

# 18 An Integrative Theory of Anterior Cingulate Cortex Function: Option Selection in Hierarchical Reinforcement Learning

Clay B. Holroyd and Nick Yeung

The anterior cingulate cortex (ACC), a region on the medial surface of the frontal lobes, has received substantial research attention in recent years as a key neural substrate of cognitive control. In part this interest derives from neuroanatomical evidence that ACC is uniquely well positioned to collate information about the motivational significance of ongoing events—based on its inputs from the limbic system, orbitofrontal cortex (OFC), and midbrain dopamine system—and to use this information to guide behavior via its dense interconnections with primary motor, premotor, and lateral prefrontal cortex.<sup>41</sup> In accord with the idea that ACC plays this central role in behavioral control, human neuroimaging studies have shown that the area consistently coactivates with regions in the lateral prefrontal and parietal cortex as part of an “executive” network responding to a diverse range of cognitive demands.<sup>19</sup> However, although there is widespread agreement that ACC plays an important role in motivational and cognitive control, far less agreement exists as to what that role might be.

Theories of ACC function fall into four main categories: *Performance monitoring* theories emphasize its role in evaluating ongoing behavior, detecting errors or conflicts in response execution, and implementing remedial actions as appropriate.<sup>9,44</sup> *Action selection* theories focus on the contribution of ACC to the internal, willful generation of behavior.<sup>18,41</sup> *Reinforcement learning* (RL) theories propose a role for ACC in learning action values that can be used to select appropriate goal-directed behaviors.<sup>25,45</sup> Finally, *motivational* theories underscore the sensitivity of ACC to affect, effort, and costs.<sup>52,56</sup> These theoretical frameworks are not mutually exclusive and in key respects share important claims. For example, performance monitoring and RL theories agree that ACC is sensitive to the efficiency and effectiveness of chosen actions (see also chapter 17, this volume), while both action selection and motivational theories emphasize the contribution of ACC to the willed generation of behavior. Nevertheless, despite this overlap, a unifying theory that spans the wealth of existing anatomical, neurophysiological, neuroimaging, and lesion data has yet to be developed.

In this chapter we propose a new account of ACC function that integrates several salient features of existing theories while aiming to reconcile their inconsistencies. Our account focuses specifically on the dorsal region of ACC believed to be involved in cognitive control, rather than on the rostral-ventral subdivision that is more involved in emotional processing.<sup>16</sup> We propose that dorsal ACC supports the selection and execution of coherent behaviors over extended periods, an idea we formalize in terms of recent advances in the theory of RL that use a hierarchical mechanism for action selection to choose between *options*—sequences of primitive actions associated with particular goals.<sup>7</sup> This proposal builds on existing theories rather than representing a radical departure from them. We therefore begin with a review of these theories.

## **Current Theories of ACC Function**

### **Performance Monitoring**

An influential hypothesis within the human neuroimaging literature is that ACC monitors for signs of inefficient or suboptimal performance to signal when increased cognitive control is required.<sup>6,9,44</sup> Initial evidence for this view came from primate neurophysiology<sup>22</sup> and human electroencephalographic<sup>20</sup> evidence of ACC activity following errors. In the scalp-recorded electroencephalogram (EEG), for example, a component labeled the error-related negativity (ERN or Ne) is elicited by errors in speeded decision-making tasks,<sup>20</sup> while a related feedback ERN (fERN) is seen following error feedback in trial-and-error learning tasks.<sup>33</sup> Converging evidence from dipole modeling, EEG-informed fMRI, and intracranial recording studies indicates that the ERN and fERN are generated in ACC.

However, ACC activity is also apparent in conditions of increased cognitive demand in the absence of errors.<sup>9</sup> For example, increased ACC activity is observed in the Stroop task when the presented word is incongruent with the required color-naming response, even when participants ultimately respond correctly. Although such findings do not rule out that ACC is involved in error processing—for example, ACC may predict error likelihood rather than specifically detect errors as they occur<sup>8</sup>—they have nevertheless motivated a prominent alternative account of ACC function. According to this view, ACC monitors for occurrences of conflict between incompatible actions, such as the competing responses cued by color and word information in the Stroop task, to signal the need for increased cognitive control by dorsolateral prefrontal cortex (DLPFC).<sup>6</sup> Formal computational models have shown this theory to account for a range of findings regarding conflict- and error-related activity in ACC. Subsequent neuroimaging studies have confirmed key predictions of the theory. For example, in the Stroop task, high levels of conflict-related

ACC activity on one trial are predictive of increased DLPFC activity on the next trial, leading to reduced conflict and improved performance on that trial, consistent with the notion that ACC and DLPFC form a regulatory feedback loop in cognitive control.<sup>28</sup>

These successes notwithstanding, conflict monitoring has been challenged as a comprehensive account of ACC function. First, human neuroimaging evidence indicates that ACC not only shows transient responses to experienced conflict as predicted by the theory, but also shows sustained activation during task preparation and execution.<sup>17</sup> This finding suggests that ACC plays a broader role in cognitive control than the conflict-monitoring theory proposes. Second, although some patients with ACC lesions exhibit deficits in Stroop task performance and impaired conflict adaptation,<sup>53</sup> these effects are far less consistently observed than the conflict-monitoring theory predicts.<sup>21</sup> Finally, single-unit recording studies in monkeys have failed to find convincing evidence of conflict-related activity in ACC.<sup>36</sup> Instead, it has been suggested that conflict modulates movement-specific neural activity and that conflict-related activity seen in human neuroimaging studies may be an artifact created by averaging activity across populations of neurons that selectively code for the conflicting responses.

### Action Selection

Whereas performance monitoring accounts suggest that ACC plays a primarily evaluative role, an alternative class of theories suggests that ACC contributes directly to the generation and control of behavior. In support of this view, neuroanatomical evidence indicates that ACC comprises several distinct premotor areas,<sup>18</sup> while human neuroimaging studies demonstrate functional specialization within ACC for movements of different effectors.<sup>43</sup> Single-unit recordings in monkeys indicate that neuronal firing in ACC may precede overt movements by several hundred milliseconds.<sup>48</sup> Converging evidence from this primate work and from human neuroimaging indicates that this movement-related activity increases when actions are internally selected rather than externally instructed.<sup>15,48</sup> Further, patients with cingulate lesions exhibit characteristic reductions in their spontaneous speech and movement.<sup>13</sup> These deficits in self-initiated movements are seen in extreme form in akinetic mutism following bilateral damage to ACC and surrounding cortex, in which the awake patient remains immobile and unresponsive to external stimuli beyond simple eye movements.<sup>37</sup> Taken together, these findings are suggestive of a role for ACC in voluntary, or “willed,” action selection.<sup>41</sup>

However, despite compelling evidence that ACC is involved in the generation of voluntary actions, the precise functional role of the region remains unclear: whether ACC selects particular actions or provides a generalized motivating arousal signal; what specific computations ACC might perform in order to drive or guide behavior;

and how the role of ACC complements the operations of other motor regions such as the supplementary motor area (SMA) and basal ganglia. Moreover, the role of ACC does not appear to be limited to voluntary action selection: As reviewed in the preceding section, ACC activity is robustly observed during performance of stimulus-driven as well as self-initiated actions,<sup>9</sup> and in association with the evaluation (e.g., through feedback) as well as the selection of actions.<sup>33</sup>

### **Reinforcement Learning**

RL theories agree that ACC plays an important role in action selection, but propose a computationally specific account of this role that explains its sensitivity to performance feedback and action outcomes. According to RL approaches, actions are selected on the basis of stored values derived from their past association with positive and negative outcomes.<sup>11,34,38</sup> ACC has been proposed to play a pivotal role in linking actions and their outcomes according to RL principles.<sup>25,45</sup> Consistent with this view, ACC neurons in primates are sensitive to the degree and magnitude of expected rewards,<sup>2,47</sup> code for reward prediction errors associated with action selection,<sup>29</sup> and fire in relation to both actions and rewards in a manner that appears to link these events.<sup>24</sup> Disturbances of normal ACC function impair animals' ability to switch to alternative behaviors following the reduction of an expected reward,<sup>49</sup> and disrupt the utilization of outcome information for learning about action-reward relationships.<sup>23,27,58</sup>

Human functional neuroimaging studies also suggest that ACC learns about the consequences of internally generated actions.<sup>26,55</sup> These learning-related changes appear to be instigated by the activity of the midbrain dopamine system, which projects to, and reaches its highest density over, medial regions of the frontal cortex,<sup>57</sup> and which conveys so-called reward prediction error signals to its neural targets.<sup>46</sup> It has been proposed that ACC uses these signals in adaptive decision making, and further, that this learning process elicits the ERN and fERN.<sup>25</sup>

As with performance monitoring theories, however, lesion studies in animals and patients provide only partial support for RL accounts. Thus, whereas some lesion studies in animals indicate deficits related to fast trial-to-trial learning,<sup>49</sup> others suggest that the deficits relate to integration of reward information across multiple trials.<sup>27</sup> More troubling still, ACC damage in humans appears to spare feedback-based learning in the Wisconsin Card Sort Test (WCST) despite disrupting spontaneous movement production.<sup>13</sup> Human neuroimaging evidence that ACC activity is consistently observed in tasks such as the Stroop task—in which no reward is provided and responses are instructed rather than learned—is likewise problematic for RL theories: This evidence suggests that ACC implements a specific computational function beyond simply associating actions with outcomes. A similar conclusion follows from an important conceptual challenge to RL theories: that the function

ascribed to ACC by these theories is more commonly attributed to the basal ganglia.<sup>39</sup> The fact that human behavior in standard decision-making tasks can be accounted for without recourse to simulating ACC<sup>11</sup> suggests strongly that ACC implements a function that is not exploited by these tasks (see also chapter 17, this volume).

### **Motivation and Effort**

ACC has been associated with motivation and emotion since Papez first identified this structure within the limbic circuit.<sup>40</sup> Subsequent articulations of this idea held that ACC monitors the motivational significance of stimuli and events<sup>31</sup> and integrates hedonic value with action plans,<sup>52</sup> positions that have been supported by observations of converging limbic connections onto the cingulate motor areas.<sup>35</sup> Such considerations have led to the view that ACC does not directly mediate performance monitoring or action selection, but may instead produce affective responses and concurrent autonomic activity to salient events as they take place.<sup>14</sup> Likewise, ACC may contribute to motivational control during task execution by supplying a “global energizing factor” that facilitates neural processes underlying decision making.<sup>50</sup> According to this view, akinetic mutism results from the withdrawal of this energizing factor.

A specific role for ACC in effort-related decision making is suggested by evidence that rats with ACC lesions tend to shift from selecting effortful actions that yield large rewards to choices that yield less reward but require less effort.<sup>54</sup> It has further been shown that dopaminergic input to ACC is essential for this function, apparently by facilitating response selection based on the relative values associated with different actions.<sup>3</sup> These observations parallel reports that human akinetic mutism is ameliorated by administration of dopamine agonists.<sup>1</sup>

However, a criticism of motivational theories is that they lack computational specificity<sup>59</sup>: It remains to be demonstrated how ACC computes qualities such as affect and effort and how these constructs mediate action production and selection. Potentially instructive in this regard are theories of dopamine function that emphasize its role in carrying “incentive salience” signals that are said to “transform the neural representation of a stimulus into an object of attraction that animals will work to acquire.”<sup>4</sup> This idea has been formalized using computational RL models in which dopaminergic signaling of action values serves to boost the probability of those actions being selected.<sup>30</sup> In the following, we suggest that these ideas might provide a framework for understanding the dual role of ACC in learning and motivation.

### **Evaluation of Current Theories**

Current theories of ACC function therefore emphasize its role in four key aspects of behavior: monitoring ongoing performance, selecting and initiating voluntary actions, learning about the consequences of actions, and motivating effortful

behavior. Crucially, however, no single theory seems capable of explaining the full range of existing findings. Thus, performance monitoring theories cannot easily explain ACC activity observed as people prepare to act, while action selection accounts provide no ready explanation of ACC activity observed following delivery of reward and feedback. Conversely, both RL and motivational theories struggle to account for ACC activity in stimulus-driven cognitive tasks that have little direct affective significance or motivational content.

Moreover, a deeper concern with all current theories is that they seemingly fail to capture adequately the unique contribution of ACC to behavior. As a striking illustration of this point, each of the theories ascribes to ACC a function that seems vital to normal cognitive processing; yet, aside from rare reports of akinetic mutism, the deficits induced by ACC lesions tend to be subtle and limited. For example, ACC-lesioned patients often show broadly intact executive functioning and learning from feedback,<sup>13</sup> and deficits in conflict paradigms such as the Stroop task are not universally observed even in patients with large bilateral lesions.<sup>21</sup> Interpretation of this neuropsychological evidence is complicated by the heterogeneity of function within ACC, variability in lesion extent and location across patients, and effects of neural reorganization and behavioral compensation. Nevertheless, taken with the limitations in existing theories noted earlier, this evidence suggests that current theories and the experimental paradigms used to test them may not effectively capture the core functions of ACC.

In the second part of this chapter, we therefore outline a new theory that aims to reconcile the central claims of existing theories while addressing their principal weaknesses. The starting point for our proposal is a desire to provide an integrative account of the key findings already discussed. To account for this range of findings, we extend previous RL theories by proposing that ACC contributes specifically to reinforcement learning of high-level, temporally extended behaviors. In so doing, we hope to explain why, despite its apparently central role in motivational and cognitive control, lesions to ACC tend to have rather subtle behavioral consequences, and thereby begin to outline new tasks that might more accurately target the proposed functions of ACC.

## **ACC and the Hierarchical Control of Action**

### **Option Selection in ACC**

Drawing on recent advances in RL theory<sup>7</sup> (see also chapter 16, this volume), we propose that ACC implements a mechanism for selecting high-level behavioral plans, or *options*, that comprise structured sequences of actions directed toward specified subgoals. Within this framework, options are defined in terms of *initiation*

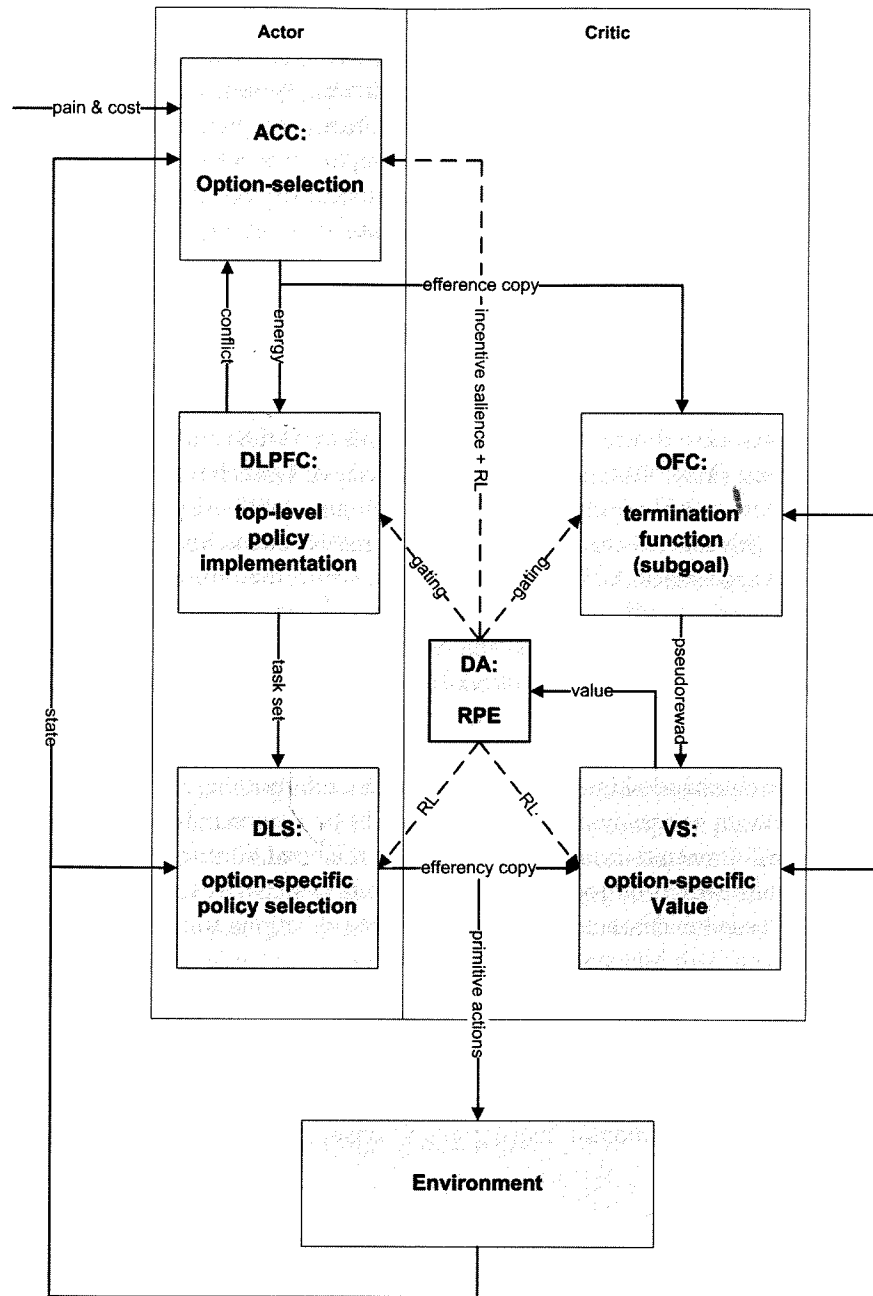
*sets* specifying states of the world in which they can be selected, option-specific *policies* specifying the individual actions they comprise, and *termination functions* specifying when option execution has been completed. Options are learned and selected according to established RL principles. In turn, completion of an option serves as a “pseudo-reward” that reinforces preceding lower-level actions according to those same principles. This hierarchical formulation of the RL problem (hierarchical reinforcement learning, or HRL) can increase computational efficiency in situations in which multiple permutations of potential courses of action can lead to a combinatorial explosion.

We specifically frame the role of ACC in HRL within an extension of the well-known actor-critic architecture, in which an actor component selects and executes behaviors and a critic evaluates the appropriateness of those actions.<sup>51</sup> Neurally, the actor is typically associated with the dorsolateral striatum (DLS) and the critic with the ventral striatum (VS).<sup>39</sup> Recent conceptualizations of HRL have extended the domains of the actor and critic to include DLPFC and OFC, respectively.<sup>7</sup> Our proposal extends this framework further by placing ACC at the apex of the actor (figure 18.1). We suggest that ACC stores (or has access to) the option-specific policies, their initiation states, and their termination functions, and uses this information in the probabilistic selection of options. The output of the option selection process is then mediated by two primary routes: via the actor (consisting of DLPFC and DLS) and the critic (consisting of OFC and VS), as detailed in the following.

According to this proposal, ACC supports the selection and execution of high-level, temporally extended sequences of responses underpinning complex, long-term behaviors. As an everyday example, ACC might be responsible for a jogger's decision to run up a mountain and for seeing that this goal is ultimately fulfilled, rather than, say, the would-be jogger staying home to watch TV. Lesions to ACC would result in behavior characterized by immediate reactions to external events rather than by extended, internally driven actions (e.g., the decision to watch TV rather than run). By contrast, ACC should be less important in standard laboratory paradigms that require learning about simple stimulus-response contingencies (such as the WCST) or that involve instructed stimulus-driven responding (such as the Stroop task). Yet even in these stimulus-dependent tasks, ACC may be responsible for compliance with experimental instructions to ensure that task performance is fast and accurate.

### **Interactions Involving ACC**

Our HRL account suggests that ACC performs complementary functions in learning, selecting, and motivating high-level behavioral options. These functions depend in turn on its interactions with other core components of the actor-critic architecture (figure 18.1). First, by way of the actor route, options selected by ACC provide



**Figure 18.1**

Schematic illustration of the proposed role of ACC in the hierarchical actor-critic reinforcement learning architecture. ACC, anterior cingulate cortex; DLPFC, dorsolateral prefrontal cortex; OFC, orbitofrontal cortex; DA, dopamine; RPE, reward prediction error; DLS, dorsolateral striatum; VS, ventral striatum.



excitatory input to the DLPFC, which in turn implements option-specific policies (i.e., sets or sequences of lower-level actions) via its connections with the DLS and other motor structures. In the example of our would-be jogger, ACC would be responsible for the decision to run up the mountain, whereas DLPFC and DLS would be responsible for implementing this decision as a specific sequence of actions. This excitatory effect of ACC option selection on DLPFC policy implementation provides one important route by which ACC can be said to energize or motivate behavior.

By way of the critic route, the OFC receives information about and assigns value to options received as “efference copy” from ACC.<sup>12</sup> Value signals from OFC serve three important functions. First, they facilitate option selection and implementation by acting as an incentive salience signal that communicates whether the option under consideration is a good one (e.g., “Yes, you should jog up the mountain”). Second, these signals provide context to the VS during option execution (“All behaviors consistent with jogging up the mountain are good”). Third, they serve as a “pseudo-reward” once the termination function has been satisfied, indicating that the goal has been achieved (“You made it to the top—good for you!”). In turn, the VS maintains a separate set of values for each option based on external state input, efference copy from the DLS, and contextual input from the OFC. For example, if contextual information from OFC indicates that the task is to jog up the hill, then the VS evaluates individual steps in that direction as being good.

The midbrain dopamine system is the lynchpin of this HRL system. First, dopaminergic signals communicate reward prediction errors, defined as instantaneous changes in value plus pseudo-reward, to induce synaptic plasticity in target structures according to RL principles.<sup>34,38</sup> This learning serves to optimize evaluative predictions of the critic (OFC and VS) and option/action selection processes of the actor (ACC, DLS). In this way, successfully completed action plans (which elicit pseudo-reward on goal obtainment) will be more likely to be selected again in the future. Second, dopaminergic projections provide incentive salience signals that facilitate option selection by ACC and policy (task-set) implementation by DLPFC: During option selection, value signals that are generated by OFC and signaled via the dopamine system cause the specifics of the associated policy to be coordinated across multiple neural structures (especially ACC, DLPFC, and OFC) by gating relevant information into working memory.<sup>10</sup> This incentive salience mechanism provides a second key route by which option selection in ACC can be said to motivate or energize behavior.

### Comparison with Existing Theories

We next outline how our theory explains key empirical findings previously taken to support the four main accounts of ACC function discussed earlier, while drawing out important differences in how our theory accounts for these results.

reinforcement learning  
x; OFC, orbitofrontal  
VS, ventral striatum.

### Performance Monitoring

Our proposal shares with performance monitoring accounts an emphasis on the sensitivity of ACC to suboptimal task performance, and on the importance of its interactions with DLPFC.<sup>32</sup> However, our theory recasts ACC primarily as a recipient rather than source of monitoring signals, and suggests that ACC directly guides behavior through selecting options that DLPFC translates into specific sets of actions. This framework provides two ways to explain neuroimaging evidence of conflict-related activity in ACC.<sup>9</sup> First, ACC might monitor for conflict between options just as it has been proposed to monitor conflict between individual actions.<sup>5</sup> According to the HRL framework,<sup>7</sup> task sets studied in the laboratory (e.g., the set of stimulus-response mappings employed in the Stroop task) are a special case of options wherein individual actions are deployed in an entirely stimulus-driven manner. Given this identity between options and task sets, it follows that ACC might be sensitive to the performance costs associated with option-level conflict—the simultaneous activation of incompatible task sets in DLPFC—a process that could lead to stronger policy implementation.

However, our framework is also consistent with the view that conflict-related activity is simply an epiphenomenon of ACC's primary role in action selection<sup>36</sup>. This activity may reflect the summed activation of multiple options in the presence of incongruent stimulus information. Alternatively, ACC may be active in high-conflict conditions because, by definition, the required response is only weakly cued by the stimulus, leading to greater or more prolonged involvement of high-level energizing input from ACC. In either view, conflict adaptation (i.e., increased control following conflict) could reflect the benefit of repeated option selection: Executing the correct task and response in the face of conflict would lead to strengthening of task-relevant associations such that the option would subsequently be executed more efficiently. This increase in efficiency would result in reduced ACC activity (reflecting the reduced need to constrain action by option context) and perhaps also increased DLPFC activity (reflecting increased effectiveness of imposing the option).

### Action Selection

According to the HRL theory, ACC input is required primarily when behavior is guided by high-level, internal constraints rather than directly cued by the environment. This interpretation provides a straightforward account of ACC activity in voluntary action selection tasks requiring unpredictable or irregular movement sequences<sup>15</sup>: Such sequences are, by definition, unconstrained with respect to environmental stimuli, precisely the conditions that demand option selection from among the many potential option-specific policies afforded. In contrast, ACC should be less important, and less active, when the range of possible behaviors is strongly

constrained (e.g., because of experimental instruction). In such situations, options and actions essentially become preselected, rendering ACC less important for effective behavior. As a consequence, standard stimulus-driven laboratory tasks (such as the Stroop task) may be relatively insensitive to ACC lesions, whereas underconstrained real-world tasks dependent on longer-term courses of action (such as jogging up a mountain) would rely on ACC to a greater extent.

Nevertheless, even for instructed tasks, ACC may (via its interactions with DLPFC) play a facilitatory role when the task has hierarchical structure, for example, by linking successive actions or by providing contextual input that weakens selection of task-irrelevant actions in the DLS. In this view, sustained ACC activity during task selection and execution<sup>17</sup> reflects its role in selecting, maintaining, and evaluating behavioral options. This process is enforced through a division of labor between ACC and DLPFC, with the former being responsible for selecting and energizing task-relevant representations, and the latter implementing those representations by providing top-down biasing signals that facilitate execution of appropriate stimulus-response mappings in lower-level structures in motor cortex and the basal ganglia.

### Reinforcement Learning

The present theory extends previous RL accounts to propose that ACC and the DLS play complementary roles in learning and selecting high-level options and individual actions, respectively. Our theory thus inherits from these accounts its explanation of ACC sensitivity to reward (primarily via dopaminergic input) and pain and costs (perhaps via input from the cingulate gyrus and insula<sup>42</sup>). However, it is distinguished by its emphasis on reinforcement of options by goal achievement (pseudo-reward) rather than of individual actions by primary reward. Learning by ACC is thus related mainly to sequences of actions (i.e., which options to choose) as opposed to the primitive actions that comprise the option: ACC learns whether to run up the mountain, not how to put one foot in front of the other. Performance feedback may therefore elicit ACC activity even in the absence of task-related behavior, because dopaminergic signals can reinforce the option itself (e.g., the decision to participate in the task) in addition to the primitive actions that comprise the task sequence.<sup>60</sup>

This reasoning implies that behavior following ACC lesions should be driven primarily by outcomes of individual actions rather than by temporally extended behavioral strategies. This prediction provides a straightforward account of primate work showing that ACC lesions cause responding to become fragmented and reliant on recent reward history rather than on rewards integrated over longer behavioral sequences.<sup>27</sup> Disruption of ACC function should similarly reduce animals' ability to represent transitions between individual actions, and hence impair their ability to switch plans flexibly when reduced reward indicates the need to change strategy<sup>49</sup>:

Without higher-level structure, new actions will not be attempted (and learned) until previous contingencies are largely extinguished. However, because ACC primarily cares about the value of actions, lesions to this region should have much less impact in stimulus-driven tasks for which contingencies reflect the value or relevance of particular stimulus attributes, as is the case for the WCST in human neuropsychology.<sup>13</sup>

### Motivation and Effort

Our theory suggests that ACC contributes to motivation and effort in two principal ways. First, options selected in ACC provide excitatory input to DLPFC, energizing the adoption of particular behavioral policies (task sets). Second, efference copy of selected options elicits value signals in OFC that are broadcast via the midbrain dopamine system to facilitate option selection by ACC and policy implementation by the DLPFC and DLS. These ACC contributions are important when goal-oriented actions are not strongly driven by the environment, thus capturing the essence of motivation and effort in a variety of contexts: when no stimuli are provided and action must be internally generated (as in voluntary selection paradigms), when stimuli are presented but strongly cue a different response from the one required (as in conflict paradigms), and when the value of the current option does not outweigh that of competing options (as when the desire to reach the top of the hill matches the desire to rest one's aching legs and lungs).

According to HRL theory, then, ACC lesions should result in biases away from extended action sequences accruing long-term reward, toward less effortful but more immediately rewarding actions. Corresponding deficits in incentive salience will cause a shift toward overall behavioral inhibition. Thus, in contrast to other formulations,<sup>5</sup> our hypothesis does not imply that ACC specifically codes for effort. If this were the case, one might expect ACC-lesioned animals to be insensitive to effort (i.e., factoring this cost poorly into their action choices). Instead, these animals show hypersensitivity and aversion to effortful options. Human patients with ACC lesions exhibit similar reductions in spontaneous speech and behavior.

### Future Directions

The option-selection theory provides a unified account of the role of ACC in performance monitoring, selecting and initiating action, learning about the consequences of actions, and motivating effortful behavior. It extends previous RL accounts<sup>25</sup> to attribute a specific computational function to ACC that distinguishes it from other neural systems—such as the basal ganglia and DLPFC—that are likewise integral to adaptive decision making.

Unsurprisingly, the theory raises as many questions as it answers. Foremost among these is how to test it: What types of task might probe the proposed functions of ACC? We have argued that many existing paradigms rely too heavily on stimulus-dependent tasks that can be solved without the advantages of HRL,<sup>11</sup> thus explaining why patients with ACC lesions show only mild and subtle deficits. Our theory allows us to sketch the outlines of a task that might be more sensitive to the contributions of ACC: The crucial elements are that responding should be underconstrained by the stimuli presented and that the task should have a hierarchical structure in which the reinforcement value of individual actions provides a weak, or even misleading, guide to the appropriate overall strategy. A valuable approach in future research would be to develop HRL-based computational models to identify appropriate task designs and to predict human performance, ACC activity, and ACC lesion effects in these contexts.

A second question concerns how the dual-level neural architecture we propose (figure 18.1) might deal with real-world problems that typically involve multiple embedded goals. For example, the desire to lead a long, healthy life might encompass as a subgoal the intent to exercise regularly, which itself might contain the subgoal of running up a mountain on a particular day, which in turn would depend on a subgoal of placing one foot in front of the other, and so on. We suggest that in practice these multilevel problems are translated into series of two-level problems, the top level of which at any given moment is mediated by the contents of working memory. For example, our jogger might one morning choose between running and relaxing, but executing the former option itself entails choosing among suboptions (e.g., regarding which route to run). During this choice, the highest-level option (the decision to run) might be temporarily cleared from working memory, to be reinstated later on completion of the lower-level decision.

A third question concerns the implications of the proposal that action selection can be shaped by the pseudo-reward associated with each option's termination function. Specifically, departing somewhat from standard treatments of HRL, our framework allows the possibility that failure to achieve the termination state of an option may reduce the strength of that option, implying a reduced preference for its corresponding subgoal. This implication seems antithetical to the fundamental assumption in the artificial intelligence approach to RL that the values of rewards are intrinsic and immutable.<sup>51</sup> But perhaps this flexibility is in fact a lineament of human behavior. In Aesop's fable, the hungry fox is unable to jump high enough to reach some grapes hanging high on the vine. Turning away in disgust, the fox grumbles: "You aren't even ripe yet! I don't need any sour grapes!" This fable serves as an apposite reminder that human goals change in response to our evolving beliefs about what we can and cannot achieve. In this way, ACC may provide the impetus for the seemingly infinite variety of individual preferences

that characterizes human culture, including the pursuits of music, science, art, and sport.

### Outstanding Questions

- What kind of experimental task will best test the role played by ACC in option selection?
- How does the brain carry out multilevel HRL problems? Must hierarchical task structure be replicated in a hierarchical neural architecture, or can multilevel problems be solved using flexible switching within a dual-level architecture as outlined here?
- How are option strengths initialized and modified?

### Further Reading

Botvinick MM, Niv Y, Barto AC. 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113: 262–280. This paper provides the conceptual foundations of the present chapter, laying out the computational framework of HRL and suggesting a possible neural implementation.

Paus T. 2001. Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci* 2: 417–424. An influential review of anatomical and functional evidence that ACC plays a pivotal role in the motivational control of behavior.

Vogt BA. 2009. *Cingulate Neurobiology and Disease*. Oxford: Oxford University Press. A comprehensive handbook on the functional neuroanatomy of cingulate cortex, with an emphasis on its cytoarchitecture and connectivity.

### References

1. Alexander MP. 2001. Chronic akinetic mutism after mesencephalic-diencephalic infarction: remediated with dopaminergic medications. *Neurorehabil Neural Repair* 15: 151–156.
2. Amiez C, Joseph J-P, Procyk E. 2005. Anterior cingulate error related activity is modulated by predicted reward. *Eur J Neurosci* 21: 3447–3452.
3. Assadi SM, Yucel M, Pantelis C. 2009. Dopamine modulates neural networks involved in effort-based decision-making. *Neurosci Biobehav Rev* 33: 383–393.
4. Berridge KC, Robinson TE. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28: 309–369.
5. Botvinick MM. 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* 7: 356–366.
6. Botvinick MM, Braver TS, Carter CS, Barch DM, Cohen JD. 2001. Conflict monitoring and cognitive control. *Psychol Rev* 108: 624–652.
7. Botvinick MM, Niv Y, Barto AC. 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113: 262–280.
8. Brown JW, Braver TS. 2005. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 307: 1118–1121.

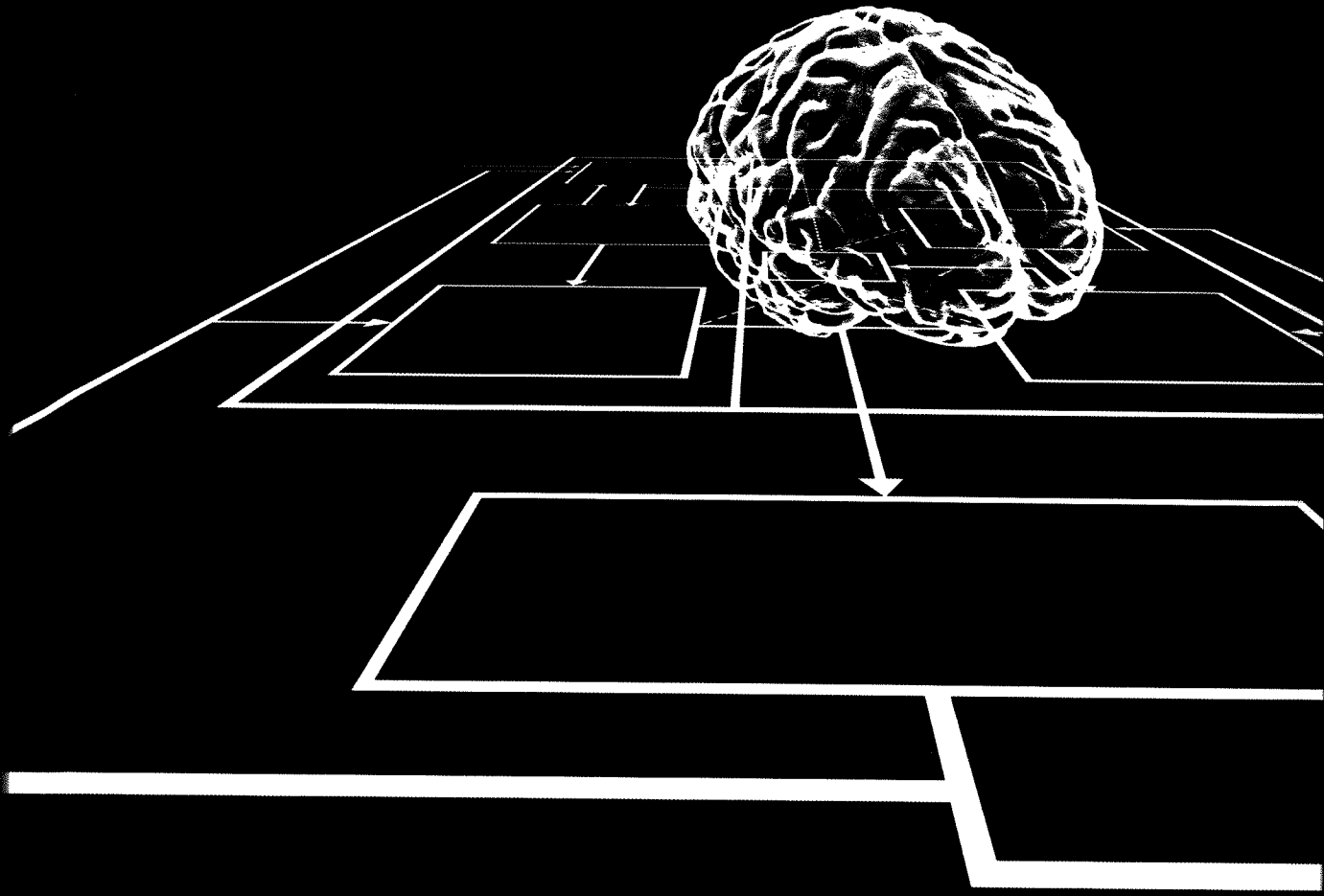
9. Carter
- error de
10. Coh
- frontal
11. Coh
- memory
12. Coh
- orbitofr
13. Coh
- of inten
- Clin Neu*
14. Crit
- RJ. 2003
- Brain* 12
15. Deib
- external
- 81: 3065
16. Dev
- Brain* 11
17. Dose
- Grimes
- 50: 799–
18. Dum
- Thalamu
19. Dum
- demands
20. Falke
- on the E
- Gaillard
21. Fello
- 788–796.
22. Gem
- 24) of m
23. Hadl
- guided se
24. Haye
- and actio
25. Holr
- dopamin
26. Holr
- anterior
- 497–498.
27. Kenn
- and the a
28. Kern
- conflict m
29. Mats
- tion erro
30. McCl
- Neurosci*

9. Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280: 747–749.
10. Cohen JD, Braver TS, Brown JW. 2002. Computational perspectives on dopamine function in prefrontal cortex. *Curr Opin Neurobiol* 12: 223–229.
11. Cohen MX, Frank M. 2008. Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* 199: 141–156.
12. Cohen MX, Heller AS, Ranganath C. 2005. Functional connectivity with anterior cingulate and orbitofrontal cortices during decision-making. *Brain Res Cogn Brain Res* 23: 61–70.
13. Cohen RA, Kaplan RF, Zuffante P, Moser DJ, Jenkins MA, Salloway S, Wilkinson H. 1999. Alteration of intention and self-initiated action associated with bilateral anterior cingulotomy. *J Neuropsychiatry Clin Neurosci* 11: 444–453.
14. Critchley HD, Mathias CJ, Josephs O, O'Doherty J, Zanini S, Dewar BK, Cipolotti L, Shallice T, Dolan RJ. 2003. Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain* 126: 2139–2152.
15. Deiber M-P, Honda M, Ibanez V, Sadato N, Hallett M. 1999. Mesial motor areas in self-initiated versus externally triggered movements examined with fMRI: effect of movement type and rate. *J Neurophysiol* 81: 3065–3077.
16. Devinsky O, Morrell MJ, Vogt BA. 1995. Contributions of anterior cingulate cortex to behaviour. *Brain* 118: 279–306.
17. Dosenbach NUF, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HSC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. 2006. A core system for the implementation of task sets. *Neuron* 50: 799–812.
18. Dum RP, Strick PL. 1993. Cingulate motor areas. In: *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook* (Vogt BA, Gabriel M, eds), pp 415–441. Boston: Birkhauser.
19. Duncan J, Owen AM. 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci* 23: 475–483.
20. Falkenstein M, Hohnsbein J, Hoorman J, Blanke L. 1990. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: *Psychophysiological Brain Research* (Brunia CHM, Gaillard AWK, Kok A, eds), pp 192–195. Tilburg, Netherlands: Tilburg University Press.
21. Fellows LK, Farah MJ. 2005. Is anterior cingulate cortex necessary for cognitive control? *Brain* 128: 788–796.
22. Gemba H, Sasaki K, Brooks VB. 1986. "Error" potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neurosci Lett* 70: 223–227.
23. Hadland KA, Rushworth MFS, Gaffan D, Passingham RE. 2003. The anterior cingulate and reward-guided selection of actions. *J Neurophysiol* 89: 1161–1164.
24. Hayden BY, Platt ML. 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. *J Neurosci* 30: 3339–3346.
25. Holroyd CB, Coles MGH. 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109: 679–709.
26. Holroyd CB, Nieuwenhuis S, Yeung N, Nystrom LE, Mars RB, Coles MGH, Cohen JD. 2004. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat Neurosci* 7: 497–498.
27. Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. 2006. Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9: 940–947.
28. Kerns JG, Cohen JD, MacDonald AW, 3rd, Cho RY, Stenger VA, Carter CS. 2004. Anterior cingulate conflict monitoring and adjustments in control. *Science* 303: 1023–1026.
29. Matsumoto M, Matsumoto K, Abe H, Tanaka K. 2007. Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10: 647–656.
30. McClure SM, Daw ND, Montague PR. 2003. A computational substrate for incentive salience. *Trends Neurosci* 26: 423–428.

31. Mesulam M. 1981. A cortical network for directed attention and unilateral neglect. *Ann Neurol* 10: 309–325.
32. Miller EK, Cohen JD. 2001. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24: 167–202.
33. Miltner WHR, Braun CH, Coles MGH. 1997. Event-related potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *J Cogn Neurosci* 9: 788–798.
34. Montague PR, Hyman SE, Cohen JD. 2004. Computational roles for dopamine in behavioral control. *Nature* 431: 760–767.
35. Morecraft RJ, van Hoesen GW. 1998. Convergence of limbic input to the cingulate motor cortex in the rhesus monkey. *Brain Res Bull* 45: 209–232.
36. Nakamura K, Roesch MR, Olson CR. 2005. Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *J Neurophysiol* 93: 884–908.
37. Nemeth G, Hegedus K, Molnar L. 1988. Akinetic mutism associated with bicingular lesions: clinicopathological and functional anatomical correlates. *Arch Psychiatry Clin Neurosci* 237: 218–222.
38. Niv Y. 2009. Reinforcement learning in the brain. *J Math Psychol* 53: 139–154.
39. O’Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304: 452–454.
40. Papez JW. 1937. A proposed mechanism of emotion. *Arch Neurol Psychiatry* 38: 725–743.
41. Paus T. 2001. Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci* 2: 417–424.
42. Peyron R, Laurent B, Garcia-Larrea L. 2000. Functional imaging of brain responses to pain. a review and meta-analysis. *Neurophysiol Clin* 30: 263–288.
43. Picard N, Strick PL. 1996. Motor areas of the medial wall: a review of their location and functional activation. *Cereb Cortex* 6: 342–353.
44. Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. 2004. The role of medial frontal cortex in cognitive control. *Science* 306: 443–447.
45. Rushworth MFS, Behrens TE, Rudebeck PH, Walton ME. 2007. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn Sci* 11: 168–176.
46. Schultz W. 2002. Getting formal with dopamine and reward. *Neuron* 36: 241–263.
47. Shidara M, Richmond BJ. 2002. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* 31: 1709–1711.
48. Shima K, Aya K, Mushiake H, Inase M, Aizawa H, Tanji J. 1991. Two movement-related foci in primate cingulate cortex observed in signal-triggered and self-paced forelimb movements. *J Neurophysiol* 65: 188–202.
49. Shima K, Tanji J. 1998. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282: 1335–1338.
50. Stuss DT, Alexander MP, Shallice T, Picton TW, Binns MA, Macdonald R, Borowiec A, Katz DI. 2005. Multiple frontal systems controlling response speed. *Neuropsychologia* 43: 396–417.
51. Sutton RS, Barto AG. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
52. Tucker DM, Luu P, Pribram KH. 1995. Social and emotional self-regulation. *Ann N Y Acad Sci* 769: 213–239.
53. Turken AU, Swick D. 1999. Response selection in the human anterior cingulate cortex. *Nat Neurosci* 2: 920–924.
54. Walton ME, Bannerman DM, Alterescu K, Rushworth MF. 2003. Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J Neurosci* 23: 6475–6479.
55. Walton ME, Devlin JT, Rushworth MF. 2004. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat Neurosci* 7: 1259–1265.



56. Walton ME, Kennerley SW, Bannerman DM, Phillips PEM, Rushworth MFS. 2006. Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw* 19: 1302–1314.
57. Williams SM, Goldman-Rakic PS. 1993. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb Cortex* 3: 199–222.
58. Williams ZM, Bush G, Rauch SL, Cosgrove GR, Eskandar EN. 2004. Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat Neurosci* 7: 1370–1375.
59. Yeung N. 2004. Relating cognitive and affective theories of the error-related negativity. In: *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring* (Ullsperger M, Falkenstein M, eds), pp 63–70. Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.
60. Yeung N, Holroyd CB, Cohen JD. 2005. ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cereb Cortex* 15: 535–544.



# **Neural Basis of Motivational and Cognitive Control**

edited by

Rogier B. Mars, Jérôme Sallet,  
Matthew F. S. Rushworth, and Nick Yeung

# **Neural Basis of Motivational and Cognitive Control**

**edited by Rogier B. Mars, Jérôme Sallet, Matthew F. S. Rushworth, and  
Nick Yeung**

**The MIT Press  
Cambridge, Massachusetts  
London, England**

© 2011 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email [special\\_sales@mitpress.mit.edu](mailto:special_sales@mitpress.mit.edu)

This book was set in Times Roman by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Neural basis of motivational and cognitive control / edited by Rogier B. Mars . . . [et al.].  
p. ; cm.

Includes bibliographical references and index.

ISBN 978-0-262-01643-8 (hardcover : alk. paper)

1. Motivation (Psychology)—Physiological aspects. 2. Cognition—Physiological aspects. 3. Frontal lobes. I. Mars, Rogier B.

[DNLM: 1. Motivation—physiology. 2. Cognition—physiology. 3. Frontal Lobe—physiology.

QP 409]

QP409.N48 2012

612.8'233—dc22

2011010089

10 9 8 7 6 5 4 3 2 1