Prefrontal Cortex in Control: Broadening the Scope to Identify Mechanisms

William H. Alexander[1], Joshua W. Brown[2], Anne G. E. Collins[3],
Benjamin Y. Hayden[4], and Eliana Vassena[5]

Abstract

■ Sometime in the past two decades, neuroimaging and behavioral research converged on pFC as an important locus of cognitive control and decision-making, and that seems to be the last thing anyone has agreed on since. Every year sees an increase in the number of roles and functions attributed to distinct subregions within pFC, roles that may explain behavior and neural activity in one context but might fail to generalize across the many behaviors in which each region is implicated. Emblematic of this ongoing proliferation of functions is dorsal ACC (dACC). Novel tasks that activate dACC are followed by novel interpretations of dACC function, and each new interpretation adds to the number of functionally specific processes contained within the region. This state of affairs, a recurrent and persistent behavior followed by an illusory and transient relief, can be likened to behavioral pathology. In *Journal of Cognitive Neuroscience*, 29:10 we collect contributed articles that seek to move the conversation beyond specific functions of subregions of pFC, focusing instead on general roles that support pFC involvement in a wide variety of behaviors and across a variety of experimental paradigms. ■

In the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5), obsessive–compulsive disorder (OCD) is diagnosed based on the presence of obsessions, compulsions, or both (American Psychiatric Association & DSM-5 Task Force, 2013). Obsession is defined as a recurrent, persistent thought that may be intrusive and unwanted and that an individual attempts to suppress through an alternate thought or action—a compulsion. Compulsions are signified by repetitive behaviors undertaken to prevent or reduce anxiety or distress. The DSM further specifies that obsessions and compulsions are time-consuming and cannot be attributed to the physiological properties of a substance.

Considering these criteria and noting that the authors are not clinicians, it may be reasonably proposed that the field of cognitive neuroscience is experiencing a form of OCD in which the recurrent, intrusive thought concerns how best to explain cognitive control and decision-making as a product of the function of dorsal ACC (dACC). This thought manifestly produces distress, leading to ongoing debate and disagreement regarding the role of dACC across an array of paradigms and subfields of cognitive neuroscience. To prevent or reduce this distress, time-consuming behaviors, such as the umpteenth replication of the Eriksen flanker task or the proposal of a new computational model describing the role of dACC in a particularly constrained setting, are performed. The relief provided by such acts is ultimately temporary as new variations on the same idea arise, and the resulting unease must be ameliorated by additional behavioral performances.

Perhaps coincidentally, dACC, in addition to a vast range of other effects, is implicated in the neuroetiology of OCD in individuals (Fitzgerald et al., 2005). In extreme cases, the symptoms of OCD may be alleviated by cingulotomy: the complete or partial removal of the cingulate (Dougherty et al., 2002). This procedure is invasive in the extreme and generally only invoked when all other forms of treatment have been exhausted. Considering the pernicious nature of the *idée fixe* in our case—dACC activity is ubiquitously observed in studies of brain activity, even when it is not the target system (Alexander & Brown, 2015b)—and the increasingly drastic steps taken to resolve the discomfort—the number of scientific articles referring to dACC continues to increase on a yearly basis (Gage, Parikh, & Marzullo, 2008), one begins to wonder whether similarly drastic measures are called for. Does cognitive neuroscience need a cingulotomy, and what does that even mean?

Certainly, it is not feasible to remove all considerations of dACC from attempts to understand brain function and behavior—that the region is implicated in diverse modes of behavior and cognition is without question. Instead, we propose that the emphasis on examining differences of behavior and brain function in the context of behaviorally narrow control and decision-making tasks may be misplaced, resulting in a persistent focus on the functional

[1]Ghent University, [2]Indiana University, [3]University of California, Berkeley, [4]University of Rochester, [5]Radboud University Nijmegen

specialization of subregions of pFC rather than providing insight into general mechanisms and the architecture of control across pFC. Rather than being the principal generator of control signals in the brain, dACC may be one especially obvious center for control that nonetheless recruits a distributed neural network. Rather than calculating functionally specific signals that depend on the nature of a particular task, dACC may compute general signals that apply broadly over various experimental paradigms. Rather than investigating relatively simple control tasks to infer the nature of specific brain regions, more complex tasks may be required to derive more general principles by which the brain might operate. Instead of removing cingulate, then, the articles contributed to the articles in 29:10 remove the focus on identifying the function of particular regions of pFC, including dACC, and illustrate how a broader theoretical and empirical approach may inform our understanding.

Eisenreich, Akaishi, and Hayden (2017) look at this issue from a broad perspective. In their review, they discuss the idea that control is a broadly distributed phenomenon—and contrast that idea with more conventional modular control systems (Eisenreich et al., 2017 review). As a way of introducing this idea, they discuss several cases of distributed control systems from nature—but not from brains. For example, colonies of ants can make speed–accuracy trade-offs in their collective decision-making procedures (Franks, Dornhaus, Fitzsimmons, & Stevens, 2003). When making these trade-offs, there is no special control-sensitive ant that influences all the others. Instead, each ant has its own set of strategies and responses to environmental parameters. The speed–accuracy trade-off, like the decision itself, is an emergent phenomenon of the rules the group follows. They introduce some key principles of distributed control systems and then use these ideas as a lens through which to view the function of the dACC. Their focus is especially on reconciling the ideas of control with the known physiological properties of the structure (Heilbronner & Hayden, 2016). Ultimately, these ideas support the idea that executive control may reside in the dACC—but not exclusively there. Instead, it may be a more general property of brain systems that lies between the input and output ends of decision-making. These ideas align with recent work suggesting that economic decision-making—another function associated with dACC—may also be implemented in a distributed manner (Hunt & Hayden, 2017).

Brown and Alexander (2017) approach the question of cingulate function by taking one of the current controversies in the literature head on. They explore the dialectic of the foraging theory of dACC (Kolling et al., 2012) versus the theories of choice difficulty and expected value of control (Shenhav et al., 2014, 2016). From this dialectic, the authors derive a synthesis (Alexander & Brown, 2015b) by which their PRO model can be extended to account for what seemed to be contradictory findings. They propose a reinterpretation of choice difficulty signals as

reflecting surprise—when the choice is most difficult, then one of the options is always very likely to have been chosen (e.g., 50% likely), but it does not get chosen. The consistent surprise averaged across trials leads to maximal activity when the choice is most difficult and the option that will be chosen is most uncertain.

With respect to foraging value, the authors simulate foraging as a choice to overcome a default response of engaging with the current decision options. The decision to forage therefore requires effort. This can be exerted by the dACC in two ways, one of which may be characterized as proactive; and one, reactive. The proactive signals are trained by experiences in which the model fails to forage, despite the high value of foraging, and therefore experiences negative consequences. This in turn trains control signals to drive foraging under similar conditions in the future. Likewise, a surprising failure trains the PRO model surprise signals to temporarily suppress the actions that led to surprising failure for the next few trials. This allows the model to simulate a number of control effects.

Moreover, the Brown and Alexander (2017) article highlights a theme—that there is no single scalar control signal generated by the dACC. Instead, the control signals are rich and specific. They are rich because there are multiple kinds of control signals: proactive, reactive, and consisting of many different kinds of particular control signals that are specific to particular control requirements. They are specific in that they may be activated by one kind of context, or one kind of surprise, but not by another. Indeed, some control signals may compensate for the loss of others. In the end, Brown and Alexander argue that it may be simplistic to argue for a single control function in the dACC and that we should consider the richness of the signals.

Vassena, Deraeve, and Alexander (2017) adopt an empirical perspective, translating existing computational models of pFC function to the domain of motivation and effort-based decision-making and task performance. Wide empirical evidence assigns a key role to medial pFC (including dACC) in computing predictions regarding stimulus–outcome associations and updating these predictions according to prediction error signals (discrepancy between predicted and actual outcome; Jahn, Nee, Alexander, & Brown, 2014; Vassena, Krebs, Silvetti, Fias, & Verguts, 2014). The PRO model (Alexander & Brown, 2011) formalizes this in a computational framework, able to account for an extensive amount of empirical data. However, this framework in its current version cannot accommodate effects of motivation and effort-based behavior that are also reliably found in dACC (Botvinick & Braver, 2015; Vassena, Silvetti, et al., 2014). In this article, the authors reframe behavioral problems involving effort allocation in terms of prediction and prediction error and test the validity of the original PRO model in this new context. Model-based simulations reveal behavioral patterns consistent with participants' behavior and generate

novel predictions regarding dACC activity in this context. Furthermore, the same framework is used to suggest how behavior and brain activity in psychiatric disorders characterized by motivational impairments, such as depression, may be altered. Novel model-based behavioral and neural predictions are generated, to be tested in clinical populations. This approach might shed light on the neural mechanism underlying motivational impairments.

More recent computational work by Alexander and Brown (2015a) extended the explanatory scope of the PRO model using a hierarchical structure, still based on the principles of prediction and prediction error. The resulting hierarchical error representation (HER) model includes dorsolateral pFC (DLPFC) and accounts for empirical effects measured in this area. However, this framework has also not been translated to the domain of effort-based behavior, while situations that require cognitive effort typically result in increased DLPFC activity (Vassena, Silvetti, et al., 2014). In this article, the authors discuss the application of the HER model in effort-based behavior. Novel behavioral predictions are generated that suggest that order of presentation of motivational information influences DLPFC coding and, when manipulated, might affect decision-making as well as performance accuracy. Furthermore, the authors discuss the relevance of these predictions for clinical contexts. Overall, this article underlines the importance of testing theoretical frameworks beyond the scope they were conceived for, which comes with two main advantages: testing the validity and generalizability of a theory of brain function and providing theory-based qualitatively and quantitatively precise predictions for further empirical testing.

Alexander, Vassena, Deraeve, and Langford (2017) take the central proposal of the PRO model that dACC computes a multidimensional prediction and error signal, as a general role of cingulate, and examine how elaborating on this role may provide a foundation by which the organization and function of additional regions of the brain might be understood. In the previous two contributions, the focus on particular and unique functions of dACC in specific tasks is discarded in favor of examining how the principal signals of prediction and prediction error may be broadly applied over various forms of control and across different experimental contexts, thus removing the need for additional, functionally specific modules that must be added to the list of dACC functions. Here, Alexander et al. describe how prediction and error signals generated by dACC provide critical constraints on how the region might interact with additional areas in the brain underlying behavior and how these constraints informed the development of additional computational models describing the function of DLPFC in supporting sophisticated cognitive behaviors (Alexander & Brown, 2015a). Together, these models provide a uniquely comprehensive account of the function and organization of large regions within pFC at multiple levels of description.

The authors suggest how the PRO and HER models provide further constraints on future integrative models of pFC based on the computation and manipulation of prediction error representations. The proposed framework includes dACC as one region among many involved in processing quantities related to error and thus removes the cingulate as the focus of control and decision-making.

Collins's (2017) contribution moves away from the specific role of ACC to look more broadly at the contributions of fronto-BG networks to learning and how executive functions contribute to flexible hierarchical reinforcement learning. Previous work showed that the hierarchical organization of lateral and medial pFC plays an important role in using reward prediction errors to update representations at multiple abstraction levels (Alexander & Brown, 2015a; Donoso, Collins, & Koechlin, 2014; Collins & Koechlin, 2012), affording simultaneous learning of simple stimulus–action contingencies and more complex context–rule associations (Collins & Frank, 2013; Frank & Badre, 2012). This ability often allows us to simplify an existing problem by finding structure in the environment that allows information to be generalized. However, we also showed that humans engage in structure learning proactively, even in environments where this does not provide a learning benefit (Werchan, Collins, Frank, & Amso, 2015; Collins & Frank, 2012, 2013).

Here, Collins extends this finding to show that humans' tendency to build structure is not only present when such involvement of executive functions is unhelpful but that it carries heavy behavioral costs. Indeed, participants exhibited significantly slower RTs and worse performance when learning a problem that could be hierarchically structured, compared with a problem of matched difficulty that could not be structured. Thus, as predicted by computational models, structure learning occurred in the absence of any benefit to learning and in the presence of a strong behavioral cost. Furthermore, repeated exposure to this lack of benefit and cost did not decrease participants' tendency to create structure. Collins argues that this highlights the strength of our prior study for structure learning, putatively reflecting the fact that our environment is overwhelmingly structured, and thus affords opportunities for generalizing knowledge when we do structure our learning into hierarchical rules. This benefit might be general enough that the structure of hierarchical fronto-BG networks reflects this bias (Collins & Frank, 2016), as might be inferred by the fact that 8-month-old infants exhibit similar structure learning (Werchan et al., 2015), dependent on pFC function (Werchan, Collins, Frank, & Amso, 2016). Thus, our ability to overcome initial lack of benefit and behavioral costs would reflect an adaptive process in our environment where generalizing our behavior in different contexts is usually a good strategy.

Taken together, the contributed articles suggest how the obsessive–compulsive cycle that has led to a profusion of specific functional roles for subregions of pFC, and

the attendant debates in the literature, might be halted. First, although dACC may indeed be a region critical for cognitive control, it is by no means the only region involved in control. Hence, attempts to pack every control-related behavior into a single area may be counter-productive and overlook regions that may also be implicated in control. Second, roles proposed for the function of neural regions should look beyond the details of a single study or paradigm. pFC is implicated in a vast range of behaviors, from simple stimulus–response contingencies to complex decision-making involving multiple dimensions and stages; however, the diversity of function need not imply a diversity of mechanisms. General mechanisms that can account for the involvement of subregions of pFC in both simple and complex contexts are more parsimonious and elegant than accounts that invoke specialized operating modes. Finally, our understanding of the brain, both at the microlevel of individual neurons and the macrolevel of the behaviors to which they contribute, depends on the development of theoretical accounts of pFC function and its subregions that are capable of generating precise, testable predictions.

Reprint requests should be sent to Eliana Vassena, Donders Institute for Brain Cognition and Behaviour, Radboud Universiteit, Montessorilaan 3, Nijmegen, Gelderland, The Netherlands 6525 EN, or via e-mail: e.vassena@donders.ru.nl.

REFERENCES

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

Alexander, W. H., & Brown, J. W. (2015a). Hierarchical error representation: A computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation, 27,* 2354–2410.

Alexander, W. H., & Brown, J. W. (2015b). Reciprocal interactions of computational modeling and empirical investigation. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 321–338). New York: Springer.

Alexander, W. H., Vassena, E., Deraeve, J., & Langford, Z. D. (2017). Integrative modeling of prefrontal cortex. *Journal of Cognitive Neuroscience, 29,* 1674–1683.

American Psychiatric Association & DSM-5 Task Force. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5.* Washington, DC: American Psychiatric Association.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology, 66,* 83–113.

Brown, J. W., & Alexander, W. H. (2017). Foraging value, risk avoidance, and multiple control signals: How the anterior cingulate cortex controls value-based decision-making. *Journal of Cognitive Neuroscience, 29,* 1656–1673.

Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology, 10,* e1001293.

Collins, A. G. E. (2017). The cost of structure learning. *Journal of Cognitive Neuroscience, 29,* 1646–1655.

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience, 35,* 1024–1035.

Collins, A. G. E., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. *Psychological Review, 120,* 190–229.

Collins, A. G. E., & Frank, M. J. (2016). Motor demands constrain cognitive rule structures. *PLoS Computational Biology, 12,* e1004785.

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

Dougherty, D. D., Baer, L., Cosgrove, G. R., Cassem, E. H., Price, B. H., Nierenberg, A. A., et al. (2002). Prospective long-term follow-up of 44 patients who received cingulotomy for treatment-refractory obsessive–compulsive disorder. *American Journal of Psychiatry, 159,* 269–275.

Eisenreich, B. R., Akaishi, R., & Hayden, B. Y. (2017). Control without controllers: Towards a distributed neuroscience of executive control. *Journal of Cognitive Neuroscience, 29,* 1684–1698.

Fitzgerald, K. D., Welsh, R. C., Gehring, W. J., Abelson, J. L., Himle, J. A., Liberzon, I., et al. (2005). Error-related hyperactivity of the anterior cingulate cortex in obsessive–compulsive disorder. *Biological Psychiatry, 57,* 287–294.

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

Franks, N. R., Dornhaus, A., Fitzsimmons, J. P., & Stevens, M. (2003). Speed versus accuracy in collective decision making. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 270,* 2457–2463.

Gage, G. J., Parikh, H., & Marzullo, T. C. (2008). The cingulate cortex does everything. *Annals of Improbable Research, 14,* 12–15.

Heilbronner, S. R., & Hayden, B. Y. (2016). Dorsal anterior cingulate cortex: A bottom-up view. *Annual Review of Neuroscience, 39,* 149–170.

Hunt, L. T., & Hayden, B. Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews Neuroscience, 18,* 172–182.

Jahn, A., Nee, D. E., Alexander, W. H., & Brown, J. W. (2014). Distinct regions of anterior cingulate cortex signal prediction and outcome evaluation. *Neuroimage, 95,* 80–89.

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science, 336,* 95–98.

Lieberman, M. D., & Eisenberger, N. I. (2015). The dorsal anterior cingulate cortex is selective for pain: Results from large-scale reverse inference. *Proceedings of the National Academy of Sciences, U.S.A., 112,* 15250–15255.

Shenhav, A., Straccia, M. A., Botvinick, M. M., & Cohen, J. D. (2016). Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cognitive, Affective & Behavioral Neuroscience, 16,* 1127–1139.

Shenhav, A., Straccia, M. A., Cohen, J. D., & Botvinick, M. M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature Neuroscience, 17,* 1249–1254.

Vassena, E., Deraeve, J., & Alexander, W. H. (2017). Predicting motivation: Computational models of PFC can explain neural coding of motivation and effort-based decision-making in health and disease. *Journal of Cognitive Neuroscience, 29,* 1633–1645.

Vassena, E., Krebs, R. M., Silvetti, M., Fias, W., & Verguts, T. (2014). Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia, 59,* 112–123.

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

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2015). 8-Month-old infants spontaneously learn and generalize hierarchical rules. *Psychological Science, 26,* 805–815.

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2016). Role of prefrontal cortex in learning and generalizing hierarchical rules in 8-month-old infants. *Journal of Neuroscience, 36,* 10314–10322.