



Published in final edited form as:

Trends Cogn Sci. 2011 April ; 15(4): 143–151. doi:10.1016/j.tics.2011.02.002.

Posterior Cingulate Cortex: Adapting Behavior to a Changing World

John M. Pearson^{1,2}, Sarah R. Heilbronner^{1,2}, David L. Barack^{2,3}, Benjamin Y. Hayden^{1,2}, and Michael L. Platt^{1,2,4}

¹ Department of Neurobiology, Duke University Medical Center, Box 3209, Durham, NC 27710

² Center for Cognitive Neuroscience, Duke University, B203 Levine Science Research Center, Box 90999, Durham, NC 27708

³ Department of Philosophy, Duke University, Box 90743, Durham, NC 27708

⁴ Department of Evolutionary Anthropology, Duke University, 104 Biological Sciences Building, Box 90383, Durham, NC 27708-0680

Abstract

When has the world changed enough to warrant a new approach? The answer depends upon current needs, behavioral flexibility, and prior knowledge about the environment. Formal approaches solve the problem by integrating the recent history of rewards, errors, uncertainty, and context via Bayesian inference to detect changes in the world and alter behavioral policy. Neuronal activity in posterior cingulate cortex (CGp)—a key node in the default network—is known to vary with learning, memory, reward, and task engagement. We propose that these modulations reflect the underlying process of change detection and motivate subsequent shifts in behavior.

Learning in a changing world

Most days, you drive home along a familiar route. But today, something unexpected happens. The city has opened a new street, offering the possibility of a shortcut. Another day, a new intersection sends you down a road you didn't intend. Often a traffic jam severely alters the time it takes you to reach your destination. Whether you are a human driving home, a monkey foraging for food, or a rat navigating a maze, unexpected changes in the world necessitate a shift in behavioral policy—rules that guide decisions based on prior knowledge—and potentially promote learning. Changes force agents to engage learning systems, switch mental states, and shift attention, among other adjustments, and recent work has examined their physiological substrates [1,2]. Yet the loci of change detection within the brain remain unidentified. Here we propose that the posterior cingulate cortex (CGp) plays a key role in altering behavior in response to unexpected change. It may be that this region, which consumes more energy than any other cortical area [3], must do so just to keep pace with a dynamically changing world.

Corresponding author: Platt, M.L., platt@neuro.duke.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Learning, Change Detection, and Policy Switching

Since the discovery that dopaminergic neurons respond to rewarding events by signaling the difference between expected and received rewards—the so-called “reward prediction error” (RPE)—theories of reinforcement learning (RL) have come to dominate discussions of learning and conditioning [4,5]. In typical RL algorithms, learning is incremental and only slowly converges on stable behavior [5]. In an environment rapidly alternating among several fixed but distinct reward contingencies, crude RL agents might find themselves forever playing catch-up, unable to do more than gradually adjust in response to abrupt transitions. At present, we know little about the neural processes underlying rapid change detection, both those that switch between and those that initiate the learning of behavioral policies. Such processes should accommodate changes in the contingency structure of the environment (state space, Markov structure) and in its distribution of returns (volatility, outcomes, outliers). In the driving example above, the new or diverted streets correspond to changes in environmental structure, while the traffic jam corresponds to a change in the distribution of returns.

Standard reinforcement learning fails to capture the ability of animals to implement a wide array of strategies, each learned independently, with minimal switching costs. For example, drivers stuck in unexpected traffic do not gradually modify the typical route home, but re-select from among multiple alternatives based on expected delays, time of day, etc. While several classical theories of conditioning [6] posit surprise (formalized as the absolute value of RPE) or similar violations of expectation to dynamically adjust behavior and learning rate, such models are still based on the idea of a single policy subject to gradual updates. By contrast, a change detection framework allows for *multiple* behavioral policies by using statistical inference to distinguish expected variation from underlying environmental shifts [7–11], allowing only the behaviorally relevant policy to be deployed and learned. As a result, agents capable of change detection, rather than forever reshaping a single policy in the face of a dynamic environment, can instead adapt rapidly by switching between behavioral policies.

This idea accords well with the rapid adjustment of behavior to sudden changes in reward contingencies [12,13], and operates similarly to theories of conditioning that invoke Bayesian mechanisms to dynamically adjust learning rates [7,14](Box 2). In such scenarios, outcome tracking not only adjusts the current strategy (and learning rate), but also determines whether or not environmental change warrants a switch to a different strategy entirely. Reinforcement learning then operates as a sub-process within the change detection system, utilizing Bayesian inference along with a suite of inborn or learned models of the world [7,11,15].

Box 2

Bayesian Learning

Recent studies have provided evidence that both humans and non-human animals often employ sophisticated, model-based assumptions when learning about their environments [7,11,15]. That is, agents first determine an appropriate set of constructs by which to model the world, and then update the parameters of these models via Bayesian inference.

The Bayesian approach to learning proves particularly useful when modeling behavior in environments where the underlying parameters are subject to change. The underlying link between stimuli and outcomes may either change gradually in time [7,10] or shift suddenly [11–13,74]. Because such models include full predictive distributions for

observed quantities, they are able to incorporate both prediction error and surprise, thereby uniting aspects of both reinforcement learning and attentional learning [6].

While we remain agnostic regarding specific implementations of Bayesian learning, several key features are common to all models [7]: First, updates to the means of model parameters are proportional to the prediction error—the difference between observed and expected outcomes. Second, these updates are scaled by the learning rate, which depends on uncertainty—the width of the distribution of parameter estimates (Figure I). This uncertainty itself depends upon the current estimate of environmental rate of change, allowing agents to learn more slowly in stable worlds and incorporate new information more readily in frequently changing contexts. Finally, agents are expected to incorporate prior information through Bayes' rule, forcing them to fall back on model-based assumptions in the presence of limited information or recent change.

Yet full Bayesian approaches require intensive nonlinear computation, highlighting the need for simple, online update rules that approximate the full solution. Though models such as probabilistic population codes solve the Bayesian inference problem via neurons tuned to specific outcome variables [75], in regions without such well-characterized tuning curves, memory and processing constraints argue for trial-to-trial inference. In many cases, techniques such as the Kalman filter offer optimal solutions [8], while in others maximum likelihood estimates serve as useful approximations [11,74]. We believe that the brain uses a multitude of such approaches, which can be matched to environmental dynamics to produce “good enough” Bayesian behavior.

Past hypotheses regarding Posterior Cingulate Cortex function

Despite extensive interconnections with memory, attentional, and decision areas (see Box 1), the primary function of the posterior cingulate cortex (CGp) remains mysterious. Indeed, no commonly recognized unified theory for its role exists (Table 1). Clinical evidence demonstrates that early hypometabolism and neural degeneration in CGp predict cognitive decline in Alzheimer's disease [16,17], and CGp hyperactivity predicts cognitive dysfunction in schizophrenia [18]. While a host of experiments have identified the anterior cingulate cortex (ACC) as critical for processing feedback from individual choices and subsequent alterations in behavior [9,19–27], the functional role of CGp remains relatively obscure.

Box 1

Posterior Cingulate Cortex Anatomy & Physiology

Cingulate cortex as a whole has long been recognized as an important site integrating sensory, motor, visceral, motivational, emotional, and mnemonic information [73]. Posterior cingulate (CGp) is the portion of cingulate cortex caudal to the central sulcus. Though poorly understood, this brain structure nevertheless consumes large amounts of energy, making it one of the most metabolically expensive regions of the brain [3]. CGp is reciprocally connected with areas involved in attention—areas 7a, LIP, and 7, or PGM—as well as with brain areas involved in learning and motivation, including the anterior and lateral thalamic nuclei, the caudate nucleus, orbitofrontal cortex, and anterior cingulate cortex (ACC) (Figure I). CGp also forms strong, reciprocal connections with the medial temporal lobe, especially the parahippocampal gyrus, long known to be crucial for associative learning and episodic memory (for a thorough review of neuroanatomical connections, see [73]).

To date, disparate evidence for CGp involvement in cognitive and behavioral processes has thwarted any simple functional characterization. Anatomical connections to medial temporal lobe areas necessary for learning and memory, as well as to neocortical areas responsible for movement planning, suggest that CGp links these areas in order to store spatial and temporal information about the consequences of action [76,77]. The involvement of CGp in learning and memory, however, seems to extend far beyond the scope of movements and spatial orientation. Neuroimaging experiments have shown that CGp is activated by emotional stimuli, particularly when they have personal significance [28,29]. Neurons in CGp signal reward size, and some studies have suggested that they encode the subjective value of a chosen option [30–32]. The bulk of recent CGp studies, however, have focused on its role in the so-called “default mode network”, a set of interconnected brain regions showing elevated BOLD signal at rest and suppression during active task engagement [3].

Yet none of these proposed functional roles—mnemonic, attentional, spatial, default—successfully captures the full range of phenomena shown to modulate activity in CGp. Taking a broader view based on both electrophysiological and functional imaging evidence (summarized below), we conjecture that many of these observed modulations reflect the contribution of CGp to signaling environmental change and, when necessary, relevant shifts in behavioral policy. In our scheme, suppressed CGp activity favors operation within the current cognitive set, whereas increased activity reflects a change in either environmental structure or internal state and promotes flexibility, exploration, and renewed learning. This proposed role for CGp is consistent with converging evidence from imaging studies that default network regions play an active role in cognition when tasks require mental simulation and strategic planning [33,34], and lesion studies demonstrating deficits in multi-tasking and set-shifting [35,36]. While this hypothesis pertains primarily to the contributions of CGp to learning and decision-making, we also provide evidence to suggest that such an approach offers clues to the broader role of the default network in cognition.

Change Detection and Policy Control in Posterior Cingulate Cortex

CGp neurons not only respond in graded fashion to the magnitude of liquid reward associated with orienting, but also respond to the unpredicted omission of these same rewards [37]. In a choice task with one “safe” option delivering a fixed liquid reward and one “risky” option randomly delivering larger and smaller rewards with equal probability [32,38], CGp neurons encoded not only reward size but also reward variance [32]. Thus CGp neurons track policy-relevant variables like reward value as well as estimates of variability within the current environment.

In this probabilistically rewarded choice task, monkeys’ behavior followed a win-stay lose-shift (WSLS) heuristic [38]. After choosing the risky option, monkeys were more likely to choose it again if they received a larger reward but more likely to switch to the safe option if they received the smaller reward. Firing rates of CGp neurons were correspondingly higher following smaller rewards than large rewards, and variability in responses predicted the likelihood the monkey would switch his choice on the next trial. Brief microstimulation in CGp increased the likelihood monkeys would switch to the safe option after receiving a large reward—as if they had erroneously detected a bad outcome [38]. Just as reward outcomes occurring several trials in the past weakly influenced monkeys’ choices, tonic firing rates in CGp maintained information about previous rewards for several trials, and these modulations predicted future choices (Figure 1) [32,38]. Taken together, these findings suggest CGp neurons encode environmental outcomes (rewards and variance), maintain this information online (in a leaky fashion), and contribute to adjusting subsequent behavior.

A follow-up study probed whether representation of the need to switch applied beyond simple option switching to policy changes more generally [39]. Monkeys performed a variant of the k-armed bandit task in which reward amounts for four targets varied independently on each trial and slowly changed over time (Figure 2A–B)[8]. Behavior was characterized as following two distinct policies—explore and exploit—which depended on the recent history of rewards experienced for choosing each target [8]. Firing rates of CGp neurons not only signaled single-trial reward outcomes, but also predicted the probability of shifting between policies in graded fashion (Figure 2C–E). These observations endorse the idea that CGp participates in a circuit that monitors environmental outcomes for purposes of change detection and subsequent policy switching [39]. Thus, even in a more complex environment where changes in returns are gradual rather than sudden, the activity of CGp neurons both tracks and maintains strategically relevant information used to implement a change in behavioral policy.

A Model of Change Detection and Policy Switching in Cingulate Cortex

In order for agents to learn and switch between distinct strategies in response to environmental change, they must detect the change through inference over behavioral outcomes and adjust parameters within a given strategy. Figure 3 depicts a schematic of one such change detection and learning process.

In the model, outcome data from single events are passed to the change detection system, which recombines these variables into strategy-specific measures of Bayesian evidence that the environment has changed. As in other models of information accumulation [40,41], this signal, representing the log posterior odds favoring a given hypothesis (in this case, environmental change) increases until reaching a threshold, after which a “change” signal is broadcast, learning rates increase, and the agent switches strategy. In contrast with standard models of sensory evidence accumulation, these decision signals are maintained across multiple outcomes, and possess only a single threshold, as is appropriate for an all-or-none switching process that allows full Bayesian inference to be reduced to a simple update model [11]. Equally important, the decision variables accumulated by the change detector may vary between strategies, depending upon the expected distribution of outcomes from the environment. That is, the correct statistical test for agents to perform in change detection depends not only on the environment, but also on current and alternative strategies. Thus in an environment where the appropriate strategy depends heavily on the relative frequency of outliers, the correct tracking statistic may be neither the mean nor the variance of outcomes, but a simple proportion of occurrence above a threshold. In such a framework, the best statistic is the one that maximizes the area under the receiver operating characteristic (ROC) curve for the strategy switching problem, balancing false positives against false negatives in accord with the cost-benefit analysis that obtains in the current environment.

Physiologically, we tentatively identify areas such as ACC, amygdala, and basal ganglia as encoding individual event-related outcomes necessary for altering behavior within a given strategy. These variables function as inputs to both the gradual within-strategy (RL) learning system and the change-detection system, the latter of which may include key default network areas like CGp and the ventromedial prefrontal cortex (vmPFC). These identifications are suggested by the demonstrated sufficiency of midbrain dopaminergic signals for classical conditioning [42], as well as the role played by cortical targets of these signals in associative learning. For instance, the orbitofrontal (OFC) and anterior cingulate cortices (ACC), which maintain strong reciprocal connections with the basal ganglia [43,44], are necessary for representing links between outcomes, actions, and predictive cues [26,45–48] as well as facilitating changes in action [19,24]. Furthermore, we hypothesize that individual policies and policy selection are most likely computed in dorsolateral

prefrontal cortex (DLPFC), already implicated in strategic decision-making and action planning [49,50], while long-term associative learning is implemented via the hippocampus and surrounding structures. Clearly, however, much work remains to validate both the model and the identification of its individual functions with specific anatomical areas.

Evidence for such a process is depicted in Figure 2F–G. There, we present a suggestive reanalysis of behavioral and neural data from the above-mentioned bandit task [39] using an expanded Kalman filter model (see Online Supplementary Material). This Kalman filter improved our behavioral fits in all cases, and firing in CGp neurons significantly tracked variability in the chosen targets (equivalent to uncertainty; $n=8/83, 9/83$), reward prediction error (RPE; $n=15/83, 17/83$) and learning rate ($n=10/83, 10/83$). These correlations (variability: $R=0.10$ chosen option, RPE: $R=0.10$, learning rate: $R=0.10$; mean absolute value of partial correlation for significant cells), proved both stronger and, in some cases, more numerous than those signaling the difference between explore and exploit policies [39]. Indeed, such results are surprising, since the task itself contained no underlying noise to learn (apart from the random walk). Additionally, cells that coded RPE maintained this information for the duration of the trial, suggesting that single-trial outcome variables encoded in ACC [51, 52] are buffered for purposes of uncertainty estimation and within-policy adjustment. These data not only reaffirm that CGp encodes variability in options [32, 37, 38], but also situate these signals within a broader online learning framework. And while we note that such results (both effect size and frequency) are strongly model-dependent, they bolster the hypothesis that CGp encodes and maintains online variables related to the statistical structure of dynamic environments. Nevertheless, direct tests are needed in tasks where abrupt contingency changes occur, and in which change detection is necessary for optimal behavior.

Change Detection: An Active Role for the Default Mode Network?

Functionally, CGp in humans belongs to the “default network” of cortical areas, which includes ventromedial prefrontal cortex and temporo-parietal junction (TPJ). These areas show high metabolic and hemodynamic activity at rest that is suppressed during task engagement [53]. Activation in the default network is typically anti-correlated with activation in the dorsal fronto-parietal network, a set of brain areas implicated in selective attention and its concomitant benefits in accuracy and task performance [54–56]. Monkeys show strikingly similar patterns of spontaneous BOLD activity [57] and single unit activity [58,59] within this network.

What has often gone missing in the discussion, however, is an active role for the default network in cognition. One suggestion for this role is retrieval of information – especially personally-relevant information – from long-term memory [60], a function that CGp, with its strong connections to parahippocampal areas, is well situated to perform. Another is that default network may be associated with divergent thinking patterns that lead to insight and creative problem solving – a process that may interfere with performance on many simple standard laboratory tasks [61]. This possibility is closely linked with the idea that default network activity is more associated with an “exploratory” mode of cognition, and that default suppression is associated with an “exploitative” mode [8,39]. This accords well with a view in which changing between world models and task sets requires withdrawal from task performance and a re-deployment of internally directed cognition for purposes of strategy retrieval and selection.

The centrality of CGp within the default network invites the possibility that change detection may be a key sub-function of default mode processing. Indeed, large-scale environmental changes are likely to signal the need for exploratory behavior. We recently reported that

baseline activity of CGp neurons is suppressed during task performance, and that spontaneous firing rates predict subsequent task engagement on a trial-by-trial basis [59]. Specifically, higher firing rates predicted poorer performance on simple orienting and memory tasks [59], while cued rest periods, in which monkeys were temporarily liberated from exteroceptive vigilance, evoked the highest firing rates. Importantly, local field potentials in the gamma band, which has been closely linked to synaptic activity (and by extension, the fMRI BOLD signal), were also suppressed by task engagement. These results fit BOLD signal measurements showing lower CGp activity during task engagement [53,62,63,64]. Firing rates of CGp neurons were likewise suppressed when monkeys were explicitly cued to switch tasks, but activity gradually increased on subsequent trials, suggesting relaxation of cognitive control [58]. The fact that CGp neurons track task engagement suggests that the characteristic functions of default network, including monitoring, are suppressed when operating within a stable, well-learned environment. By contrast, periods of rest may be accompanied by more generalized exploratory behavior requiring reduced focus and maximum flexibility. Even brief pauses in task performance liberate agents—monkey or human—from the need for focused task engagement, thus permitting self-directed cognition, cognitive housekeeping, and mind-wandering [59,65].

Thus CGp and the default network may play a broader role in basic cognitive processes typically suppressed during performance of well-learned tasks, including memory retrieval, internal monitoring, and the global balance of internal versus external information processing. As a result, we predict CGp would respond most strongly during the initial phases of learning, in response to sudden environmental changes, and during self-initiated switches between behavioral policies. Conversely, we predict greatest deactivations during performance of outwardly directed effortful tasks that demand attention and engagement, and a return to baseline during breaks between trials [59], and even on repetitions of trials of the same type, which presumably demand less attention [58]. These responses are quite distinct from those seen in other cortical areas, especially the frontoparietal attention network.

The prevalence of increased default network activity outside task conditions in most studies would appear, on its surface, to militate against our model, as does the observation that firing in CGp neurons *decreases* following changes in the task at hand [58]. In fact, these responses further bolster the common view that default activation is simply the complement of activity in the frontoparietal attention network [56]. Yet more recent studies have shown that in tasks requiring goal-directed introspection, default network activity shows strong positive correlation with activity in frontoparietal networks [33,34], suggesting that repetitive performance in well-learned tasks may underrepresent the role of default areas in active cognition. CGp lesions in both humans [35] and rodents [36] result in deficits in tasks requiring implementation of new strategies in multitasking scenarios and changes of cognitive set. Indeed, in richer environments where altering strategy requires periodically withdrawing from the current cognitive set in order to evaluate and decide, we expect to see higher co-activation of default and attentional networks. In the same way, we hypothesize that the deactivation observed in animals switching between two well-learned tasks [58] results from the fact that changes in task demands were explicitly cued by the environment—no evaluation of evidence or internal query was required. In a task where behavioral change requires first inferring the presence of a switch, we predict increased activity in CGp.

Extending the Model

We have proposed that CGp is a key node in the network responsible for environmental change detection and subsequent alterations in behavioral policy. This proposed network sits atop the reinforcement learning module in the cognitive hierarchy, enabling organisms to

learn and implement a variety of behavioral responses to diverse environmental demands, refining each independently and employing them adaptively in response to change.

This model makes several key predictions. First, CGp activity should show pronounced enhancement in scenarios that demand endogenously-driven, as opposed to exogenously-cued, changes in behavior. That is, when statistical inference becomes necessary to detect environmental change and alter behavior, CGp should show a concomitant rise in firing rate as evidence mounts, followed by a gradual fall-off as behavior crystallizes into a single policy. Naturally, this behavior requires that information be maintained and integrated across trials, and thus that firing rates in the present exhibit correlations with outcomes in the past. Likewise, set shifts resulting from inferred change points should elicit stronger modulations in CGp activity than cued change points, since the latter require no integration of evidence. Consistent with this prediction, in a task-switching paradigm with random, explicitly cued switches [58], suppression of CGp firing reflected both task engagement and task switching, presumably because change detection could not rely on statistical inference.

Second, the process of learning entirely new associations should enhance CGp activity. This follows not only from anatomical connections between CGp and parahippocampal gyrus, which is necessary for long-term memory formation [66,67], but also from conditioning experiments showing enhanced CGp activity during learning [68]. We predict that CGp activity will be more strongly modulated by new cues that predict environmental changes that require a cognitive set switch than by new cues that are irrelevant to set shifts.

Finally, the change detection hypothesis opens up new avenues for probing default network function. If the modulations in BOLD activity in fMRI studies merely represent, as we predict, one end on a continuum of resource allocation, the idea of changes in cognitive set may offer a new perspective on default mode activation. In this framework, default network areas may be crucial for initiating transitions between basic modes of behavior, or even overriding them. This may be particularly relevant in schizophrenia, in which patients exhibit hyper-vigilance to behaviorally salient events in the environment, but simultaneously show a diminished ability to “turn down” the internal milieu [69,70]. Likewise, early degeneration in CGp in Alzheimer’s disease may cripple a key node in the interface between cognitive set and memory networks, resulting in disorientation and impaired memory access [71,72]. Together, these observations indicate a healthy CGp is necessary for organizing flexible behavior in response to an ever-changing environment, by mediating learning, memory, control, and reward systems to promote adaptive behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors were supported by NIH grants EY103496 (MLP), EY019303 (JP), F31DA028133 (SRH), and by a career development award (DA027718) and a NIDA post-doctoral fellowship (DA023338) (BYH), as well as by the Duke Institute for Brain Sciences.

Glossary

Reinforcement learning (RL)	a computational learning model in which organisms adapt to an environment by incrementally altering behavior in response to rewards and punishments
------------------------------------	---

Classical or Pavlovian conditioning	the process by which environmental stimuli become associated, via learning, with the prediction of reward
State space	the collection of distinct possible conditions for an agent-environment system in formal models of learning. States may be distinguished by, among other factors: the number and type of available choices, the information available to the agent, and the response properties of the environment
Markov property	the mathematical assumption that transitions between states depend only upon the current state, not a decision agent's entire history: $p(s_i s_{i-1}, s_{i-2}, \dots) = p(s_i s_{i-1})$, where s_i is a state and $p(s_i)$ is the probability of transitioning to that state
Markov structure	the set of transition probabilities $p(s_i s_j)$ that defines the relationship between states in a space with the Markov property
World model or model	a set of states of the world, their transition relations, and the distributions of outcomes from those states
Bayes' Rule	a rule of statistical inference used for updating uncertainty about statistical parameters (θ) based on prior information ($p(\theta)$) and new observations (x_i)
	$p(\theta x_i) = \frac{p(x_i \theta)p(\theta)}{p(x_i)}$
Bayesian learning	a set of learning models in which agents maintain knowledge as probability distributions updated by Bayes' Rule
Policy	a mapping (potentially probabilistic) between environmental variables like states and values, and actions. Policies implicitly depend on world models, since not all world models share the same sets of variables
Cognitive set or set	the combination of world model, policy, and attentional factors that govern performance in a task

References

1. Dosenbach N, et al. A dual-networks architecture of top-down control. *Trends Cogn Sci.* 2008; 12:99–105. [PubMed: 18262825]
2. Nakahara K, et al. Functional MRI of macaque monkeys performing a cognitive set-shifting task. *Science.* 2002; 295:1532–1536. [PubMed: 11859197]
3. Gusnard DA, Raichle ME. Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience.* 2001; 2:685–694.
4. Schultz W. Behavioral dopamine signals. *Trends in Neurosciences.* 2007; 30:203–210. [PubMed: 17400301]
5. Sutton, RS.; Barto, AG. *Reinforcement Learning: An Introduction.* MIT Press; 1998.
6. Pearce J, Bouton M. Theories of associative learning in animals. *Annual Review of Psychology.* 2001; 52:111–139.
7. Courville AC, et al. Bayesian theories of conditioning in a changing world. *Trends Cogn Sci.* 2006; 10:294–300. [PubMed: 16793323]

8. Daw ND, et al. Cortical substrates for exploratory decisions in humans. *Nature*. 2006; 441:876–879. [PubMed: 16778890]
9. Behrens TE, et al. Learning the value of information in an uncertain world. *Nat Neurosci*. 2007; 10:1214–1221. [PubMed: 17676057]
10. Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. *Neuron*. 2005; 46:681–692. [PubMed: 15944135]
11. Nassar M, et al. An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. *Journal of Neuroscience*. 2010; 30:12366–12378. [PubMed: 20844132]
12. Gallistel C, et al. The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes*. 2001; 27:354–372. [PubMed: 11676086]
13. Daw N, Courville A. The rat as particle filter. *Advances in Neural Information Processing*. 2007; 20:369–376.
14. Hampton AN, et al. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*. 2008; 105:6741–6746.
15. Green CS, et al. Alterations in choice behavior by manipulations of world model. *Proceedings of the National Academy of Sciences*. 2010; 107:16401–16406.
16. Minoshima S, et al. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol*. 1997; 42:85–94. [PubMed: 9225689]
17. Yoshiura T, et al. Diffusion tensor in posterior cingulate gyrus: correlation with cognitive decline in Alzheimer's disease. *NeuroReport*. 2002; 13:2299–2302. [PubMed: 12488815]
18. Whitfield-Gabrieli S, et al. Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *Proceedings of the National Academy of Sciences*. 2009; 106:1279–1284.
19. Walton ME, et al. Adaptive decision making and value in the anterior cingulate cortex. *Neuroimage*. 2007; 36(Suppl 2):T142–154. [PubMed: 17499161]
20. Rudebeck PH, et al. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci*. 2008; 28:13775–13785. [PubMed: 19091968]
21. Quilodran R, et al. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*. 2008; 57:314–325. [PubMed: 18215627]
22. Kennerley SW, et al. Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *J Cogn Neurosci*. 2008; 21:1162–1178. [PubMed: 18752411]
23. Holroyd CB, Coles MG. Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex*. 2008; 44:548–559. [PubMed: 18387587]
24. Hayden BY, et al. Fictive reward signals in the anterior cingulate cortex. *Science*. 2009; 324:948–950. [PubMed: 19443783]
25. Seo H, Lee D. Behavioral and neural changes after gains and losses of conditioned reinforcers. *J Neurosci*. 2009; 29:3627–3641. [PubMed: 19295166]
26. Hayden BY, Platt ML. Neurons in anterior cingulate cortex multiplex information about reward and action. *J Neurosci*. 2010; 30:3339–3346. [PubMed: 20203193]
27. Kennerley SW, Wallis JD. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *European Journal of Neuroscience*. 2009; 29:2061–2073. [PubMed: 19453638]
28. Maddock RJ, et al. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*. 2001; 104:667–676. [PubMed: 11440800]
29. Maddock RJ, et al. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Hum Brain Mapp*. 2003; 18:30–41. [PubMed: 12454910]
30. Levy I, et al. Neural Representation of Subjective Value Under Risk and Ambiguity. *Journal of neurophysiology*. 2010; 103:1036. [PubMed: 20032238]
31. Kable JW, Glimcher PW. The neural correlates of subjective value during intertemporal choice. *Nat Neurosci*. 2007; 10:1625–1633. [PubMed: 17982449]

32. McCoy AN, Platt ML. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience*. 2005; 8:1220–1227.
33. Spreng RN, et al. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage*. 2010; 53:303–317. [PubMed: 20600998]
34. Gerlach KD, et al. Solving future problems: Default network and executive activity associated with goal-directed mental simulations. *NeuroImage*. 2011 In Press, Uncorrected Proof.
35. Burgess P, et al. The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*. 2000; 38:848–863. [PubMed: 10689059]
36. Ng C-W, et al. Double Dissociation of Attentional Resources: Prefrontal Versus Cingulate Cortices. *J Neurosci*. 2007; 27:12123–12131. [PubMed: 17989278]
37. McCoy AN, et al. Saccade reward signals in posterior cingulate cortex. *Neuron*. 2003; 40:1031–1040. [PubMed: 14659101]
38. Hayden BY, et al. Posterior cingulate cortex mediates outcome-contingent allocation of behavior. *Neuron*. 2008; 60:19–25. [PubMed: 18940585]
39. Pearson JM, et al. Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr Biol*. 2009; 19:1532–1537. [PubMed: 19733074]
40. Gold JJ, Shadlen MN. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*. 2002; 36:299–308. [PubMed: 12383783]
41. Gold JJ, Shadlen MN. The neural basis of decision making. *Annu Rev Neurosci*. 2007; 30:535–574. [PubMed: 17600525]
42. Tsai H, et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 2009; 324:1080–1084. [PubMed: 19389999]
43. Haber S. The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy*. 2003; 26:317–330. [PubMed: 14729134]
44. Haber S, et al. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *Journal of Neuroscience*. 2000; 20:2369–2382. [PubMed: 10704511]
45. Lee D, et al. Functional specialization of the primate frontal cortex during decision making. *J Neurosci*. 2007; 27:8170–8173. [PubMed: 17670961]
46. Schoenbaum G, Esber G. How do you (estimate you will) like them apples? Integration as a defining trait of orbitofrontal function. *Current Opinion in Neurobiology*. 2010; 20:205–211. [PubMed: 20206497]
47. Schoenbaum G, et al. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience*. 2009; 10:885–892.
48. Matsumoto M, et al. Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci*. 2007; 10:647–656. [PubMed: 17450137]
49. Barraclough DJ, et al. Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci*. 2004; 7:404–410. [PubMed: 15004564]
50. Lee D, Seo H. Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. *Ann N Y Acad Sci*. 2007; 1104:108–122. [PubMed: 17347332]
51. Seo H, Lee D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci*. 2007; 27:8366–8377. [PubMed: 17670983]
52. Rushworth MF, Behrens TE. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci*. 2008; 11:389–397. [PubMed: 18368045]
53. Buckner RL, et al. The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci*. 2008; 1124:1–38. [PubMed: 18400922]
54. Corbetta M, et al. Neural systems for visual orienting and their relationships to spatial working memory. *J Cogn Neurosci*. 2002; 14:508–523. [PubMed: 11970810]
55. Hopfinger JB, et al. The neural mechanisms of top-down attentional control. *Nat Neurosci*. 2000; 3:284–291. [PubMed: 10700262]
56. Fox MD, et al. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:9673–9678. [PubMed: 15976020]

57. Vincent JL, et al. Coherent spontaneous activity identifies a hippocampal-parietal memory network. *Journal of Neurophysiology*. 2006; 96:3517–3531. [PubMed: 16899645]
58. Hayden B, et al. Cognitive Control Signals in Posterior Cingulate Cortex. *Frontiers in Human Neuroscience*. 2010:4. [PubMed: 20198130]
59. Hayden BY, et al. Electrophysiological correlates of default-mode processing in macaque posterior cingulate cortex. *Proc Natl Acad Sci U S A*. 2009; 106:5948–5953. [PubMed: 19293382]
60. Kim H, et al. Overlapping brain activity between episodic memory encoding and retrieval: Roles of the task-positive and task-negative networks. *NeuroImage*. 2010; 49:1045–1054. [PubMed: 19647800]
61. Kounios J, et al. The Prepared Mind. *Psychological Science*. 2006; 17:882–890. [PubMed: 17100789]
62. Andrews-Hanna JR, et al. Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*. 2010; 65:550–562. [PubMed: 20188659]
63. Margulies DS, et al. Precuneus shares intrinsic functional architecture in humans and monkeys. *Proceedings of the National Academy of Sciences*. 2009; 106:20069–20074.
64. Gusnard D, et al. Role of medial prefrontal cortex in a default mode of brain function. *Neuroimage*. 2001; 13:S414–S414.
65. Mason MF, et al. Wandering minds: The default network and stimulus-independent thought. *Science*. 2007; 315:393–395. [PubMed: 17234951]
66. Vincent JL, et al. Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*. 2007; 447:83–84. [PubMed: 17476267]
67. Yukie, M.; Shibata, H. Temperocingulate interactions in the monkey. In: Vogt, B., editor. *Cingulate Neurobiology and disease*. Oxford University Press; 2009. p. 145-162.
68. Gabriel M, et al. Hippocampal control of cingulate cortical and anterior thalamic information processing during learning in rabbits. *Experimental Brain Research*. 1987; 67:131–152.
69. Broyd SJ, et al. Default-mode brain dysfunction in mental disorders: A systematic review. *Neuroscience & Biobehavioral Reviews*. 2009; 33:279–296. [PubMed: 18824195]
70. Garrity AG, et al. Aberrant “default mode” functional connectivity in schizophrenia. *American Journal of Psychiatry*. 2007; 164:450–457. [PubMed: 17329470]
71. Nestor PJ, et al. Retrosplenial cortex (BA 29/30) hypometabolism in mild cognitive impairment (prodromal Alzheimer's disease). *European Journal of Neuroscience*. 2003; 18:2663–2667. [PubMed: 14622168]
72. Zhou Y, et al. Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment. *Alzheimer's and Dementia*. 2008; 4:265–270.
73. Vogt, B. *Cingulate Neurobiology and Disease*. Oxford University Press; 2009.
74. Wilson R, et al. Bayesian online learning of the hazard rate in change-point problems. *Neural computation*. 2010; 22:2452–2476. [PubMed: 20569174]
75. Ma W, et al. Bayesian inference with probabilistic population codes. *Nature Neuroscience*. 2006; 9:1432–1438.
76. Vogt BA, et al. Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb Cortex*. 1992; 2:435–443. [PubMed: 1477524]
77. Sutherland RJ, et al. Contributions of cingulate cortex to two forms of spatial learning and memory. *J Neurosci*. 1988; 8:1863–1872. [PubMed: 3385478]
78. Gron G, et al. Brain activation during human navigation: gender-different neural networks as substrate of performance. *Nat Neurosci*. 2000; 3:404–408. [PubMed: 10725932]
79. Bussey TJ, et al. Dissociable effects of anterior and posterior cingulate cortex lesions on the acquisition of a conditional visual discrimination: facilitation of early learning vs. impairment of late learning. *Behav Brain Res*. 1996; 82:45–56. [PubMed: 9021069]
80. Hirono N, et al. Hypofunction in the posterior cingulate gyrus correlates with disorientation for time and place in Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*. 1998; 64:552–554.

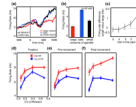


Figure 1.

CGp encodes reward outcomes over multiple trials and predicts changes in strategy. **(a)** PSTH for example neuron following reward delivery when monkeys choose between variably rewarded outcomes and deterministically rewarded outcomes with the same mean reward rate. Firing rates were significantly greater following small or medium rewards than following large rewards. **(b)** Bar graph showing the average firing of all neurons in the population following delivery of large, medium, and small rewards. Firing rates are averaged over a 1s epoch beginning at the time of reward offset ($t=0$). Tick marks indicate one standard error. **(c)** Average effect of reward outcome on neuronal activity up to five trials in the future. Bars indicate one standard error. **(d–f)** Average neuronal firing rates as a function of coefficient of variation (CV) in reward size plotted for 3 trial epochs in the probabilistic reward task. Firing increased for saccades into the neuronal response field (RF), and was positively correlated with reward CV (standard deviation/mean). Bars represent s.e.m. Adapted from [38] and [32].

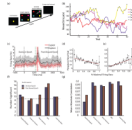


Figure 2.

CGp neurons encode variance/learning rate and change probability in a volatile environment. **(a)** Schematic of the four-armed bandit task. Four targets appear each trial, each baited with an independently varying juice reward. **(b)** Sample payouts and choices for the four options over a single block. Reward values for each target independently follow a random walk process. Black diamonds indicate the monkey's choices during the given block. **(c)** PSTH for example neuron in the 4-armed bandit task, showing significant differences in firing for exploratory and exploitative strategies in both the decision and evaluation epochs. Exploit trials are in red, explore trials in black. The task begins at time 0. Onset of the "go" cue (dashed green line), reward delivery (solid red line), beginning of intertrial interval (dashed gray line), and end of trial (rightmost dashed black line) are mean times. Dashed red lines indicate \pm one standard deviation in reward onset. Shaded areas represent s.e.m. of firing rates. **(d,e)** Neurons in CGp encode probability of exploring on the next trial. Points are probabilities of exploring next trial as a function of percent maximal firing rate in the decision epoch, averaged separately over negatively- and positively-tuned populations of neurons (d and e, respectively). **(f)** Numbers of neurons encoding relevant variables in a Kalman filter learning model of behavior [8,39]. Bars indicate numbers of significant partial correlation coefficients for each variable with mean firing rate in each epoch when controlling for others. The decision epoch (blue) lasted from trial onset to target selection; the post-reward epoch (red) lasted from the end of reward delivery to the end of the inter-trial interval. Dotted line indicates a statistically significant population of neurons, assuming a $p=0.05$ false positive rate. Variance chosen indicates the estimated variance in the value of the chosen option. Mean variance indicates the mean of this quantity across all options. **(g)** Mean absolute partial correlation coefficients for significant cells. Colors are as in (f). Adapted from [39].



Figure 3.

A simplified schematic of change detection and policy selection. Sensory feedback from reward outcomes is divided into task-specific variables and passed on to both a reinforcement learning module and a change detector. The learning module computes an update rule based on the difference between expectations and outcomes in the current world model, and updates the policy accordingly. The change detector calculates an integrated log probability that the environment has undergone a change to a new state. If this variable exceeds a threshold, the policy selection mechanism substitutes a new behavioral policy, which will be updated according to subsequent reward outcomes.

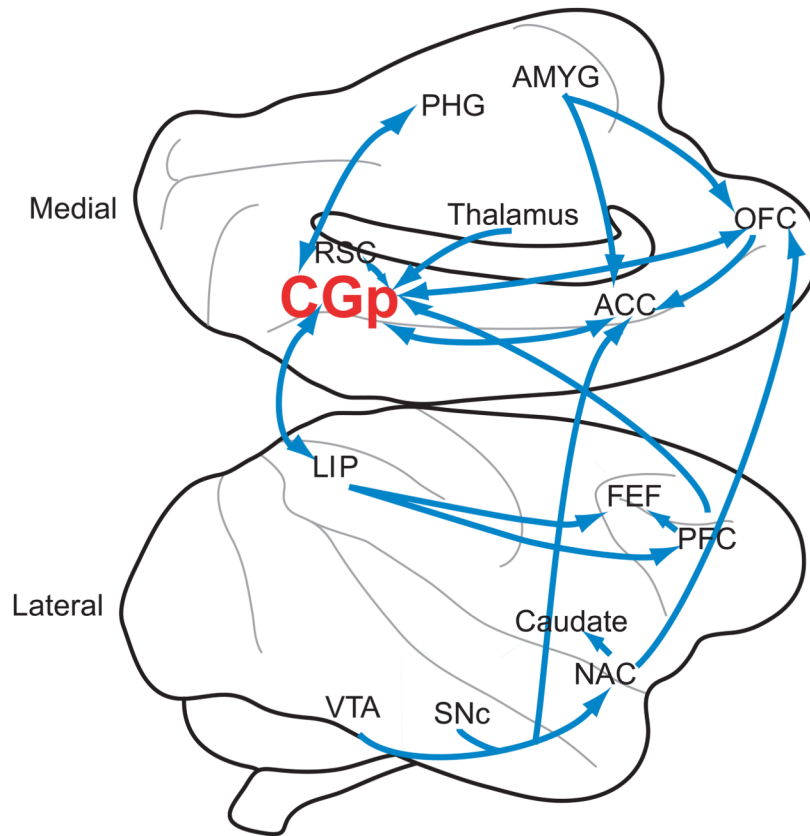
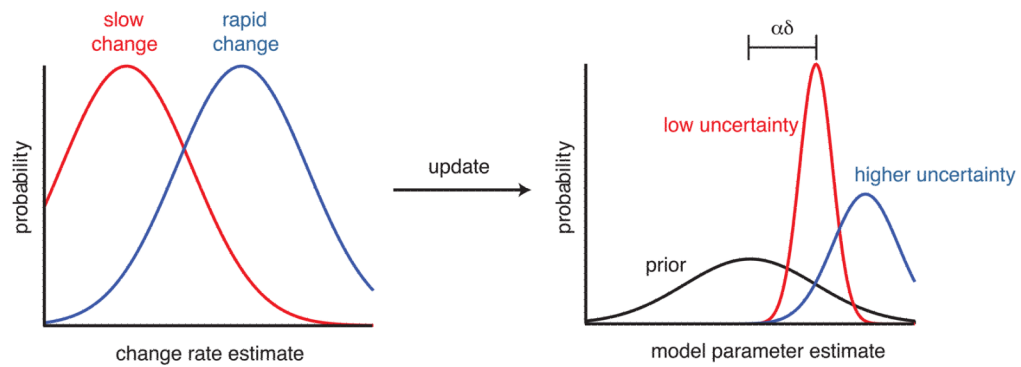
**Box 1, Figure I.**

Figure shows medial (top) and lateral (bottom) views of the macaque brain. Significant neuroanatomical connections to and from CGp are shown. Note that the figure does not represent a thorough diagramming of innervations, as CGp connects to a large number of brain regions. Generally, CGp neuroanatomy is similar in humans and macaques [73]. Abbreviations: VTA=ventral tegmental area; SNc=Substantia nigra pars compacta; NAC=nucleus accumbens; PFC=prefrontal cortex; FEF=frontal eye field; LIP=lateral intraparietal area; CGp=posterior cingulate cortex; ACC=anterior cingulate cortex; RSC=retrosplenial cortex; PHG=parahippocampal gyrus; OFC=orbitofrontal cortex; AMYG=amygdala

**Box 2, Figure I.**

Bayesian update rules for learning. Model parameters are initially estimated as prior distributions (black). When outcomes are observed, distribution means are shifted by the product of learning rate, α , and prediction error, δ , while variances change based on the estimated rate of environmental change (left). When the estimated rate of environmental change is low (red), both mean updates and uncertainty (proportional to variance) are likewise small. When change is rapid (blue), means are updated rapidly, but uncertainty remains high.

Table 1

Proposed functions of CGp, with selected references.

Functional claims	Refs.
Evaluating sensory events and behavior to guide movement and memory	[76]
Spatial learning and navigation	[77,78]
Late stages of reinforcement learning	[79]
Orientation in time and space	[80]
Autobiographical memory retrieval	[28]
Emotional stimulus processing	[29]
Reward outcome monitoring	[37]
Representation of subjective value	[30–32]
Problem-solving via insight	[61]
Mind-wandering	[65]
Action evaluation and behavioral modification	[38]
Goal-directed cognition	[33]
Problem-solving via insight	[61]
Mind-wandering	[65]