Neuron
Review

Dopamine Does Double Duty in Motivating Cognitive Effort

Andrew Westbrook[1,*] **and Todd S. Braver**[1,2]
[1]Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, MO 63130, USA
[2]Department of Radiology, Washington University in St. Louis, St. Louis, MO 63110, USA
*Correspondence: jawestbrook@wustl.edu

http://dx.doi.org/10.1016/j.neuron.2015.12.029

Cognitive control is subjectively costly, suggesting that engagement is modulated in relationship to incentive state. Dopamine appears to play key roles. In particular, dopamine may mediate cognitive effort by two broad classes of functions: (1) modulating the functional parameters of working memory circuits subserving effortful cognition, and (2) mediating value-learning and decision-making about effortful cognitive action. Here, we tie together these two lines of research, proposing how dopamine serves "double duty", translating incentive information into cognitive motivation.

Why is thinking effortful? Unlike physical exertion, there is no readily apparent metabolic cost (relative to "rest", which is already metabolically expensive) (Raichle and Mintun, 2006). And yet, we avoid engaging in demanding activities even when doing so might further valuable goals. This appears particularly true when goal pursuit requires extended allocation of working memory for cognitive control. One hypothesis is that cognitive effort avoidance is intended to minimize opportunity costs incurred by the allocation of working memory (Kurzban et al., 2013). If this is true, it suggests not only that working memory is allocated opportunistically, but also that allocation policies entail sophisticated cost-benefit decision-making that is sensitive to as yet unknown cost and incentive functions. In any case, the phenomenon raises a number of questions: How do brains track effort costs? What information is being tracked? How can incentives overcome such costs? What mechanisms mediate adaptive working memory allocation?

Working memory capacity is sharply limited, especially in the domain of cognitive control, involving abstract, flexible, hierarchical rules for behavior selection. Optimizing working memory allocation is thus critical for optimizing behavior. Prevalent computational frameworks have proposed reward- or expectancy-maximization algorithms for working memory allocation (Botvinick et al., 2001; Donoso et al., 2014; O'Reilly and Frank, 2006). Yet, these frameworks largely neglect that working memory allocation itself carries affective valence. High subjective costs drive disengagement, whereas sufficient incentive drives engagement. That is, allocation of working memory is a *motivated* process. In this review, we argue that modulatory functions of the midbrain dopamine (DA) system translate cost-benefit information into adaptive working memory allocation.

DA has been implicated in numerous processes including, but not limited to, motivation, learning, working memory, and decision-making. There are two largely independent literatures that ascribe disparate functional roles to DA with relevance to motivated cognition. First, DA influences the allocation of working memory directly by modulating the functional parameters of working memory circuits. For example, DA tone in the prefrontal cortex (PFC) influences the stability of working memory representations, with higher extrasynaptic tone promoting greater stability, to a limit (Seamans and Yang, 2004). Phasic DA efflux may also push beyond the limit and toggle the PFC into a labile state such that working memory representations can be flexibly updated (Braver et al., 1999). Additionally, DA may support the learning of more sophisticated (and hierarchical) allocation policies via synaptic depression and potentiation in corticostriatal loops (Frank et al., 2001; O'Reilly and Frank, 2006). Second, DA is critical for action selection. Specifically, DA trains value functions for action selection via phasic reward prediction error dynamics potentiating behaviors that maximize reward with respect to effort in a given context (see Niv, 2009 for a review). DA tone in the striatum and the medial PFC also promotes preparatory and instrumental behaviors in response to conditioned stimuli and particularly effortful behavior (Kurniawan et al., 2011; Salamone and Correa, 2012).

Here, we tie together these largely independent lines of research by proposing how the very same functional properties of DA encoding incentive information translate incentives into cognitive motivation by regulating working memory. Specifically, we propose that DA dynamics encoding incentive state promote subjectively costly working memory operations experienced as conscious, phenomenal effort. As we detail below, our proposal makes use of the concept of a "control episode" during goal pursuit (cf. "attentional episodes", see Duncan, 2013), involving stable maintenance of the goal state at higher-levels of the control hierarchy, along with selective updating of lower level rules for guiding behavior during completion of subgoals, as progress is made toward the ultimate goal state. We review the ways in which DA dynamics encoding a net cost-benefit of goal engagement and persistence result in adaptive working memory allocation. As such, DA translates incentive motivation into cognitive effort.

Motivated Cognition

Why Cognitive Effort Matters

Cognitive effort is an everyday experience. The subjective costliness of cognitive effort is consequential, sometimes driving disengagement from otherwise highly valuable goals.

Yet, surprisingly little is known about this phenomenon. It is neither clear what makes tasks effortful, nor why task engagement is apparently aversive in the first place (Inzlicht et al., 2014; Kurzban et al., 2013).

Beyond a quizzical influence over goal-directed behavior, there are numerous reasons to care about cognitive effort. First, expenditure is critical for career and educational success, economic decision-making, and attitude formation (Cacioppo et al., 1996; von Stumm et al., 2011). Second, deficient effort may be a significant component of neuropsychiatric disorders for which avolition, anhedonia, and inattention feature prominently, such as attention deficit hyperactivity disorder (ADHD) (Volkow et al., 2011), depression (Hammar et al., 2011), and schizophrenia (Strauss et al., 2015). Effort avoidance may also contribute to declining cognitive performance in healthy aging (Hess and Ennis, 2012; Westbrook et al., 2013). Engagement with certain kinds of cognitive tasks appears negatively valenced, indicating a subjective cost. Subjectively inflated effort costs might undermine cognitive engagement and thereby performance.

Control-Demanding Tasks Are Valenced

Not all tasks are effortful. Tasks requiring allocation of working memory for cognitive control, however, appear to be (Botvinick et al., 2009; Dixon and Christoff, 2012; Dreisbach and Fischer, 2012; Kool et al., 2010; Massar et al., 2015; McGuire and Botvinick, 2010; Schouppe et al., 2014; Westbrook et al., 2013). Individuals allowed to select freely between tasks differing only in the frequency with which working memory must be reallocated for cognitive control express a progressive preference for the option with lower reallocation demands (Kool et al., 2010; McGuire and Botvinick, 2010). Critically even when offered larger reward, decision-makers discount reward as a function of effort costs, thus selecting smaller reward with lower demands over larger reward with higher demands (Massar et al., 2015; Westbrook et al., 2013).

Under what conditions might cognitively demanding tasks acquire affective valence? By one account, tasks demanding cognitive control involve response conflict (Botvinick et al., 2001) or frequent errors (Brown and Braver, 2005; Holroyd and Coles, 2002) and as such are less likely to be successful, thus engendering avoidance learning to bias behavior toward tasks with higher chances of success (Botvinick, 2007). Multiple lines of evidence suggest that conflict is aversive. First, conflict in the context of a Stroop task predicts overt avoidance (Schouppe et al., 2012). Also, trial-wise variation in subjective frustration with a stop-signal task predicts BOLD signal in the anterior cingulate cortex (ACC), otherwise implicated in conflict detection (Spunt et al., 2012). In another study (McGuire and Botvinick, 2010), participant ratings of their desire to avoid a conflict-inducing task correlated positively with individual differences in recruitment of ACC and also dorsolateral PFC, putatively involved in working memory maintenance of task sets. Moreover, the dorsolateral PFC correlation remained after controlling for performance differences (reaction time, RTs, and error rates), indicating that the desire to avoid the task did not simply reflect perceived failure. Finally, interesting interactions between affect and cognitive control also support the notion that conflict is aversive (Dreisbach and Goschke, 2004; Saunders and Inzlicht, 2015; Shackman et al., 2011). For example, individuals respond faster to affectively negative, and slower to affectively positive stimuli, following priming by conflicting versus non-conflicting Stroop trials (Dreisbach and Fischer, 2012).

Avoidance learning to minimize loss may partly explain aversion to working memory allocation for cognitive control. Yet, it cannot be the full story. On the one hand, individuals avoid cognitive demand, even controlling for reward likelihood (Kool et al., 2010; McGuire and Botvinick, 2010; Westbrook et al., 2013). On the other, opportunity costs may reflect more than just the likelihood of failure during the current control episode; namely, they may reflect the value of missed opportunities (Kurzban et al., 2013). Finally, an adaptive system must also be judicious, and avoidance of all goals requiring cognitive control is clearly maladaptive. Decision-making must consider both costs and benefits. Indeed, there is growing evidence that the ACC is as important for biasing engagement with effortful, control-demanding tasks as it is for biasing avoidance (Shenhav et al., 2013).

Incentives Motivate Cognitive Control

If control is avoided because of subjective costs, increased incentives could offset costs, promoting control. Indeed, incentives yield control-mediated performance enhancements (see Botvinick and Braver, 2015; Pessoa and Engelmann, 2010 for review). Incentives enhance performance in control-demanding tasks encompassing visuospatial attention (Krebs et al., 2012; Small et al., 2005), task-switching (Aarts et al., 2010), working memory (Jimura et al., 2010), and context maintenance (Chiew and Braver, 2014; Locke and Braver, 2008), among others. Furthermore, incentives predict greater activity in control-related regions, including medial and lateral PFC. For example, incentives yield increased BOLD signal in the ACC, propagating to dorsolateral PFC, corresponding well with the canonical model by which the ACC monitors for control demands and recruits lateral PFC to implement control (Kouneiher et al., 2009). This particular study showed that incentives yielded an additive increase in BOLD signal, on top of demand-driven control signals. However, more recent work has shown that incentive information is not merely additive, but interactive: with increasing incentive-related activity under high task-demand conditions, thus more directly implicating incentives in the enhancement of cognitive control (Bahlmann et al., 2015), cf. Krebs et al. (2012). Beyond mean activity, incentives also enhance the fidelity of working memory representations. Task set representations are more distinctive, as revealed by multivariate pattern analysis of BOLD data, during incentivized working memory trials (Etzel et al., 2015). Interestingly, increased distinctiveness predicts individual differences in incentive-driven behavioral enhancement.

Incentives not only drive *more* control-related activity, or higher fidelity task set representations, but they also affect the selection of more costly control strategies. For example, cognitive control may be recruited proactively, in advance of imperative events, or reactively, concurrent with event onset (Braver, 2012). Proactive control has behavioral advantages, but also incurs opportunity costs that bias reliance on reactive control. Incentives appear to offset costs, increasing proactive relative to reactive control, as reflected in sustained increases

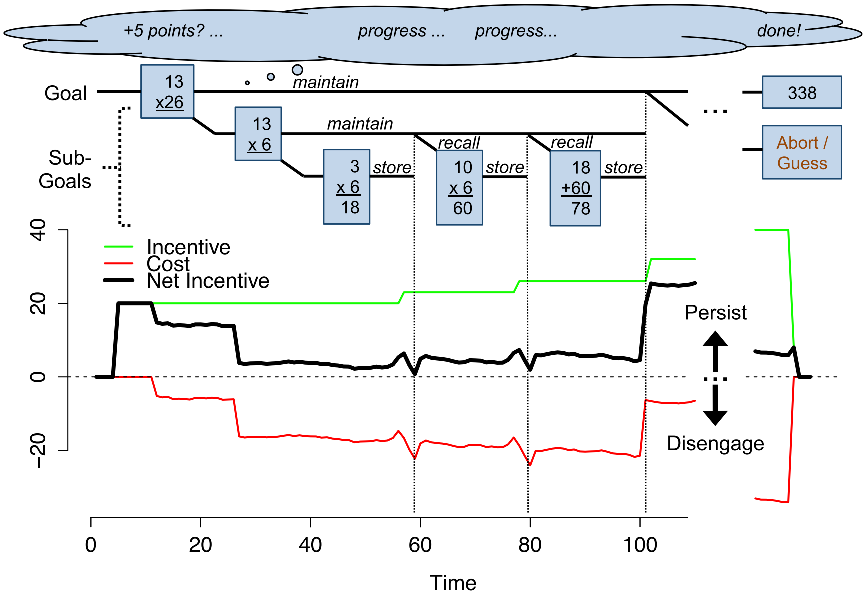

Figure 1. Incentive State Dynamics during a Control Episode

State dynamics as exemplified by succession through mental multiplication task operations. In the image, points incentivize initial engagement. The costs (red line) mount with time-on-task and increasing maintenance and updating demands. The actors persist while the net incentive value of engagement (black line) remains positive, which occurs when costs are offset by incremental progress (e.g., at subgoal completion) and other incentives (green line). If the net incentive value goes negative, the actors are prone to disengagement.

in BOLD signal prior to imperative events, and attenuated phasic responses at event onsets, and this shift to proactive control predicts performance enhancements (see Jimura et al., 2010). Moreover, incentive-driven shifts to proactive control are larger among highly reward-sensitive individuals (Jimura et al., 2010).

In sum, working memory operations are treated as subjectively costly. Whether apparent costliness reflects avoidance learning of behaviors with low likelihood of success, or opportunity costs, incentives can counterbalance costs, promoting working memory operations. Cost-benefit decision-making thus underlies working memory allocation for cognitive control. We propose that during goal pursuit, individuals engage in costly control episodes, remaining engaged to the extent that benefits outweigh costs. Moreover, we propose that DA solves a core computational problem of control episodes: namely, value-based management of working memory for cognitive control that reflects not only prior reward learning, but also instantaneous effects of current incentive state.

To illustrate, we consider an example control episode involving the demanding task of finding the product of two two-digit numbers, incentivized by points on an examination (without calculators; Figure 1). Control episodes may be initiated by incentive-driven (point-value cued) allocation of working memory to represent the goal state (finding the product). Throughout an episode, the actor must maintain high-level goal information (e.g., the original numbers), resisting interference from distractors, while flexibly updating targeted, lower-level representations of subgoals in a hierarchical fashion. Subgoals in our example include: (1) multiplying the ones column digits; (2) carrying the tens-digit value of that product; (3) adding that value to the product of the tens-digits, etc. Maintaining each subgoal is subjectively costly and thus the stability of goal representations should reflect the value of those goals. Similarly, updating operations, as required when subgoals are completed, are also subjectively costly. As each stage has its own costs, and costs may accumulate in excess of perceived benefits, any stage may result in disengagement. We consider the mental multiplication example for illustrative purposes only; the general notion of a control episode should apply broadly to any hierarchically structured, temporally extended sequence of goal-directed behaviors that require working memory allocation (e.g., planning, problem-solving, and reasoning).

In the sections that follow, we describe how DA mediates value-based working memory management during control episodes. Figure 2 provides an overview of critical functions that will be reviewed. Tonic DA, for example, influences the stability of working memory contents by direct action in PFC (Figure 2B), while phasic DA efflux in the striatum trains policies for value-based updating of working memory contents that reflect both the reward value of the goals to which they correspond and effort (updating and maintenance) costs (Figure 2C). While cached value-functions reflect past experience, their implementation is subject to instantaneous modulation by incentive state. Accordingly, we describe how DA and its projection targets encode net incentive state, dynamically accounting for goal state revaluation and generalized motivation. Such information is used to bias policies for working memory allocation actions (Figure 2D). Hence, DA does double duty in translating incentive information into cognitive effort both by functional modulation of working memory circuits (Figures 2B and 2C) and by influencing value-learning and decision-making about effortful action (Figures 2C and 2D). We take up each of these key duties in turn.

DA and Working Memory Management

Successful control episodes demand stable maintenance and also targeted, flexible updating of working memory, with DA appearing to play an important role in both processes. In the PFC, DA influences the stability of recurrent networks (Brunel and Wang, 2001; Seamans and Yang, 2004) and, thereby, the stability of short-term configurations that constitute control-related working memory representations (Cools and D'Esposito, 2011; Robbins and Arnsten, 2009). In the striatum, DA trains gating policies that come to determine the kinds of information that become represented in the PFC and the stimulus signals that drive updating of specific PFC subregions (Frank et al., 2001;

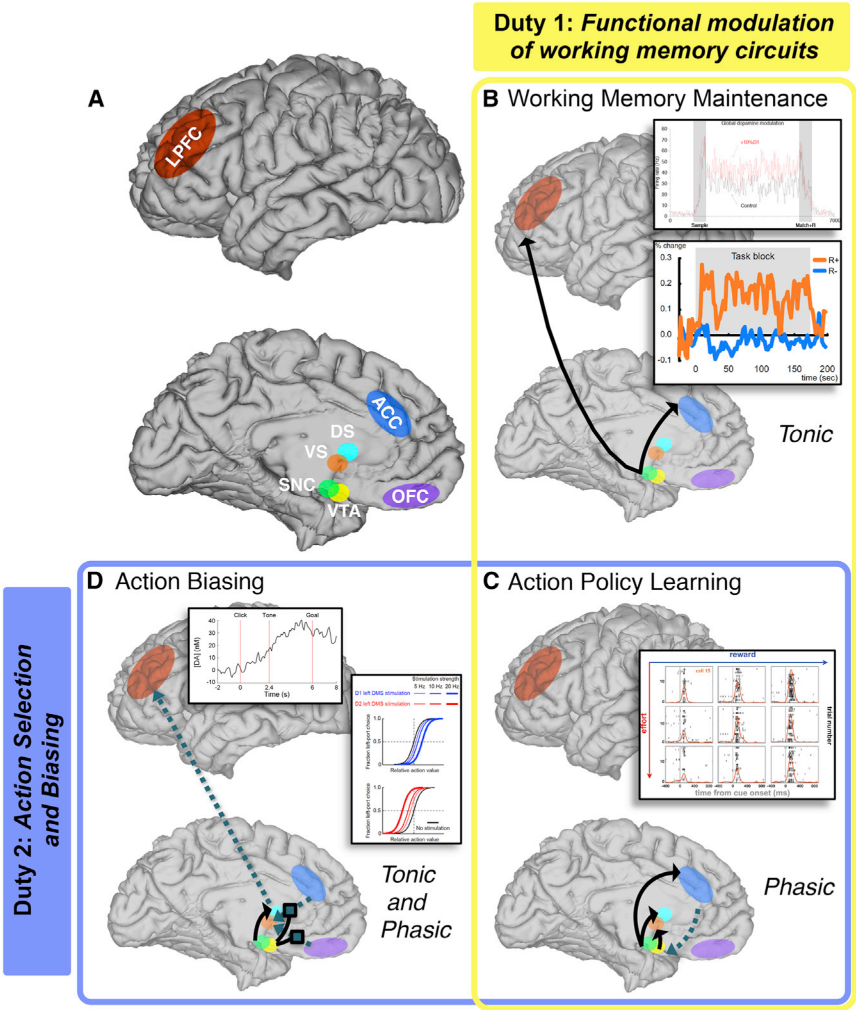

Figure 2. Dopamine Does Double Duty during Control Episodes

Double duty for DA in cognitive effort includes: (1) modulating the functionality of WM circuits including maintenance stability and specific flexibility for updating WM contents (yellow), and (2) developing and biasing value-based policies for WM allocation (blue). The images clockwise from upper left: (A) key anatomical loci of DA circuitry regulating control episodes; (B) tonic DA promotes stable and robust WM maintenance via PFC modulation; (C) phasic DA release encoding effort-discounted reward trains allocation policies in striatum and ACC; and (D) phasic DA release and ramping tone in the striatum bias action selection toward costly WM updating in the lateral PFC, by potentiating updating generally, and updating in accordance with PFC-based action policy signals, in particular. The top-down policy signals reflect hierarchically higher-level goals and thus favor gating of contextually appropriate subgoals into WM. The insets are described in subsequent figures.

O'Reilly and Frank, 2006). Thereby, DA plays key roles in initiating and sustaining control episodes by functionally promoting both working memory stability and targeted flexibility.

Promoting Stability of Higher-Order Goal Representations

Working memory representations in the PFC (Miller and Cohen, 2001) (though see Riggall and Postle, 2012) are instantiated as temporarily stable, recurrent cortical pyramidal networks (Brunel and Wang, 2001). Extracellular DA promotes recurrent dynamics by increasing excitatory N-methyl-D-aspartate (NMDA) drive and also pruning firing external to such networks by exciting inhibitory gamma-Aminobutyric acid (GABA) interneurons (Berridge and Arnsten, 2013; Cools and D'Esposito, 2011; Seamans and Yang, 2004). The net effect of increasing DA (to a point) is to increase network-specific recurrent firing rates (Figure 3) and thus signal-to-noise ratio of working memory representations (Brunel and Wang, 2001). Evidence includes that, DA D1 receptor agonism sharpens spatial tuning in task-relevant PFC neurons in monkeys performing a spatial working memory task (Vijayraghavan et al., 2007).

Importantly, PFC DA changes dynamically, precisely when needed, to promote working memory maintenance. Salient, cognitive task-relevant events have been shown to drive mesocortical DA neuron firing that can increase extrasynaptic DA concentration in the PFC (Figure 2B) (reviewed in Bromberg-Martin et al., 2010; Phillips et al., 2008). In humans, BOLD dynamics in the ventral tegmental area (VTA) support the hypothesis that DA neurons respond to cognitive task demands, independently of reward (see Boehler et al., 2011), as well as the interaction of reward and task complexity (Krebs et al., 2012). The effect of this VTA activation may be to promote maintenance of task sets in lateral PFC regions, e.g., in those demonstrated (by reversible TMS lesion) to be critical for supporting rule-guided behavior (D'Ardenne et al., 2012). Importantly, a positron emission tomography (PET) study has revealed increased D2 receptor binding in ventrolateral PFC in humans performing a verbal working memory task, relative to a simpler sustained attention task (Aalto et al., 2005) (Figure 4A).

Incentive cues also drive PFC DA release (reviewed in Bromberg-Martin et al., 2010; Phillips et al., 2008). To the extent that incentive-related DA promotes robust maintenance, such effects help explain motivational enhancements of memory- and rule-guided behavior. It could, for example, explain why incentives predict stronger proactive, maintenance-related BOLD signal in the lateral PFC during a Sternberg-type working memory task that mediates better performance (Jimura et al., 2010) (Figure 4B). It could also explain performance enhancements following pharmacological COMT inhibition (boosting PFC DA tone, in particular) in an exploration/exploitation task, which requires the tracking of multiple value signals in working memory (Kayser et al., 2014).

Conversely, while increasing DA promotes maintenance, flexible shifting may require decreased DA. In one study, set-shifting

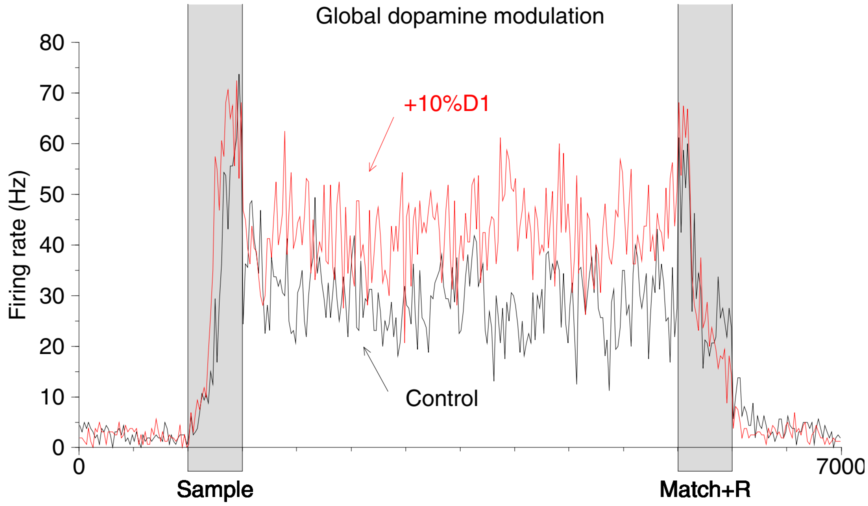

Figure 3. Influence of Dopamine on Neuronal Dynamics in a Model Recurrent Network

Increased PFC DA tone (upper line) boosts firing in task-selective neurons in recurrent networks during WM maintenance (e.g., delayed match-to-sample), relative to a baseline (control) dopaminergic state (lower line). In this computational simulation of neural dynamics, DA-linked increase in NMDA and GABA currents boosts persistent, recurrent firing, enhancing the stability and distractor resistance of task-relevant WM representations (x axis units are arbitrary time; Brunel and Wang, 2001).

performance was modulated in humans dosed with l-dopa. fMRI evidence localized these effects to the PFC (Shiner et al., 2015). Specifically, when participants were dosed with l-dopa, the difference between better performance on incentivized, and worse performance on non-incentivized trials, was removed. Critically, this mirrored the attenuation of BOLD signal deactivation in the ventromedial PFC typical on incentivized versus non-incentivized trials. The result was interpreted as evidence that set maintenance was under dopaminergic control, and this control must be transiently removed to shift task sets.

Too much extrasynaptic DA, on the other hand, may destabilize working memory representations (Berridge and Arnsten, 2013; Cools and D'Esposito, 2011; Seamans and Yang, 2004). One potential mechanism of supraoptimal DA effects is increasing stimulation of relatively low-affinity DA D2 receptors (Durstewitz and Seamans, 2008). This D2 stimulation leads to decreased GABA and NMDA currents, thus counteracting D1 activation effects. Blocking D2 action, therefore, could enhance PFC representations. In a recent demonstration, DA D2 receptor blockade by amisulpride, relative to placebo, enhanced PFC representations as indexed by sharper multivariate pattern discrimination of PFC BOLD data between incentive conditions during an incentive learning task (Kahnt et al., 2015).

According to one proposal, task-based DA release yielding supraoptimal DA may provide a local task-switching mechanism in the PFC (Braver et al., 1999). Specifically, DA release may toggle PFC lability, by pushing DA tone from optimal to supraoptimal levels, increasing the likelihood of context updating during task performance. However, as noted, this kind of updating would have diffuse influence and lacks the temporal and spatial specificity required for targeted updating of, for example, a subcomponent of a task-set hierarchy (O'Reilly and Frank, 2006). Even if phasic PFC DA efflux does not support selective updating, it may be useful to serve as a general updating or disengagement signal.

We close this section by noting that while increasing incentive can drive higher PFC DA tone, a recent study has shown conflicting results. Notably, the investigators found higher PFC DA release in anticipation of less subjectively valued outcomes in monkeys (Kodama et al., 2014). The authors interpreted this unexpected result as reflecting stress-driven DA release observed in other studies (see Butts et al., 2011). Thus incentive can promote PFC DA tone, but stress may be another affective determinant. In any case, there is a growing consensus that affective stimuli influence PFC DA tone which, in turn, modulates the stability of recurrent networks and, thereby, the contents of working memory.

Promoting Targeted, Flexible Updating of Task Sets

The need for both stability and flexibility of working memory, during control episodes, creates opposing demands that DA acting by the PFC alone cannot resolve. Indeed, DA-mediated increases in stability undermine flexibility, as reflected in higher task-switch costs (Herd et al., 2014; van Schouwenburg et al., 2010). There is evidence, however, that DA can increase cognitive flexibility via D2 signaling in the ventral striatum (VS) (Aarts et al., 2010; Samanez-Larkin et al., 2013; Shiner et al., 2015; van Holstein et al., 2011). Incentives can enhance task switching, and this effect is stronger among individuals with a variant of the DA transporter gene DAT1 predicting lower transporter density, and therefore higher synaptic and extrasynaptic DA tone, particularly in the striatum (Aarts et al., 2010). This result supports the hypothesis that striatal DA release mediates incentive enhancement of cognitive flexibility. Evidence of D2 receptor involvement comes from a study comparing the effects of the DA agonist bromocriptine and the DA D2-selective antagonist sulpiride on task-switching (van Holstein et al., 2011). Critically, those individuals with DAT1 coding for higher transporter density (lower striatal DA tone) showed reduced switch costs after being dosed with DA agonist bromocriptine, and this improvement was blocked by the D2-selective antagonist sulpiride.

Successful control episodes require not simply generalized increases in flexibility, but targeted, context-specific updating. In the mental arithmetic example, it is critical to maintain a representation of the full problem (13 × 26), while updating specific subgoals as they are completed (e.g., shifting to multiply 10 × 6, after storing the 3 × 6 result). While DA in the PFC lacks the temporal or spatial specificity to support targeted updating, DA can effect specific updating via the basal ganglia (Frank et al., 2001; O'Reilly and Frank, 2006) (Figure 2C). A well-supported model holds that phasic DA release in the dorsal striatum (DS) trains "so-called 'Go'" cell (D1-expressing medium spiny

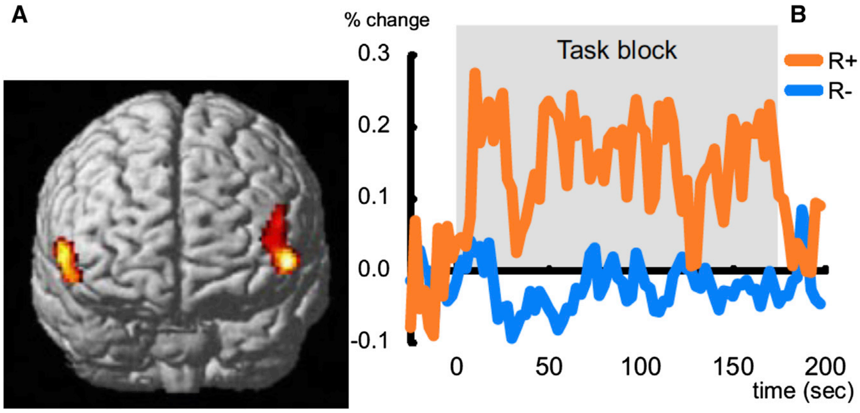

Figure 4. Control Demand and Incentive Effects of Dopamine in Lateral PFC

(A) Decreased binding potential of D2 receptors in ventrolateral PFC indicates increased DA tone during a verbal WM (2-back) task relative to a less demanding sustained attention (0-back) task (Aalto et al., 2005).

(B) In a high-incentive context (orange; R+), sustained BOLD signal is enhanced in right lateral PFC during a WM (Sternberg) task, relative to a low incentive context (blue; R−) (Jimura et al., 2010).

neurons) synapses, through long-term potentiation (LTP), which increase the likelihood of contextual information being gated to the PFC. DA dips, on the other hand, are proposed to train "so-called 'NoGo'" cells (D2-expressing medium spiny neurons) synapses, through long-term depression (LTD), decreasing the likelihood of context gating (Frank et al., 2001). When stimuli evoke activity in relatively more Go than NoGo cells, information is gated (by transient removal of tonic inhibition of the thalamus) for representation into the PFC. Thus, by training striatal synapses to reflect reward history, phasic DA dynamics generate cached policies governing context-specific updating of working memory.

The gating model has been extended to support the hierarchical structure of control episodes (Chatham and Badre, 2015). Corticostriatal loops may support DA-mediated hierarchical reinforcement learning, in which content is selected for updating at different levels of a hierarchy (Badre and Frank, 2012; Frank and Badre, 2012). Reciprocal connections allow the basal ganglia (BG) to not only direct which information gets gated into working memory, but also for higher-level PFC representations of context to direct what lower-level representations get out-gated, when they are no longer useful (Chatham et al., 2014). Thus, higher-level representations may interact in a top-down manner with bottom-up gating mechanisms to adaptively target content at a hierarchically lower level.

Successful control episodes are enabled by both: (1) DA-trained cached, value-based gating policies in cortico-striatal circuits that bias adaptive updating in hierarchical environments, and (2) DA-mediated stability (in the PFC) and flexibility (in the striatum) of working memory as a function of incentive information. Thus, DA appears to translate incentives into cognitive motivation by direct modulation of the cortico-striatal working memory network supporting control episodes.

Cost-Benefit Decision Making

Control episodes are treated as subjectively costly. Behavioral evidence suggests cost-benefit decision-making, balancing the value of the desirable goal against an underlying cost function (Dixon and Christoff, 2012; Kool et al., 2010; Massar et al., 2015; Westbrook et al., 2013) and DA likely plays a key role. Indeed, DA has long been implicated not only in working memory (WM) and motivation, but also in both value-learning and decision-making. Specifically, DA trains functions mapping value to external states along with cognitive and motor actions (Li and Daw, 2011; Wickens et al., 2007), and temporally extended action sequences (Holroyd and Yeung, 2012; O'Reilly et al., 2014). Incentive salience models propose that DA may further bias action selection at the time of choice by modulating value signals, e.g., as a function of motivational state (McClure et al., 2003; Zhang et al., 2009). So, for example, DA could mediate the decision to engage in a temporally extended sequence of cognitive actions required for multiplication of two-digit numbers, as well as to execute all subgoals in sequence, as a function of the point value on an examination. In contrast, if the task were insufficiently incentivized, or if a lower-effort strategy was available (e.g., using a calculator), the decision process may instead resolve against control episode engagement. In the next sections, we review evidence for the role of DA in training value functions and also instantaneously biasing the selection of, and persistence with, costly cognitive actions.

DA and Action Policy Learning

Reward-Prediction Errors

A rich literature implicates the firing of midbrain DA cells in encoding the momentary difference between expected and actual reward (Schultz et al., 1997). The remarkable functional similarity between these reward prediction errors (RPEs) and temporal difference values in computational reinforcement learning (RL) has led to the hypothesis that phasic DA dynamics train the system to bias behaviors that increase context-based reinforcement likelihood (Montague et al., 1996). Mechanistically, DA does so by potentiating synapses linking representations of the current state to specific behaviors (Wickens et al., 2007). Synaptic weights acquired through this process can be thought of as value functions, in the sense of stronger weights biasing actions that maximize the likelihood of reward (i.e., those actions with greatest expected value). This extends to cognitive actions. Indeed, it is precisely these phasic DA RPEs that are thought to train WM gating policies described in the previous section (Frank et al., 2001) (Figure 2C).

Critically, the functional capacity of RPE signals extends beyond simple stimulus-response pairings, to action-outcome association learning in the PFC (Gläscher et al., 2009). From an action selection standpoint, this is enormously powerful. Foremost, action-outcome associations are necessary for calculating net incentive value: the expected benefits of outcomes less the cost of actions. Moreover, action-outcome associations can not only support selecting the most highly rewarded action in a given state, they also enable "looking forward": selecting actions based upon an internal model of the environment, its states and action-contingent state transitions. An agent acting in a "model-based" fashion may select actions that also take into account its state motivation for particular outcomes (Daw et al., 2011; Gläscher et al., 2009). Indeed, sensitivity to outcome devaluation (e.g., devaluation by selective satiation) is used as the benchmark of model-based decision-making (Dolan and Dayan, 2013).

There is evidence that RPEs can reflect internal models of actions and subsequent states (Hiroyuki, 2014). Hence, value functions may be learned for allocating WM, if doing so implements a mental state that increases the probability of reward, given subsequent actions (Chatham and Badre, 2013; Dayan, 2012). Thus, RPEs may train value functions governing WM allocation. Evidence of value-based WM allocation comes from an fMRI study of humans selecting among task sets manipulated to have variable utility (Chatham and Badre, 2013). The expected value of task sets was varied systematically over trials, and a RL model of choice behavior was used to predict trial-wise subjective values of task sets. Subjective value estimates predicted BOLD dynamics in a fronto-striatal network, supporting that task set values are tracked according to a value-updating algorithm that is likely mediated by phasic DA RPE signals. It is worth noting here that, although PFC DA is thought have slow clearance, which would preclude the temporal resolution and specificity required for precise DA-based training, per se, there is reason to believe that co-release of glutamate from DA cells innervating the PFC could provide the mechanism for synaptic learning effects (Seamans and Yang, 2004). DA cells may thus direct learning, whether by the functional consequences of glutamate in the PFC or DA in the striatum. As discussed above, however, DA release in the PFC may have further consequences in the PFC in terms of promoting WM stability, by modulating the dynamics of recurrently firing networks of pyramidal cells.

The functionality of RPE signals may also extend to hierarchical RL, whereby value functions describe actions sequences rather than individual actions (Frank and Badre, 2012; O'Reilly et al., 2014; Ribas-Fernandes et al., 2011). Selection across sequences is critical for overcoming individually costly actions that are only justifiable given the value of desirable outcomes at sequence conclusion (Holroyd and Yeung, 2012). In the mental arithmetic example, updating WM with a ones-digit multiplication subgoal is costly, but may be justifiable with regard to the progress it incurs toward the ultimate, valuable goal of solving the two-digit multiplication problem. Importantly, knowledge of task hierarchy enables agents to bias such costly actions. Pseudo-RPEs (based on perceived progress rather than external reward) may train value functions regarding action sequences (Ribas-Fernandes et al., 2011). Thus RPEs may train progress-based value functions for sequences of effortful WM updating and maintenance.

As we have just reviewed, DA-mediated RL appears to train value functions with numerous properties supporting successful control episodes. Namely, RPEs can train value functions based on action-outcome associations, supporting model-based prospection, and reflecting action sequences. Such value functions may thus promote action in hierarchically structured environments where individually costly actions, like WM allocation, are justified inasmuch as they incur progress toward a goal that is more valuable than the sequence is costly. As we elaborate next, value functions within the ACC in particular, appear critical for biasing engagement and persistence with costly control episodes.

DA Cell Firing Trains Action-Outcome Associations in the ACC

The ACC and dopaminergic innervation of the ACC are critical for selecting effortful behavior (Kurniawan et al., 2011). In particular, RPE signals may train action-outcome associations in the ACC for prediction (Alexander and Brown, 2011; Donoso et al., 2014; Holroyd and Coles, 2002) and effort-based decision-making (Kennerley et al., 2011; Shenhav et al., 2013; Skvortsova et al., 2014). Action-outcome associations are necessary for cost-benefit computations. Unit recording studies in monkeys engaged in multi-attribute decision-making have uncovered ACC neurons multiplexing information about benefits and costs (including effort) in a unified value-coding scheme (Kennerley et al., 2009). This contrasts with the orbitofrontal cortex (OFC), also implicated in economic decision-making, which contains neurons encoding the value of multi-attribute outcomes, but not the cost of action to obtain such outcomes (Kennerley et al., 2011; Padoa-Schioppa, 2011). Consistent data showing multiplexed cost-benefit encoding in the ACC also come from rodent studies (Cowen et al., 2012; Hillman and Bilkey, 2012).

There is also considerable evidence supporting cost-benefit encoding in the human ACC during effort anticipation and decision-making. In tasks utilizing advance reward and demand (effort) cues, the ACC is sensitive to the anticipation of both dimensions, in both forced- and free-choice trials (Croxson et al., 2009; Kroemer et al., 2014; Kurniawan et al., 2010, 2013; Massar et al., 2015; Prévost et al., 2010; Vassena et al., 2014). Moreover, the ACC has been repeatedly linked to the conscious experience of cognitive effort. In a striking demonstration, electrical stimulation of the human ACC reliably evoked the conscious experience of a forthcoming challenge and also a "will to persevere" through that challenge (Parvizi et al., 2013).

Tonic PFC DA Strengthens Cortical Action Policy Signals

Multiplexed value information is used by the ACC to set action policies which can then be implemented via the BG, in competition with habitual biases against effortful engagement. In the domain of cognitive effort, the ACC has been proposed to subserve a specific computational function in selecting the identity of, and the intensity with which control signals are represented, as a function of the expected value of the associated outcome (Shenhav et al., 2013). In this context, DA in the ACC strengthens dynamics supporting representation and integration of action-outcome associations (as evidenced, e.g., by increasing power in gamma band oscillations, see Steullet et al., 2014) and may thereby increase the influence of ACC-based policy signals. Conversely, blocking DA diminishes the capacity of the ACC to bias the choice of greater effort for larger reward (Schweimer and Hauber, 2006; Schweimer et al., 2005). Thus, in the mental arithmetic example, incentive-driven DA release in the ACC promotes cortical action policies related to the strategy of directly computing the solution to the two-digit multiplication problem, rather than following a prepotent bias to utilize a lower-effort strategy (i.e., guessing) or otherwise disengage.

DA and the ACC Track Progress to Regulate Persistence

Following control episode initiation, an actor must decide whether to persist. Opportunity costs rise with time-on-task and so may the drive to disengage. As we propose, perceived progress implies increasing expected value, and thus may offset accruing opportunity costs. There is growing evidence that DA and the ACC regulate progress-based persistence with control episodes (Holroyd and Yeung, 2012; O'Reilly et al., 2014). In fact, in rats engaged in an effort-based decision-making task, ACC neurons multiplexing maze path, reward, and effort information were most selective after decisions were made, and their dynamics were identical across forced- and free-choice trials, suggesting greater involvement in biasing persistence than in initial selection (Cowen et al., 2012).

Control episodes are intrinsically costly, perhaps reflecting opportunity costs incurred by WM allocation (Inzlicht et al., 2014; Kurzban et al., 2013). Adaptive persistence in effortful sequences of behavior, therefore, requires ongoing computation of accruing costs and benefits (Meyniel et al., 2013). A useful metric is the rate of progress, if progress is sufficiently fast, engagement is maintained, while slow or blocked progress yield frustration and disengagement (O'Reilly et al., 2014).

The ACC, by virtue of its capacity for hierarchical RL, and reciprocal interactions with the DA midbrain (Holroyd and Yeung, 2012; Ribas-Fernandes et al., 2011), is well-positioned to track progress and regulate engagement. By this account, the ACC (perhaps in concert with the OFC, see O'Reilly et al., 2014) uses representations of hierarchical task structure to track progress toward sub and superordinate goals and conveys progress via the dopaminergic midbrain. Faster progress generates DA release, promoting value learning and engagement, while slower progress generates DA dips. Indeed, ACC unit recordings in both monkeys and rats show ramping dynamics that reflect increasing progress through action sequences (Ma et al., 2014). Importantly, this dynamic reflects internal models of task structure: rat ACC neurons track progression through a sequence of lever presses, regardless of physical lever features or of particular sequences required on a given trial (Ma et al., 2014).

The midbrain, for its part, shows RPE-like firing in response to perceived (progress-like) success in monkeys performing a visual WM task, independent of actual success, implicating model-based criteria (Matsumoto and Takada, 2013). Also, VS BOLD signal in humans performing a WM task increases transiently on correct versus incorrect trials, in the absence of performance feedback (Satterthwaite et al., 2012). Together, these results suggest not only that pseudo-RPEs report perceived goal progress, but also that one function of these pseudo-RPE signals is to modulate activity in the VS, a region proposed to serve as a key motivational hub (Mogenson et al., 1980).

DA and the ACC Track Costs to Constrain Persistence

Persistence is justifiable only inasmuch as that progress outpaces accruing costs. A normative account appeals to the opportunity costs of WM allocation (Kurzban et al., 2013). Evidence of DA encoding opportunity costs comes from a high-resolution fMRI study finding signed RPE-like increases in activity in the VTA/substantia nigra (SN) corresponding with the value of unchosen options which therefore constituted missed opportunities (D'Ardenne et al., 2013). The ACC, by virtue of its connectivity with lateral PFC WM circuits (see Kouneiher et al., 2009), and sustained activity through control-demanding tasks (see Dosenbach et al., 2006), is well-positioned to track such opportunity costs.

Regardless of the nature of control costs, however, the ACC, which has long been implicated in avoidance learning (Shackman et al., 2011), has been proposed to mediate avoidance of control demands by attenuating DA-based value-learning signals (Botvinick, 2007). The most direct evidence comes from a recent pharmaco-genetic imaging study (Cavanagh et al., 2014). The paradigm was structured such that reward and punishment cues were accompanied by either high or low decision conflict (control demands), designed to test the prediction that conflict would attenuate reward and boost punishment learning. As expected, an EEG signature of ACC activity—mid-frontal theta power—increased reliably on conflict versus non-conflict trials. Critically, conflict strengthened individual difference correlations between mid-frontal theta power and the perceived punishment value of a given stimulus, while it attenuated individual difference correlations with the perceived reward value of a given stimulus. This result supports the hypothesis that the ACC both recruits control resources and also signals the cost of recruitment, thereby attenuating reward and amplifying punishment learning. Evidence implicating DA in mediating cost signaling included that dosing with cabergoline—a D2 selective agonist that acts on presynaptic D2 autoreceptors to inhibit burst firing to reward and exaggerate burst firing to punishments in the DS—had the effect of reducing reward responsiveness and boosting punishment responsiveness during learning.

Direct midbrain recordings support the hypothesis that DA neurons encode effort costs. For example, a subset (11%) of midbrain VTA neurons in monkeys performing an effortful, incentivized reaching task fired in proportion to reward magnitude discounted by effort demands (Pasquereau and Turner, 2013). Similarly, population firing rates of SN neurons in monkeys performing an effort-based decision-making task increased during higher reward trials and decreased with increasing effort requirements (Varazzani et al., 2015) (Figure 5). Interestingly, a stronger relationship between net expected value and SN firing rates also predicted a stronger relationship between net expected value and choice behavior. This correlation suggests that midbrain dopaminergic activity has the capacity to directly influence decision-making beyond mediating value learning, a point to which we will return later.

The ACC is thus a strong candidate for regulating persistence with control episodes, via the DA midbrain, by virtue of its capacity to track not only incremental progress toward a goal, but also opportunity costs, and thereby signal control costs. We propose that ACC regulates persistence by conveying the momentary balance of accruing progress less costs via phasic DA release from the midbrain. As we discuss in the next section, these DA projections have important effects on not just value learning,

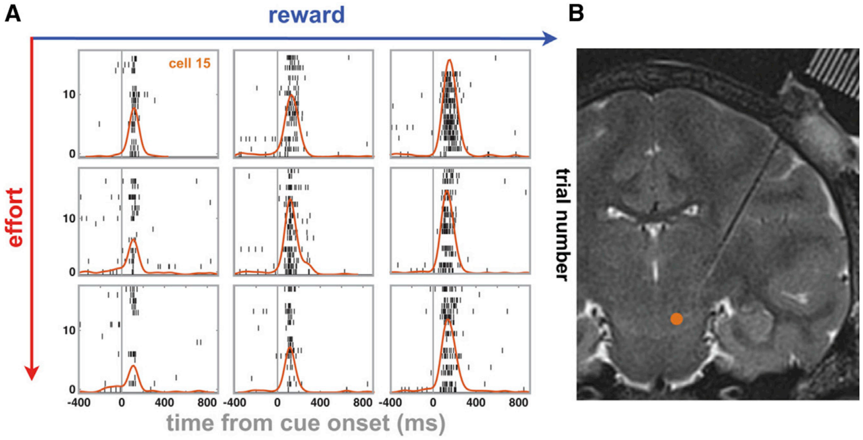

Figure 5. SN Firing Rates Encode Reward and Effort

(A) Raster plots of monkey SN cell activity to incentive cues during effort-based decision-making, firing intensifies with higher incentive values (liquid reward; rightward columns) and lower effort demands (handgrip squeeze; upper rows).

(B) Location of SN recordings (Varazzani et al., 2015).

but also on action selection, including instantaneous incentive motivation effects in the striatum.

DA and Action Selection Biasing

We propose that incentive-linked DA release promotes ACC-based action policies on engagement and persistence with control episodes over opposing action biases in the striatum (Figure 2D). The VS, and particularly the nucleus accumbens (NAcc), are regarded as a core limbic-motor interface (Mogenson et al., 1980), featuring dense reciprocal connections with both the dopaminergic midbrain and cortical regions including the ACC (Haber and Knutson, 2010). The DS, as described above, caches value functions controlling the gating of both motor behavior and WM allocation (O'Reilly and Frank, 2006). A reconceptualization of these regions, and their dopaminergic inputs, describes the VS as a "critic" evaluating states and driving DA RPE-based training of action value functions, while the DS serves as the "actor" that learns value functions for gating cognitive and motor action (Joel et al., 2002; van der Meer and Redish, 2011). Here, we highlight the role of DA in the VS in biasing action policies from cortical regions like the ACC, and DA in the DS in promoting gating of effortful cognitive actions as a function of incentive state and goal proximity.

DA RPEs Train the VS to Encode Net Incentive Value

In the VS, phasic DA RPEs train cortico-striatal synapses to reflect the net incentive value of a given state, i.e., expected reward less expected effort. Hence, fast-scan cyclic voltammetry in rats performing an effort-based decision-making task reveals NAcc DA release encoding both reward magnitude and lever-press ratio requirements of corresponding alternatives (Day et al., 2010) or the encoding of ratio requirements when demands are atypically low (Gan et al., 2009). Phasic DA RPE signals, in turn, train synapses to make VS neurons more excitable to states that signal relatively higher reward and lower effort costs.

Human fMRI studies support the hypothesis that the excitability of VS neurons encode net incentive value with respect to effort (see Croxson et al., 2009; Kurniawan et al., 2013; Schmidt et al., 2012). Striatal BOLD signal during a physical effort study increased to high versus low reward and was attenuated when it was preceded by high versus low demands for handgrip squeezes (Kurniawan et al., 2013). Similarly, in the cognitive domain, a transient VS response to reward receipt was diminished if it was preceded by high versus low demands for cognitive control (i.e., task-switching frequency) (Botvinick et al., 2009).

Importantly, the VS evaluates both model-based and model-free state features (Daw et al., 2011; van der Meer and Redish, 2011). The capacity for model-based evaluation makes the VS critical for selection of control episodes, which may involve multiple costly actions that are only justifiable with respect to ultimate goals. Hence, as we describe later, dopaminergic innervation of the VS is particularly important for selecting model-based behavior constituting control episodes.

Striatal DA Release Mediates Incentive Salience and State Motivation

Adaptive engagement with control episodes should involve not only rigid implementation of cached action values, but should also be sensitive to the current motivational state. In the arithmetic example, it would be adaptive to modulate persistence upon realizing that incentive point values were larger/smaller than first thought.

The *incentive salience* hypothesis holds that action values can be modulated instantaneously (i.e., without prior learning) by incentive cued striatal DA release (McClure et al., 2003; Phillips et al., 2008; Zhang et al., 2009). Hunger, e.g., increases instrumental lever pressing for food in rats (Phillips et al., 2008). A long-standing literature implicates striatal DA in modifying value functions and thereby promoting state willingness to expend effort (Bromberg-Martin et al., 2010; Kurniawan et al., 2011; Salamone and Correa, 2012). Alternatively, as we discuss later, incentive-cued DA release in the VS, in particular, may promote flexible approach, increasing apparent willingness to expend effort (McGinty et al., 2013; Nicola, 2010). In this section, we review evidence for DA's role in incentive state modulation of cached action values.

DA signaling in the VS appears necessary for the selection of physical effort. In a canonical paradigm, rats choose between climbing a high barrier for more reward or a low barrier for less reward, or alternatively, to select a high-ratio lever press option for more reward or a low ratio lever press option for less reward (see Bromberg-Martin et al., 2010; Kurniawan et al., 2011; Salamone and Correa, 2012 for reviews). The typical result is that DA blockade in the VS (along with antagonism in the ACC or lesions

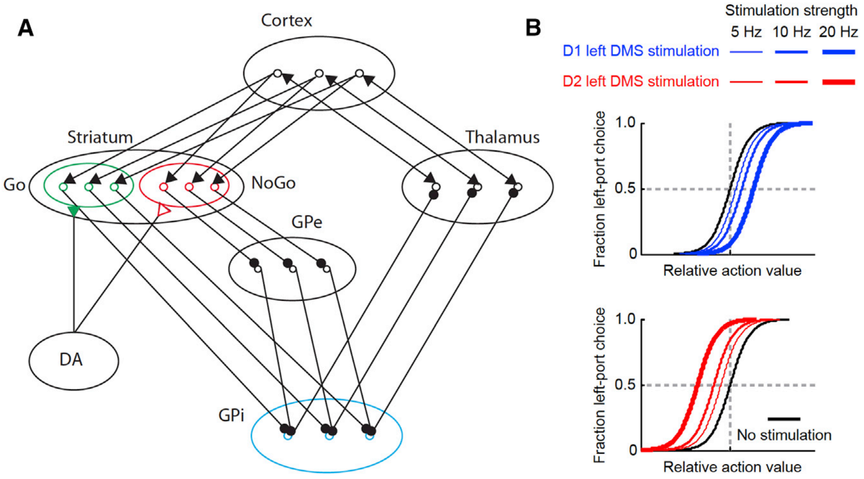

Figure 6. Striatal Dopamine Instantaneously Biases Action Selection

(A) Schematic of cortico-striatal neural network model modified in OpAL such that DA *instantaneously* promotes firing of D1-expressing, direct pathway Go cells (green region), and inhibits firing of D2-expressing, indirect pathway NoGo cells (red region) in the striatum, thereby promoting WM gating (Collins and Frank, 2014).

(B) Optogenetic stimulation of striatal D1 and D2 cells in rats, during decision-making in a two-alternative reward-learning task, produce dose-dependent shifts in preference toward the contralateral option in case of D1 cell stimulation (blue) and the ipsilateral option in case of D2 stimulation (red) mimicking shifts in subjective value functions (as reviewed in Lee et al., 2015).

of the ACC-VS loop) shifts preferences from high reward, high effort options toward low reward, low effort options.

Human studies also implicate striatal DA signaling in incentive motivation. For example, D2/D3 receptor and dopamine transporter density in the NAcc predicts trait-level achievement motivation in the individuals with ADHD (Volkow et al., 2011). In a combined fallypride-PET and d-amphetamine challenge study, human volunteers with the largest DS binding potential, and highest sensitivity to d-amphetamine during an instrumental button-pressing task were more willing to button press for reward (Treadway et al., 2012). Also, systemic DA agonism by the indirect agonist d-amphetamine ameliorates physical effort deficits among individuals with Parkinson's disease (Chong et al., 2015). In the cognitive domain, VS BOLD signal interacted with a genetic D2 receptor density marker to predict individual differences in WM performance (Nymberg et al., 2014).

The ability of incentive-cued DA release to energize behavior appears to critically depend on D2 receptor signaling in the NAcc core. In an instrumental lever-pressing task, transient GABAergic inactivation of the NAcc core, but not the shell, shifted preferences from high effort-high reward to low effort-low reward alternatives (Ghods-Sharifi and Floresco, 2010). Additionally, rats treated with a viral vector yielding acute overexpression of D2 receptors in the NAcc showed enhanced instrumental lever-pressing (Trifilieff et al., 2013). We note that studies using animal models with developmentally overexpressed D2 receptors have also shown the reverse effect, i.e., decreased incentive motivation (Krabbe et al., 2015; Ward et al., 2015). However, this reverse effect may be due to comorbid, developmental under-expression of NMDA NR1 and NR2B receptors on VTA neurons, which reduces both their firing frequency and burst firing (Krabbe et al., 2015).

DA Promotes Effortful Action by Promoting Cortical Action Policy Signals in the VS and Increasing the Likelihood of Gating in the DS

How does striatal DA bias selection of effortful action? By one proposal, action policies from the cortex, including canonical economic decision-making regions like the ACC and the OFC, are sent via axons that jointly synapse along with midbrain dopaminergic neurons in the striatum, and coincident phasic DA release boosts signal-to-noise: it enhances the contrast between strongly excited synapses corresponding to policy signals at the time of choice, relative to weakly excited synapses (Figure 2D) (Nicola et al., 2004). Thus, phasic DA efflux could instantaneously amplify cortical action policies projected to the VS (Roesch et al., 2009).

A recent computational proposal ("OPponent Actor Learning" or "OPAL"; Figure 6A) unifies value learning and incentive salience aspects of DA. In the DS, where gating policies are cached in terms of the relative strengths of cortico-striatal synapses onto D1-expressing Go and D2-expressing NoGo cells, DA should increase Go cell firing, and inhibit NoGo cells, thus modulating cached policies in favor of gating actions (Collins and Frank, 2014). According to this proposal, DA not only influences the learning of cached value functions, but can *instantaneously* modulate those value functions at the time of choice. The most direct evidence comes from an optogenetic study in which lateralized populations of DS Go and NoGo cells in rats were stimulated independently (Tai et al., 2012). Stimulation of Go cells yielded an apparent shift in preference to a contralateral option, while stimulation of NoGo cells yielded an apparent shift to an ipsilateral option. This shift mimicked additive effects in subjective value of one option over the other (Figure 6B). Hence, DS DA may translate incentive motivation into the selection of control episodes by increasing the subjective value of WM allocation.

DA in the VS Promotes Model-Based Behavior

The capacity of phasic DA in the VS to bias cortical action policies related to effortful action is most critical in the early stages of successful control episodes when behavior is necessarily model-based. For example, in the two-digit mental arithmetic problem, the actor must consider the point-valued outcome when deciding whether to engage and persist in the episode, since the immediate value of initial subgoals, e.g., computing the ones-digit product, is net-negative. Here, VS DA appears critical for promoting model-based behavior, of the kind necessary for persistence in these early stages. Indeed, higher presynaptic striatal DA, as measured by [F]DOPA PET, predicts greater

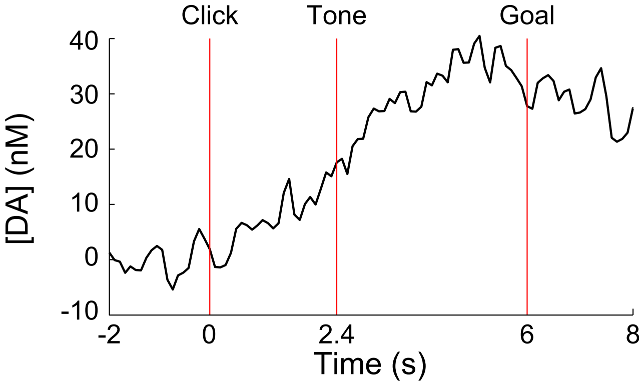

Figure 7. Ramping Dopamine Encodes Goal Progress

VS DA as measured by fast-scan cyclic voltammetry, during a single trial in which a rat progresses through a t-maze toward a final goal state.

The red vertical lines indicate, from left to right, the timing of an audible click cueing trial start, a tone indicating the direction the rat should turn, and finally successful goal attainment (chocolate milk). The ramp slopes reflected relative path distance, rather than absolute path distance or time-on-trial, and scaled with reward magnitude (Howe et al., 2013).

reliance on model-based decision-making in a two-stage sequential decision-making task and also predicted decreased reliance on habitual associations as encoded in striatal BOLD signal (Deserno et al., 2015). This could also explain why humans dosed with systemic DA agonists show more model-based relative to model-free decision-making (Wunderlich et al., 2012), especially to the extent that phasic signaling can be boosted by greater extrasynaptic tone (Dreyer et al., 2010).

The emphasis on promoting model-based behavior aligns with a reconceptualization, in which VS DA supports flexible approach or persisting in goal-directed (and therefore model-based) behavior (McGinty et al., 2013; Nicola, 2010), rather than overcoming instrumental costs per se. The flexible approach hypothesis states that during periods in which reward are not immediately available, agents are more likely to disengage and, because they can assume different positions with respect to operanda during such pauses, NAcc DA is needed to flexibly reapproach and engage.

By this account, much of the extant literature on NAcc DA promoting instrumental effort can be reinterpreted, wherein subtle task features allow more opportunities for disengagement in conditions for which effort demands are higher, placing more demands on NAcc DA to support flexible approach (Nicola, 2010). This could explain why NAcc DA depletion does not always affect effort-based decision-making about instrumental lever pressing in rats, e.g., when instrumental task design permits few opportunities for disengagement (Walton et al., 2009). It further explains the observation that NAcc DA is only necessary for initiating instrumental lever-pressing when there are longer pauses between action opportunities (Nicola, 2010).

Regardless of whether VS DA is necessary for overcoming effort costs or for flexible approach, the selection of effortful action sequences, like those comprising control episodes, requires VS DA (Nicola, 2007). Moreover, this may be particularly true when there is greater "psychological distance" from goals (Salamone and Correa, 2012), involving deeper action-outcome chaining, whether that distance is a function of space, time, or the number of subgoals in a two-digit multiplication.

DA Tone in the Striatum Reflects Goal Progress and Invigorates Action

As progress is made within a control episode, and percepts narrow in on a goal state, action invigoration becomes more important. Consider the final stage of a two-digit multiplication, when cortical representations of products for summation increasingly suggest the ultimate solution. At this stage, consideration of outcome incentives becomes less important than quick and robust execution of a retrieval action to finalize the solution. Intriguingly, recently discovered DA dynamics appear well-suited to subserve this functional shift from model-based control to invigoration (Figure 7). Namely, striatal DA tone ramps up smoothly, encoding goal progress (Howe et al., 2013). This dynamic, discovered with fast-scan cyclic voltammetry in rats navigating mazes, was found to scale with reward magnitude, and encode relative, rather than absolute distance to the goal.

The mechanism of DA ramping is not clear, whether it reflects local release or ramping firing of midbrain DA cells. Ramping may actually result from the progressive accumulation, or "spill-over" from phasic DA release (Gershman, 2014), e.g., as progress is made. In particular, phasic DA release may reflect the temporal derivative of a running average rate of progress as tracked by the ACC and OFC (O'Reilly et al., 2014) or pseudo-reward in hierarchical RL (Ribas-Fernandes et al., 2011).

An important functional consequence of rising striatal DA tone is the invigoration of behavior. Specifically, striatal DA tone is thought to encode the average rate of experienced reward and promote vigor (inverse latency to responding) adaptively, such that higher rates of reward imply a richer local resource that an actor should act more quickly to obtain (Niv et al., 2007). Hence, sufficiently fast progress toward the final goal yields ramping striatal DA tone, which can also promote action invigoration as the goal nears.

We close this section on dopaminergic mediation of value-learning and effort-based decision-making by noting conflicting evidence. First, a recent study has shown that DA-based cached values do not necessarily map onto preferred actions (Hollon et al., 2014). Specifically, fast-scan cyclic voltammetry in the rat NAcc revealed that DA tone was higher on trials in which the rat was forced to choose a dis-preferred high effort high reward option over a preferred low effort low reward option. Of course, DA may play different roles in forced- and free-choice decision-making, but this result suggests that, at least in some contexts, the rank-ordered relationship between DA and preference can be violated. Second, we note recent work aimed at developing a rodent model of cognitive effort-based decision-making (see Hosking et al., 2015). This work has provided mixed evidence so far regarding the consequences of systemic, pharmacological DA manipulation on willingness to expend cognitive (versus physical) effort. It is open for debate whether the new rodent model represents the sorts of cognitive effort-based decision-making that is of focal interest for control episodes and whether the task sufficiently discriminates effort-based from probabilistic decision-making. Nevertheless, a rodent model obviously holds great promise for more fine-grained investigation into the neural circuitry mediating decisions about cognitive effort.

DA Translates Incentives into Cognitive Motivation: Summary Proposal

Here, we recapitulate the proposal we have been building, whereby DA does double duty during costly control episodes. We define control episodes as temporally extended sequences in which WM is allocated to represent the rules needed to guide goal-directed behavior. During an episode, DA does double duty in that it: (1) influences WM contents by functional modulation of WM circuits, and (2) supports value-learning and decision-making about effortful cognitive actions (Figure 2).

Generally, we propose that:

- Phasic DA RPE signals encode goal benefits and effort costs for control episodes, caching net values in terms of LTP and LTD of cortico-striatal synapses.
- Incentive-linked DA release instantaneously augments cached values, increasing the likelihood of gating relevant task sets into WM in the striatum, thereby initiating control episodes associated with high incentive value.
- During control episodes, the ACC tracks both accruing opportunity costs and incremental progress, the balance of these is conveyed to midbrain DA neurons, where it is then transmitted to the striatum and PFC as phasic, effort-discounted, pseudo-RPE signals.
- In the PFC, rising DA tone encoding fast goal progress (or high incentive state) enhances the robustness of persistent activity, thereby stabilizing active maintenance in recurrent networks representing task goals.
- In the VS, DA release promotes drive (or flexible approach) to select extended sequences of goal-directed behavior. This is particularly critical at early stages of a control episode. As the goal state nears, ramping DA tone invigorates (potentiates) action gating, including WM allocation actions.
- In the DS, DA tone encoding sufficiently fast goal progress in a ramping fashion increases the general likelihood of task set updating. However, hierarchically structured task sets in the PFC interact with DS to target lower-level task sets for contextually appropriate out-gating. Thus, specific, lower-level flexibility is promoted while high-level goal maintenance is sustained during the control episode.
- Conversely, to the extent that opportunity costs outpace incremental progress, the likelihood of disengagement rises. This may result from falling PFC DA tone, reducing the stability of WM representations or reduced likelihood of WM gating in cortico-striatal-thalamic loops. Declining DA release undermines goal-directed flexible approach effects in the VS, further potentiating distraction.

Gaps in our account remain. We have described how rising PFC DA promotes task set stability, yet we have also pointed to evidence that supraoptimal PFC DA tone yields destabilization, and, indeed, how rapid PFC DA efflux could act as a global updating signal, indiscriminately destabilizing all current representations. However, we think that, for most operating regimes, PFC DA tone is unlikely to yield destabilization. As noted in a recent review (Spencer et al., 2015), intra-PFC injections of methylphenidate in rats, boosting DA tone, do not impair WM at concentrations that are 16- to 32-fold higher than clinically relevant methylphenidate doses. This stands in contrast to the observation that systemic administration of methylphenidate can impair WM at 4-fold concentrations higher than clinical doses.

An account of such discrepancies (between systemic and localized DA pharmacological manipulations) is that they relate to distinctions between DA modulation of PFC versus the striatum. Specifically, systemic high-dose DA manipulations may primarily act in the striatum where high DA tone can potentiate gating indiscriminately (Cools and D'Esposito, 2011). In a recent PET study demonstrating this effect in humans, the consequences of incentive motivation on performance of a Stroop task were investigated as a function of individual differences in baseline striatal DA synthesis capacity (using 6-[18F]fluoro-l-m-tyrosine uptake) (Aarts et al., 2014). The key finding was that while incentives enhanced performance for some participants, those with highest baseline synthesis capacity saw a decrement in incentivized performance. This pattern is consistent with the interpretation that incentive-cued striatal DA release for those with high baseline DA synthesis capacity yielded indiscriminant updating, undermining performance. In our proposal, striatal DA tone rises when progress outpaces opportunity costs; however, indiscriminate updating is typically prevented (under non-pharmacological conditions) by the imposition of hierarchical, targeted updating policies guided by PFC WM representations. Thus striatal DA tone interacts with targeted updating policies to maintain engagement with the current control episode.

In focusing on DA, we have neglected other potentially relevant neurotransmitter systems. Norepinephrine, for example, has similar effects on the stability of WM representations and also responds like DA to incentive cues (Sara, 2009). A recent study showed that while SN neurons encoded net cost-benefit during effort-based decision-making, locus coeruleus neurons encoded effort demands during task execution, suggesting a dissociation between DA in action selection and norepinephrine in action execution (Varazzani et al., 2015). Adenosine, for its part, appears to interact with the midbrain dopaminergic system to regulate effort-based decision-making and may account for the effects of caffeine on cognitive effort (Salamone et al., 2012). Serotonin has also been proposed to oppose DA learning effects and may subserve effort cost learning (Boureau and Dayan, 2010). Nevertheless, we think that DA in particular has a number of useful properties that position it best for mediating incentive motivation for cognitive effort.

Our proposal is similar in scope to other recent proposals. As described above, the Expected Value of Control proposal (Shenhav et al., 2013) considers cognitive control recruitment as driven by net expected value computations in ACC. A recent RL model (Holroyd and McClure, 2015) also considers the role of the ACC in value-based regulation of cognitive control and thus offers specific predictions about the influence of reward dynamics on effortful action including upregulation when reward rates are below average and downregulation when reward rates are above average. There are numerous points of theoretical overlap among our proposals. For example, in all three, the ACC biases the selection of effortful cognitive control actions via interactions with the striatum. Our proposal complements the other two accounts by articulating the varied and precise roles by which DA contributes to the value-based regulation of WM systems during control episodes. For example, in the computational model of Holroyd and McClure (2015), DA interacts with ACC signals such that when average reward, and presumably DA tone (Niv et al., 2007), are high, control signals are boosted because outcomes are more likely to fall below the average reward rate. Conversely, striatal DA blockade is posited to be computationally equivalent to low average reward, thus any reward is effectively above average and control signals dissipate. In our proposal, by contrast, striatal DA blockade also has the effect of diminishing effortful cognitive control, but it has its effects not in terms of a shift in perceived average reward, but in terms of diminished WM stability in the PFC and targeted flexibility via the striatum.

We acknowledge the tentative nature of our proposal. Computational modeling and experimental validation are required to ensure DA has the functional capacity to subserve adaptive engagement and persistence in the ways we hypothesize. Nevertheless, we hope this conceptual sketch unifies disparate literatures on DA's various functional properties and prompts development of a comprehensive theory of DA in cognitive effort. We have highlighted DA's roles in value-learning and effort-based decision-making and also the direct functional modulation of WM circuits and thereby WM contents by phasic and tonic DA modes. The integration of these two broad literatures together indicates double duty for DA in motivating cognitive effort.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institute of Mental Health (1F31 MH100855-01 to A.W. and 1R21 MH105800 to T.S.B.) and the National Institute of Aging (R01 AG043461 to T.S.B.).

REFERENCES

Aalto, S., Brück, A., Laine, M., Någren, K., and Rinne, J.O. (2005). Frontal and temporal dopamine release during working memory and attention tasks in healthy humans: a positron emission tomography study using the high-affinity dopamine D2 receptor ligand [11C]FLB 457. J. Neurosci. *25*, 2471–2477.

Aarts, E., Roelofs, A., Franke, B., Rijpkema, M., Fernández, G., Helmich, R.C., and Cools, R. (2010). Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence from genetic imaging. Neuropsychopharmacology *35*, 1943–1951.

Aarts, E., Wallace, D.L., Dang, L.C., Jagust, W.J., Cools, R., and D'Esposito, M. (2014). Dopamine and the cognitive downside of a promised bonus. Psychol. Sci. *25*, 1003–1009.

Alexander, W.H., and Brown, J.W. (2011). Medial prefrontal cortex as an action-outcome predictor. Nat. Neurosci. *14*, 1338–1344.

Badre, D., and Frank, M.J. (2012). Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. Cereb. Cortex *22*, 527–536.

Bahlmann, J., Aarts, E., and D'Esposito, M. (2015). Influence of motivation on control hierarchy in the human frontal cortex. J. Neurosci. *35*, 3207–3217.

Berridge, C.W., and Arnsten, A.F.T. (2013). Psychostimulants and motivated behavior: arousal and cognition. Neurosci. Biobehav. Rev. *37* (9 Pt A), 1976–1984.

Boehler, C.N., Hopf, J.M., Krebs, R.M., Stoppel, C.M., Schoenfeld, M.A., Heinze, H.J., and Noesselt, T. (2011). Task-load-dependent activation of dopaminergic midbrain areas in the absence of reward. J. Neurosci. *31*, 4955–4961.

Botvinick, M.M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. Cogn. Affect. Behav. Neurosci. *7*, 356–366.

Botvinick, M., and Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. Annu. Rev. Psychol. *66*, 83–113.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict monitoring and cognitive control. Psychol. Rev. *108*, 624–652.

Botvinick, M.M., Huffstetler, S., and McGuire, J.T. (2009). Effort discounting in human nucleus accumbens. Cogn. Affect. Behav. Neurosci. *36*, 74–97.

Boureau, Y.-L., and Dayan, P. (2010). Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology *9*, 16–27.

Braver, T.S. (2012). The variable nature of cognitive control: a dual mechanisms framework. Trends Cogn. Sci. *16*, 106–113.

Braver, T.S., Barch, D.M., and Cohen, J.D. (1999). Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. Biol. Psychiatry *46*, 312–328.

Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. Neuron *68*, 815–834.

Brown, J.W., and Braver, T.S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. Science *307*, 1118–1121.

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

Butts, K.A., Weinberg, J., Young, A.H., and Phillips, A.G. (2011). Glucocorticoid receptors in the prefrontal cortex regulate stress-evoked dopamine efflux and aspects of executive function. Proc. Natl. Acad. Sci. USA *108*, 18459–18464.

Cacioppo, J., Petty, R., Feinstein, J., and Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. Psychol. Bull. *119*, 197–253.

Cavanagh, J.F., Masters, S.E., Bath, K., and Frank, M.J. (2014). Conflict acts as an implicit cost in reinforcement learning. Nat. Commun. *5*, 5394.

Chatham, C.H., and Badre, D. (2013). Working memory management and predicted utility. Front. Behav. Neurosci. *7*, 83.

Chatham, C.H., and Badre, D. (2015). Multiple gates on working memory. Current Opinion in Behavioral Sciences *1*, 23–31.

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

Chiew, K.S., and Braver, T.S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. Cogn. Affect. Behav. Neurosci. *14*, 509–529.

Chong, T.T.J., Bonnelle, V., Manohar, S., Veromann, K.-R., Muhammed, K., Tofaris, G.K., Hu, M., and Husain, M. (2015). Dopamine enhances willingness to exert effort for reward in Parkinson's disease. Cortex *69*, 40–46.

Collins, A.G.E., and Frank, M.J. (2014). Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol. Rev. *121*, 337–366.

Cools, R., and D'Esposito, M. (2011). Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol. Psychiatry *69*, e113–e125.

Cowen, S.L., Davis, G.A., and Nitz, D.A. (2012). Anterior cingulate neurons in the rat map anticipated effort and reward to their associated action sequences. J. Neurophysiol. *107*, 2393–2407.

Croxson, P.L., Walton, M.E., O'Reilly, J.X., Behrens, T.E.J., and Rushworth, M.F.S. (2009). Effort-based cost-benefit valuation and the human brain. J. Neurosci. *29*, 4531–4541.

D'Ardenne, K., Eshel, N., Luka, J., Lenartowicz, A., Nystrom, L.E., and Cohen, J.D. (2012). Role of prefrontal cortex and the midbrain dopamine system in working memory updating. Proc. Natl. Acad. Sci. USA *109*, 19900–19909.

D'Ardenne, K., Lohrenz, T., Bartley, K.A., and Montague, P.R. (2013). Computational heterogeneity in the human mesencephalic dopamine system. Cogn. Affect. Behav. Neurosci. *13*, 747–756.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron *69*, 1204–1215.

Day, J.J., Jones, J.L., Wightman, R.M., and Carelli, R.M. (2010). Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. Biol. Psychiatry *68*, 306–309.

Dayan, P. (2012). How to set the switches on this thing. Curr. Opin. Neurobiol. *22*, 1068–1074.

Deserno, L., Huys, Q.J.M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A.A., Dolan, R.J., Heinz, A., and Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. Proc. Natl. Acad. Sci. USA *112*, 1595–1600.

Dixon, M.L., and Christoff, K. (2012). The decision to engage cognitive control is driven by expected reward-value: neural and behavioral evidence. PLoS ONE *7*, e51637.

Dolan, R.J., and Dayan, P. (2013). Goals and habits in the brain. Neuron *80*, 312–325.

Donoso, M., Collins, A.G., and Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. Science *344*, 1481–1486.

Dosenbach, N.U.F., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., and Petersen, S.E. (2006). A core system for the implementation of task sets. Neuron *50*, 799–812.

Dreisbach, G., and Goschke, T. (2004). How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility. J. Exp. Psychol. Learn. Mem. Cogn. *30*, 343–353.

Dreisbach, G., and Fischer, R. (2012). Conflicts as aversive signals. Brain Cogn. *78*, 94–98.

Dreyer, J.K., Herrik, K.F., Berg, R.W., and Hounsgaard, J.D. (2010). Influence of phasic and tonic dopamine release on receptor activation. J. Neurosci. *30*, 14273–14283.

Duncan, J. (2013). The structure of cognition: attentional episodes in mind and brain. Neuron *80*, 35–50.

Durstewitz, D., and Seamans, J.K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol. Psychiatry *64*, 739–749.

Etzel, J., Cole, M.W., Zacks, J.M., Kay, K.N., and Braver, T.S. (2015). Reward motivation enhances task coding in frontoparietal cortex. Cereb. Cortex. Published online January 19, 2015. http://dx.doi.org/10.1093/cercor/bhu327.

Frank, M.J., and Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb. Cortex *22*, 509–526.

Frank, M.J., Loughry, B., and O'Reilly, R.C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cogn. Affect. Behav. Neurosci. *1*, 137–160.

Gan, J.O., Walton, M.E., and Phillips, P.E.M. (2009). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nat. Neurosci. *13*, 25–27.

Gershman, S.J. (2014). Dopamine ramps are a consequence of reward prediction errors. Neural Comput. *26*, 467–471.

Ghods-Sharifi, S., and Floresco, S.B. (2010). Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. Behav. Neurosci. *124*, 179–191.

Gläscher, J., Hampton, A.N., and O'Doherty, J.P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb. Cortex *19*, 483–495.

Haber, S.N., and Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology *35*, 4–26.

Hammar, A., Strand, M., Årdal, G., Schmid, M., Lund, A., and Elliott, R. (2011). Testing the cognitive effort hypothesis of cognitive impairment in major depression. Nord. J. Psychiatry *65*, 74–80.

Herd, S.A., O'Reilly, R.C., Hazy, T.E., Chatham, C.H., Brant, A.M., and Friedman, N.P. (2014). A neural network model of individual differences in task switching abilities. Neuropsychologia *62*, 375–389.

Hess, T.M., and Ennis, G.E. (2012). Age differences in the effort and costs associated with cognitive activity. J. Gerontol. B Psychol. Sci. Soc. Sci. *67*, 447–455.

Hillman, K.L., and Bilkey, D.K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. Nat. Neurosci. *15*, 1290–1297.

Hiroyuki, N. (2014). Multiplexing signals in reinforcement learning with internal models and dopamine. Curr. Opin. Neurobiol. *25*, 123–129.

Hollon, N.G., Arnold, M.M., Gan, J.O., Walton, M.E., and Phillips, P.E.M. (2014). Dopamine-associated cached values are not sufficient as the basis for action selection. Proc. Natl. Acad. Sci. USA *111*, 18357–18362.

Holroyd, C.B., and Coles, M.G.H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. *109*, 679–709.

Holroyd, C.B., and Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. Trends Cogn. Sci. *16*, 122–128.

Holroyd, C.B., and McClure, S.M. (2015). Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. Psychol. Rev. *122*, 54–83.

Hosking, J.G., Floresco, S.B., and Winstanley, C.A. (2015). Dopamine antagonism decreases willingness to expend physical, but not cognitive, effort: a comparison of two rodent cost/benefit decision-making tasks. Neuropsychopharmacology *40*, 1005–1015.

Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575–579.

Inzlicht, M., Schmeichel, B.J., and Macrae, C.N. (2014). Why self-control seems (but may not be) limited. Trends Cogn. Sci. *18*, 127–133.

Jimura, K., Locke, H.S., and Braver, T.S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. Proc. Natl. Acad. Sci. USA *107*, 8871–8876.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Netw. *15*, 535–547.

Kahnt, T., Weber, S.C., Haker, H., Robbins, T.W., and Tobler, P.N. (2015). Dopamine D2-receptor blockade enhances decoding of prefrontal signals in humans. J. Neurosci. *35*, 4104–4111.

Kayser, A.S., Mitchell, J.M., Weinstein, D., and Frank, M.J. (2014). Dopamine, locus of control, and the exploration-exploitation tradeoff. Neuropsychopharmacology *40*, 454–462.

Kennerley, S.W., Dahmubed, A.F., Lara, A.H., and Wallis, J.D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. J. Cogn. Neurosci. *21*, 1162–1178.

Kennerley, S.W., Behrens, T.E.J., and Wallis, J.D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat. Neurosci. *14*, 1581–1589.

Kodama, T., Hikosaka, K., Honda, Y., Kojima, T., and Watanabe, M. (2014). Higher dopamine release induced by less rather than more preferred reward during a working memory task in the primate prefrontal cortex. Behav. Brain Res. *266*, 104–107.

Kool, W., McGuire, J.T., Rosen, Z.B., and Botvinick, M.M. (2010). Decision making and the avoidance of cognitive demand. J. Exp. Psychol. Gen. *139*, 665–682.

Kouneiher, F., Charron, S., and Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. Nat. Neurosci. *12*, 939–945.

Krabbe, S., Duda, J., Schiemann, J., Poetschke, C., Schneider, G., Kandel, E.R., Liss, B., Roeper, J., and Simpson, E.H. (2015). Increased dopamine D2 receptor activity in the striatum alters the firing pattern of dopamine neurons in the ventral tegmental area. Proc. Natl. Acad. Sci. USA *112*, E1498–E1506.

Krebs, R.M., Boehler, C.N., Roberts, K.C., Song, A.W., and Woldorff, M.G. (2012). The involvement of the dopaminergic midbrain and cortico-striatal-thalamic circuits in the integration of reward prospect and attentional task demands. Cereb. Cortex *22*, 607–615.

Kroemer, N.B., Guevara, A., Ciocanea Teodorescu, I., Wuttig, F., Kobiella, A., and Smolka, M.N. (2014). Balancing reward and work: anticipatory brain activation in NAcc and VTA predict effort differentially. Neuroimage *102*, 510–519.

Kurniawan, I.T., Seymour, B., Talmi, D., Yoshida, W., Chater, N., and Dolan, R.J. (2010). Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action. J. Neurophysiol. *104*, 313–321.

Kurniawan, I.T., Guitart-Masip, M., and Dolan, R.J. (2011). Dopamine and effort-based decision making. Front. Neurosci. *5*, 81.

Kurniawan, I.T., Guitart-Masip, M., Dayan, P., and Dolan, R.J. (2013). Effort and valuation in the brain: the effects of anticipation and execution. J. Neurosci. *33*, 6160–6169.

Kurzban, R., Duckworth, A., Kable, J.W., and Myers, J. (2013). An opportunity cost model of subjective effort and task performance. Behav. Brain Sci. *36*, 661–679.

Lee, A.M., Tai, L.H., Zador, A., and Wilbrecht, L. (2015). Between the primate and 'reptilian' brain: Rodent models demonstrate the role of corticostriatal circuits in decision making. Neuroscience *296*, 66–74.

Li, J., and Daw, N.D. (2011). Signals in human striatum are appropriate for policy update rather than value prediction. J. Neurosci. *31*, 5504–5511.

Locke, H.S., and Braver, T.S. (2008). Motivational influences on cognitive control: behavior, brain activation, and individual differences. Cogn. Affect. Behav. Neurosci. *8*, 99–112.

Ma, L., Hyman, J.M., Phillips, A.G., and Seamans, J.K. (2014). Tracking progress toward a goal in corticostriatal ensembles. J. Neurosci. *34*, 2244–2253.

Massar, S.A.A., Libedinsky, C., Weiyan, C., Huettel, S.A., and Chee, M.W.L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. Neuroimage *120*, 104–113.

Matsumoto, M., and Takada, M. (2013). Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. Neuron *79*, 1011–1024.

McClure, S.M., Daw, N.D., and Montague, P.R. (2003). A computational substrate for incentive salience. Trends Neurosci. *26*, 423–428.

McGinty, V.B., Lardeux, S., Taha, S.A., Kim, J.J., and Nicola, S.M. (2013). Invigoration of reward seeking by cue and proximity encoding in the nucleus accumbens. Neuron *78*, 910–922.

McGuire, J.T., and Botvinick, M.M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. Proc. Natl. Acad. Sci. USA *107*, 7922–7926.

Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J., and Pessiglione, M. (2013). Neurocomputational account of how the human brain decides when to have a break. Proc. Natl. Acad. Sci. USA *110*, 2641–2646.

Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. *24*, 167–202.

Mogenson, G.J., Jones, D.L., and Yim, C.Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. Prog. Neurobiol. *14*, 69–97.

Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. *16*, 1936–1947.

Nicola, S.M. (2007). The nucleus accumbens as part of a basal ganglia action selection circuit. Psychopharmacology (Berl.) *191*, 521–550.

Nicola, S.M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. J. Neurosci. *30*, 16585–16600.

Nicola, S.M., Hopf, F.W., and Hjelmstad, G.O. (2004). Contrast enhancement: a physiological effect of striatal dopamine? Cell Tissue Res. *318*, 93–106.

Niv, Y. (2009). Reinforcement learning in the brain. J. Math. Psychol. *53*, 139–154.

Niv, Y., Daw, N.D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl.) *191*, 507–520.

Nymberg, C., Banaschewski, T., Bokde, A.L., Büchel, C., Conrod, P., Flor, H., Frouin, V., Garavan, H., Gowland, P., Heinz, A., et al.; IMAGEN consortium (2014). DRD2/ANKK1 polymorphism modulates the effect of ventral striatal activation on working memory performance. Neuropsychopharmacology *39*, 2357–2365.

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

O'Reilly, R.C., Hazy, T.E., Mollick, J., Mackie, P., and Herd, S. (2014). Goal-driven cognition in the brain: a computational framework. arXiv, arXiv:1404.7591, http://www.arxiv.org/abs/1404.7591.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. Annu. Rev. Neurosci. *34*, 333–359.

Parvizi, J., Rangarajan, V., Shirer, W.R., Desai, N., and Greicius, M.D. (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. Neuron *80*, 1359–1367.

Pasquereau, B., and Turner, R.S. (2013). Limited encoding of effort by dopamine neurons in a cost-benefit trade-off task. J. Neurosci. *33*, 8288–8300.

Pessoa, L., and Engelmann, J.B. (2010). Embedding reward signals into perception and cognition. Front. Neurosci. *4*, 4.

Phillips, A.G., Vacca, G., and Ahn, S. (2008). A top-down perspective on dopamine, motivation and memory. Pharmacol. Biochem. Behav. *90*, 236–249.

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., and Dreher, J.-C. (2010). Separate valuation subsystems for delay and effort decision costs. J. Neurosci. *30*, 14080–14090.

Raichle, M.E., and Mintun, M.A. (2006). Brain work and brain imaging. Annu. Rev. Neurosci. *29*, 449–476.

Ribas-Fernandes, J.J.F., Solway, A., Diuk, C., McGuire, J.T., Barto, A.G., Niv, Y., and Botvinick, M.M. (2011). A neural signature of hierarchical reinforcement learning. Neuron *71*, 370–379.

Riggall, A.C., and Postle, B.R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. J. Neurosci. *32*, 12990–12998.

Robbins, T.W., and Arnsten, A.F. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. Annu. Rev. Neurosci. *32*, 267–287.

Roesch, M.R., Singh, T., Brown, P.L., Mullins, S.E., and Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. J. Neurosci. *29*, 13365–13376.

Salamone, J.D., and Correa, M. (2012). The mysterious motivational functions of mesolimbic dopamine. Neuron *76*, 470–485.

Salamone, J.D., Correa, M., Nunes, E.J., Randall, P.A., and Pardo, M. (2012). The behavioral pharmacology of effort-related choice behavior: dopamine, adenosine and beyond. J. Exp. Anal. Behav. *97*, 125–146.

Samanez-Larkin, G.R., Buckholtz, J.W., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., Arrington, C.M., Baldwin, R.M., Smith, C.E., Treadway, M.T., et al. (2013). A thalamocorticostriatal dopamine network for psychostimulant-enhanced human cognitive flexibility. Biol. Psychiatry *74*, 99–105.

Sara, S.J. (2009). The locus coeruleus and noradrenergic modulation of cognition. Nat. Rev. Neurosci. *10*, 211–223.

Satterthwaite, T.D., Ruparel, K., Loughead, J., Elliott, M.A., Gerraty, R.T., Calkins, M.E., Hakonarson, H., Gur, R.C., Gur, R.E., and Wolf, D.H. (2012). Being right is its own reward: load and performance related ventral striatum activation to correct responses during a working memory task in youth. Neuroimage *61*, 723–729.

Saunders, B., and Inzlicht, M. (2015). Vigour and fatigue: How variation in affect underlies effective self-control. In Motivation and Cognitive Control, T.S. Braver, ed. (Taylor & Francis/Routledge), pp. 211–234.

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., and Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. PLoS Biol. *10*, e1001266.

Schouppe, N., De Houwer, J., Ridderinkhof, K.R., and Notebaert, W. (2012). Conflict: run! Reduced Stroop interference with avoidance responses. Q. J. Exp. Psychol. (Hove) *65*, 1052–1058.

Schouppe, N., Demanet, J., Boehler, C.N., Ridderinkhof, K.R., and Notebaert, W. (2014). The role of the striatum in effort-based decision-making in the absence of reward. J. Neurosci. *34*, 2148–2154.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593–1599.

Schweimer, J., and Hauber, W. (2006). Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision making. Learn. Mem. *13*, 777–782.

Schweimer, J., Saft, S., and Hauber, W. (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. Behav. Neurosci. *119*, 1687–1692.

Seamans, J.K., and Yang, C.R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog. Neurobiol. *74*, 1–58.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., and Davidson, R.J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat. Rev. Neurosci. *12*, 154–167.

Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron *79*, 217–240.

Shiner, T., Symmonds, M., Guitart-Masip, M., Fleming, S.M., Friston, K.J., and Dolan, R.J. (2015). Dopamine, salience, and response set shifting in prefrontal cortex. Cereb. Cortex *25*, 3629–3639.

Skvortsova, V., Palminteri, S., and Pessiglione, M. (2014). Learning to minimize efforts versus maximizing rewards: computational principles and neural correlates. J. Neurosci. *34*, 15621–15630.

Small, D.M., Gitelman, D., Simmons, K., Bloise, S.M., Parrish, T., and Mesulam, M.M. (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. Cereb. Cortex *15*, 1855–1865.

Spencer, R.C., Devilbiss, D.M., and Berridge, C.W. (2015). The cognition-enhancing effects of psychostimulants involve direct action in the prefrontal cortex. Biol. Psychiatry *77*, 940–950.

Spunt, R.P., Lieberman, M.D., Cohen, J.R., and Eisenberger, N.I. (2012). The phenomenology of error processing: the dorsal ACC response to stop-signal errors tracks reports of negative affect. J. Cogn. Neurosci. *24*, 1753–1765.

Steullet, P., Cabungcal, J.-H., Cuénod, M., and Do, K.Q. (2014). Fast oscillatory activity in the anterior cingulate cortex: dopaminergic modulation and effect of perineuronal net loss. Front. Cell. Neurosci. *8*, 244.

Strauss, G.P., Morra, L.F., Sullivan, S.K., and Gold, J.M. (2015). The role of low cognitive effort and negative symptoms in neuropsychological impairment in schizophrenia. Neuropsychology *29*, 282–291.

Tai, L.-H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nat. Neurosci. *15*, 1281–1289.

Treadway, M.T., Buckholtz, J.W., Cowan, R.L., Woodward, N.D., Li, R., Ansari, M.S., Baldwin, R.M., Schwartzman, A.N., Kessler, R.M., and Zald, D.H. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. J. Neurosci. *32*, 6170–6176.

Trifilieff, P., Feng, B., Urizar, E., Winiger, V., Ward, R.D., Taylor, K.M., Martinez, D., Moore, H., Balsam, P.D., Simpson, E.H., and Javitch, J.A. (2013). Increasing dopamine D2 receptor expression in the adult nucleus accumbens enhances motivation. Mol. Psychiatry *18*, 1025–1033.

van der Meer, M.A., and Redish, A.D. (2011). Ventral striatum: a critical look at models of learning and evaluation. Curr. Opin. Neurobiol. *21*, 387–392.

van Holstein, M., Aarts, E., van der Schaaf, M.E., Geurts, D.E.M., Verkes, R.J., Franke, B., van Schouwenburg, M.R., and Cools, R. (2011). Human cognitive flexibility depends on dopamine D2 receptor signaling. Psychopharmacology (Berl.) *218*, 567–578.

van Schouwenburg, M., Aarts, E., and Cools, R. (2010). Dopaminergic modulation of cognitive control: distinct roles for the prefrontal cortex and the basal ganglia. Curr. Pharm. Des. *16*, 2026–2032.

Varazzani, C., San-Galli, A., Gilardeau, S., and Bouret, S. (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. J. Neurosci. *35*, 7866–7877.

Vassena, E., Silvetti, M., Boehler, C.N., Achten, E., Fias, W., and Verguts, T. (2014). Overlapping neural systems represent cognitive effort and reward anticipation. PLoS ONE *9*, e91008.

Vijayraghavan, S., Wang, M., Birnbaum, S.G., Williams, G.V., and Arnsten, A.F.T. (2007). Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. Nat. Neurosci. *10*, 376–384.

Volkow, N.D., Wang, G.-J., Newcorn, J.H., Kollins, S.H., Wigal, T.L., Telang, F., Fowler, J.S., Goldstein, R.Z., Klein, N., Logan, J., et al. (2011). Motivation deficit in ADHD is associated with dysfunction of the dopamine reward pathway. Mol. Psychiatry *16*, 1147–1154.

von Stumm, S., Hell, B., and Chamorro-Premuzic, T. (2011). The hungry mind: intellectual curiosity is the third pillar of academic performance. Perspect. Psychol. Sci. *6*, 574–588.

Walton, M.E., Groves, J., Jennings, K.A., Croxson, P.L., Sharp, T., Rushworth, M.F.S., and Bannerman, D.M. (2009). Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision making. Eur. J. Neurosci. *29*, 1678–1691.

Ward, R.D., Winiger, V., Higa, K.K., Kahn, J.B., Kandel, E.R., Balsam, P.D., and Simpson, E.H. (2015). The impact of motivation on cognitive performance in an animal model of the negative and cognitive symptoms of schizophrenia. Behav. Neurosci. *129*, 292–299.

Westbrook, A., Kester, D., and 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.

Wickens, J.R., Horvitz, J.C., Costa, R.M., and Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. J. Neurosci. *27*, 8181–8183.

Wunderlich, K., Smittenaar, P., and Dolan, R.J. (2012). Dopamine enhances model-based over model-free choice behavior. Neuron *75*, 418–424.

Zhang, J., Berridge, K.C., Tindell, A.J., Smith, K.S., and Aldridge, J.W. (2009). A neural computational model of incentive salience. PLoS Comput. Biol. *5*, e1000437.