson, N. R., & Gratton, A. (1998). Changes in medial prefrontal cortical dopamine levels associated with response-contingent food reward: An electrochemical study in rat. *The Journal of Neuroscience*, 18, 9130–9138.
- Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., & Rushworth, M. F. S. (2006). Separate neural pathways process different decision costs. *Nature Neuroscience*, 9, 1161–1168. <http://dx.doi.org/10.1038/nn1756>
- Rushworth, M. F., Buckley, M. J., Behrens, T. E., Walton, M. E., & Bannerman, D. M. (2007). Functional organization of the medial frontal cortex. *Current Opinion in Neurobiology*, 17, 220–227. <http://dx.doi.org/10.1016/j.conb.2007.03.001>
- Rushworth, M. F. S., Hadland, K. A., Gaffan, D., & Passingham, R. E. (2003). The effect of cingulate cortex lesions on task switching and working memory. *Journal of Cognitive Neuroscience*, 15, 338–353. <http://dx.doi.org/10.1162/089892903321593072>
- Rushworth, M. F. S., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: A combined fMRI and TMS study. *Journal of Neurophysiology*, 87, 2577–2592.
- Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, 8, 410–417. <http://dx.doi.org/10.1016/j.tics.2004.07.009>
- Salamone, J. D., Correa, M., Farrar, A. M., Nunes, E. J., & Pardo, M. (2009). Dopamine, behavioral economics, and effort. *Frontiers in Behavioral Neuroscience*, 3, 13. <http://dx.doi.org/10.3389/neuro.08.013.2009>
- Sallet, J., Quilodran, R., Rothé, M., Vezoli, J., Joseph, J. P., & Procyk, E. (2007). Expectations, gains, and losses in the anterior cingulate cortex. *Cognitive, Affective & Behavioral Neuroscience*, 7, 327–336. <http://dx.doi.org/10.3758/CABN.7.4.327>
- Schmidt, L., d'Arc, B. F., Lafargue, G., Galanaud, D., Czernecki, V., Grabli, D., . . . Pessiglione, M. (2008). Disconnecting force from money: Effects of basal ganglia damage on incentive motivation. *Brain: A Journal of Neurology*, 131, 1303–1310. <http://dx.doi.org/10.1093/brain/awn045>
- Schmidt, L., Lebreton, M., Cléry-Melin, M. L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biology*, 10, e1001266. <http://dx.doi.org/10.1371/journal.pbio.1001266>
- Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*, 10, 885–892.
- Schultz, W. (2013). Updating dopamine reward signals. *Current Opinion in Neurobiology*, 23, 229–238. <http://dx.doi.org/10.1016/j.conb.2012.11.012>
- Schweimer, J., Saft, S., & Hauber, W. (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. *Behavioral Neuroscience*, 119, 1687–1692. <http://dx.doi.org/10.1037/0735-7044.119.6.1687>
- Seamans, J. K., Floresco, S. B., & Phillips, A. G. (1995). Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex. *Behavioral Neuroscience*, 109, 1063–1073. <http://dx.doi.org/10.1037/0735-7044.109.6.1063>
- Seamans, J. K., Lapish, C. C., & Durstewitz, D. (2008). Comparing the prefrontal cortex of rats and primates: Insights from electrophysiology. *Neurotoxicity Research*, 14, 249–262. <http://dx.doi.org/10.1007/BF03033814>
- Seamans, J. K., & Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in Neurobiology*, 74, 1–58. <http://dx.doi.org/10.1016/j.pneurobio.2004.05.006>
- Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *The Journal of Neuroscience*, 27, 8366–8377. <http://dx.doi.org/10.1523/JNEUROSCI.2369-07.2007>
- Servan-Schreiber, D., Printz, H., & Cohen, J. D. (1990). A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science*, 249, 892–895. <http://dx.doi.org/10.1126/science.2392679>
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, 12, 154–167. <http://dx.doi.org/10.1038/nrn2994>
- Shah, A., & Barto, A. G. (2009). Effect on movement selection of an evolving sensory representation: A multiple controller model of skill acquisition. *Brain Research*, 1299, 55–73. <http://dx.doi.org/10.1016/j.brainres.2009.07.006>
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, 79, 217–240. <http://dx.doi.org/10.1016/j.neuron.2013.07.007>
- Sheynikhovich, D., Otani, S., & Arleo, A. (2011). The role of tonic and phasic dopamine for long-term synaptic plasticity in the prefrontal cortex: A computational model. *Journal of Physiology, Paris*, 105, 45–52.
- Shidara, M., Mizuhiki, T., & Richmond, B. J. (2005). Neuronal firing in anterior cingulate neurons changes modes across trials in single states of multitrial reward schedules. *Experimental Brain Research*, 163, 242–245. <http://dx.doi.org/10.1007/s00221-005-2232-y>
- Shidara, M., & Richmond, B. J. (2002). Anterior cingulate: Single neuronal signals related to degree of reward expectancy. *Science*, 296, 1709–1711. <http://dx.doi.org/10.1126/science.1069504>
- Shima, K., & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*, 282, 1335–1338. <http://dx.doi.org/10.1126/science.282.5392.1335>
- Silvetti, M., Alexander, W., Verguts, T., & Brown, J. W. (2014). From conflict management to reward-based decision making: Actors and critics in primate medial frontal cortex. *Neuroscience & Biobehavioral Reviews*, 46, 44–57. <http://dx.doi.org/10.1016/j.neubiorev.2013.11.003>
- Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience*, 5, 75. <http://dx.doi.org/10.3389/fnhum.2011.00075>

- Singh, S., Barto, A. G., & Chentanez, N. (2005). Intrinsically motivated reinforcement learning. *Advances in Neural Information Processing Systems*, 17, 1281–1288.
- Slooman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22. <http://dx.doi.org/10.1037/0033-2909.119.1.3>
- Stark, H., Rothe, T., Wagner, T., & Scheich, H. (2004). Learning a new behavioral strategy in the shuttle-box increases prefrontal dopamine. *Neuroscience*, 126, 21–29. <http://dx.doi.org/10.1016/j.neuroscience.2004.02.026>
- Stefani, M. R., & Moghaddam, B. (2005). Systemic and prefrontal cortical NMDA receptor blockade differentially affect discrimination learning and set-shift ability in rats. *Behavioral Neuroscience*, 119, 420–428. <http://dx.doi.org/10.1037/0735-7044.119.2.420>
- Stefani, M. R., & Moghaddam, B. (2006). Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *The Journal of Neuroscience*, 26, 8810–8818. <http://dx.doi.org/10.1523/JNEUROSCI.1656-06.2006>
- Stephens, D. W., Brown, J. S., & Ydenberg, R. C. (2007). *Foraging: Behavior and ecology*. Chicago, IL: The University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226772653.001.0001>
- Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., . . . Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, 43, 396–417. <http://dx.doi.org/10.1016/j.neuropsychologia.2004.06.010>
- Stuss, D. T., & Knight, R. T. (Eds.). (2002). *Principles of frontal lobe function*. New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195134971.001.0001>
- Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, 66, 449–460. <http://dx.doi.org/10.1016/j.neuron.2010.03.033>
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112, 181–211. [http://dx.doi.org/10.1016/S0004-3702\(99\)00052-1](http://dx.doi.org/10.1016/S0004-3702(99)00052-1)
- Takada, M., Tokuno, H., Hamada, I., Inase, M., Ito, Y., Imanishi, M., . . . Nambu, A. (2001). Organization of inputs from cingulate motor areas to basal ganglia in macaque monkey. *The European Journal of Neuroscience*, 14, 1633–1650. <http://dx.doi.org/10.1046/j.0953-816x.2001.01789.x>
- Thurley, K., Senn, W., & Lüscher, H. R. (2008). Dopamine increases the gain of the input-output response of rat prefrontal pyramidal neurons. *Journal of Neurophysiology*, 99, 2985–2997. <http://dx.doi.org/10.1152/jn.01098.2007>
- Torta, D. M. E., Costa, T., Duca, S., Fox, P. T., & Cauda, F. (2013). Parcellation of the cingulate cortex at rest and during tasks: A meta-analytic clustering and experimental study. *Frontiers in Human Neuroscience*, 7, 275. <http://dx.doi.org/10.3389/fnhum.2013.00275>
- Totah, N. K., Jackson, M. E., & Moghaddam, B. (2013). Preparatory attention relies on dynamic interactions between prelimbic cortex and anterior cingulate cortex. *Cerebral Cortex*, 23, 729–738. <http://dx.doi.org/10.1093/cercor/bhs057>
- Tow, P. M., & Whitty, C. W. M. (1953). Personality changes after operations on the cingulate gyrus in man. *Journal of Neurology, Neurosurgery & Psychiatry*, 16, 186–193. <http://dx.doi.org/10.1136/jnnp.16.3.186>
- Tsuchida, A., & Fellows, L. K. (2009). Lesion evidence that two distinct regions within prefrontal cortex are critical for n-back performance in humans. *Journal of Cognitive Neuroscience*, 21, 2263–2275. <http://dx.doi.org/10.1162/jocn.2008.21172>
- Tunbridge, E. M., Bannerman, D. M., Sharp, T., & Harrison, P. J. (2004). Catechol-o-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. *The Journal of Neuroscience*, 24, 5331–5335. <http://dx.doi.org/10.1523/JNEUROSCI.1124-04.2004>
- Umemoto, A., & Holroyd, C. B. (2014). Task-specific effects of reward on task switching. *Psychological Research*. Advance online publication. <http://dx.doi.org/10.1007/s00426-014-0595-z>
- Uylings, H. B. M., Groenewegen, H. J., & Kolb, B. (2003). Do rats have a prefrontal cortex? *Behavioural Brain Research*, 146, 3–17. <http://dx.doi.org/10.1016/j.bbr.2003.09.028>
- van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: Effects on perseveration and planning. *Acta Psychologica*, 113, 45–65. [http://dx.doi.org/10.1016/S0001-6918\(02\)00150-6](http://dx.doi.org/10.1016/S0001-6918(02)00150-6)
- Venables, L., & Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion*, 33, 63–74. <http://dx.doi.org/10.1007/s11031-008-9116-y>
- Vendrell, P., Junqué, C., Pujol, J., Jurado, M. A., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia*, 33, 341–352. [http://dx.doi.org/10.1016/0028-3932\(94\)00116-7](http://dx.doi.org/10.1016/0028-3932(94)00116-7)
- Venkatraman, V., & Huettel, S. A. (2012). Strategic control in decision-making under uncertainty. *European Journal of Neuroscience*, 35, 1075–1082. <http://dx.doi.org/10.1111/j.1460-9568.2012.08009.x>
- Venkatraman, V., Rosati, A. G., Taren, A. A., & Huettel, S. A. (2009). Resolving response, decision, and strategic control: Evidence for a functional topography in dorsomedial prefrontal cortex. *The Journal of Neuroscience*, 29, 13158–13164. <http://dx.doi.org/10.1523/JNEUROSCI.2708-09.2009>
- Vogt, B. A. (2009). Architecture, neurocytology and comparative organization of monkey and human cingulate cortices. In B. A. Vogt (Ed.), *Cingulate neurobiology and disease* (pp. 65–93). Oxford, UK: Oxford University Press.
- Vogt, B. A., Hof, P. R., Zilles, K., Vogt, L. J., Herold, C., & Palomero-Gallagher, N. (2013). Cingulate area 32 homologies in mouse, rat, macaque and human: Cytoarchitecture and receptor architecture. *The Journal of Comparative Neurology*, 521, 4189–4204. <http://dx.doi.org/10.1002/cne.23409>
- Vogt, B. A., & Paxinos, G. (2014). Cytoarchitecture of mouse and rat cingulate cortex with human homologies. *Brain Structure & Function*, 219, 185–192. <http://dx.doi.org/10.1007/s00429-012-0493-3>
- Vogt, B. A., & Sikes, R. W. (2009). Cingulate nociceptive circuitry and roles in pain processing: The cingulate premotor pain model. In B. A. Vogt (Ed.), *Cingulate neurobiology and disease* (pp. 311–338). Oxford, UK: Oxford University Press.
- Wager, T. D., Jonides, J., Smith, E. E., & Nichols, T. E. (2005). Toward a taxonomy of attention shifting: Individual differences in fMRI during multiple shift types. *Cognitive, Affective & Behavioral Neuroscience*, 5, 127–143. <http://dx.doi.org/10.3758/CABN.5.2.127>
- Wallis, J. D. (2012). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature Neuroscience*, 15, 13–19. <http://dx.doi.org/10.1038/nn.2956>
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, 36, 1870–1884. <http://dx.doi.org/10.1016/j.neubiorev.2012.05.008>
- Walton, M. E., Bannerman, D. M., & Rushworth, M. F. S. (2002). The role of rat medial frontal cortex in effort-based decision making. *The Journal of Neuroscience*, 22, 10996–11003.
- Walton, M. E., Croxson, P. L., Behrens, T. E., Kennerley, S. W., & Rushworth, M. F. (2007). Adaptive decision making and value in the anterior cingulate cortex. *NeuroImage*, 36(Suppl. 2), T142–T154. <http://dx.doi.org/10.1016/j.neuroimage.2007.03.029>

- Walton, M. E., Groves, J., Jennings, K. A., Croxson, P. L., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2009). Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision making. *European Journal of Neuroscience*, 29, 1678–1691. <http://dx.doi.org/10.1111/j.1460-9568.2009.06726.x>
- Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks*, 19, 1302–1314. <http://dx.doi.org/10.1016/j.neunet.2006.03.005>
- Walton, M. E., Rudebeck, P. H., Bannerman, D. M., & Rushworth, M. F. S. (2007). Calculating the cost of acting in frontal cortex. *Annals of the New York Academy of Sciences*, 1104, 340–356. <http://dx.doi.org/10.1196/annals.1390.009>
- Wanat, M. J., Kuhn, C. M., & Phillips, P. E. (2010). Delays conferred by escalating costs modulate dopamine release to rewards but not their predictors. *The Journal of Neuroscience*, 30, 12020–12027. <http://dx.doi.org/10.1523/JNEUROSCI.2691-10.2010>
- Wang, A. Y., Miura, K., & Uchida, N. (2013). The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nature Neuroscience*, 16, 639–647. <http://dx.doi.org/10.1038/nn.3377>
- Wang, S. H., Tse, D., & Morris, R. G. M. (2012). Anterior cingulate cortex in schema assimilation and expression. *Learning & Memory (Cold Spring Harbor, N. Y.)*, 19, 315–318. <http://dx.doi.org/10.1101/lm.026336.112>
- Warden, M. R., Selimbeyoglu, A., Mirzabekov, J. J., Lo, M., Thompson, K. R., Kim, S. Y., . . . Deisseroth, K. (2012). A prefrontal cortex-brainstem neuronal projection that controls response to behavioural challenge. *Nature*, 492, 428–432.
- Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2014). Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology, Paris*. Advance online publication. <http://dx.doi.org/10.1016/j.jphysparis.2014.08.008>
- Welberg, L. (2013). Psychiatric disorders: A motivating microcircuit. *Nature Reviews Neuroscience*, 14, 4–5. <http://dx.doi.org/10.1038/nrn3415>
- Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE*, 8, e68210. <http://dx.doi.org/10.1371/journal.pone.0068210>
- Wheeler, E. Z., & Fellows, L. K. (2008). The human ventromedial frontal lobe is critical for learning from negative feedback. *Brain: A Journal of Neurology*, 131, 1323–1331. <http://dx.doi.org/10.1093/brain/awn041>
- Williams, J. M., Mohler, E. G., & Givens, B. (1999). The role of the medial prefrontal cortex in attention: Altering predictability of task difficulty. *Psychobiology (Austin, Tex.)*, 27, 462–469.
- Winterer, G., Adams, C. M., Jones, D. W., & Knutson, B. (2002). Volition to action—An event-related fMRI study. *NeuroImage*, 17, 851–858. <http://dx.doi.org/10.1006/nimg.2002.1232>
- Womelsdorf, T., Johnston, K., Vinck, M., & Everling, S. (2010). Theta-activity in anterior cingulate cortex predicts task rules and their adjustments following errors. *PNAS National Academy of Sciences of the United States of America*, 107, 5248–5253. <http://dx.doi.org/10.1073/pnas.0906194107>
- Woodward, T. S., Ruff, C. C., & Ngan, E. T. C. (2006). Short- and long-term changes in anterior cingulate activation during resolution of task-set competition. *Brain Research*, 1068, 161–169. <http://dx.doi.org/10.1016/j.brainres.2005.10.094>
- Xu, T. X., Sotnikova, T. D., Liang, C., Zhang, J., Jung, J. U., Spealman, R. D., . . . Yao, W. D. (2009). Hyperdopaminergic tone erodes prefrontal long-term potential via a D2 receptor-operated protein phosphatase gate. *The Journal of Neuroscience*, 29, 14086–14099. <http://dx.doi.org/10.1523/JNEUROSCI.0974-09.2009>
- Yeung, N. (2010). Bottom-up influences on voluntary task switching: The elusive homunculus escapes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 348–362. <http://dx.doi.org/10.1037/a0017894>
- Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 455–469. <http://dx.doi.org/10.1037/0096-1523.29.2.455>
- Young, C. E., & Yang, C. R. (2005). Dopamine D1-like receptor modulates layer- and frequency-specific short-term synaptic plasticity in rat prefrontal cortical neurons. *The European Journal of Neuroscience*, 21, 3310–3320. <http://dx.doi.org/10.1111/j.1460-9568.2005.04161.x>
- Zendehrouh, S., Gharibzadeh, S., & Towhidkhah, F. (2013). Modeling error detection in human brain: A preliminary unification of reinforcement learning and conflict monitoring theories. *Neurocomputing: An International Journal*, 103, 1–13. <http://dx.doi.org/10.1016/j.neucom.2012.04.026>
- Zendehrouh, S., Gharibzadeh, S., & Towhidkhah, F. (2014). Reinforcement-conflict based control: An integrative model of error detection in anterior cingulate cortex. *Neurocomputing: An International Journal*, 123, 140–149. <http://dx.doi.org/10.1016/j.neucom.2013.06.020>

Received May 26, 2014

Revision received September 2, 2014

Accepted September 2, 2014 ■