
Faculty of Social Sciences

Faculty Publications

Computational Models of Anterior Cingulate Cortex: At the Crossroads between Prediction and Effort

Eliana Vassena, Clay B. Holroyd, and William H. Alexander

June 2017

© 2017 Vassena et al. This is an open access article distributed under the terms of the Creative Commons Attribution License. <http://creativecommons.org/licenses/by/4.0>

This article was originally published at:

<https://doi.org/10.3389/fnins.2017.00316>

Citation for this paper:

Vassena, E.; Holroyd, C.B.; & Alexander, W.H. (2017). Computational models of anterior cingulate cortex: At the crossroads between prediction and effort. *Frontiers in Neuroscience*, 11, article 316. doi: 10.3389/fnins.2017.00316



Computational Models of Anterior Cingulate Cortex: At the Crossroads between Prediction and Effort

Eliana Vassena^{1,2*}, *Clay B. Holroyd*³ and *William H. Alexander*²

¹ Donders Center for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands, ² Department of Experimental Psychology, Ghent University, Ghent, Belgium, ³ Department of Psychology, University of Victoria, Victoria, BC, Canada

In the last two decades the anterior cingulate cortex (ACC) has become one of the most investigated areas of the brain. Extensive neuroimaging evidence suggests countless functions for this region, ranging from conflict and error coding, to social cognition, pain and effortful control. In response to this burgeoning amount of data, a proliferation of computational models has tried to characterize the neurocognitive architecture of ACC. Early seminal models provided a computational explanation for a relatively circumscribed set of empirical findings, mainly accounting for EEG and fMRI evidence. More recent models have focused on ACC's contribution to effortful control. In parallel to these developments, several proposals attempted to explain within a single computational framework a wider variety of empirical findings that span different cognitive processes and experimental modalities. Here we critically evaluate these modeling attempts, highlighting the continued need to reconcile the array of disparate ACC observations within a coherent, unifying framework.

Keywords: anterior cingulate cortex (ACC), effort, prediction error, computational models of ACC, computational modeling, effortful control

OPEN ACCESS

Edited by:

Stefano Rozzi,
University of Parma, Italy

Reviewed by:

James J. Bonaiuto,
University College London,
United Kingdom
Mehdi Khamassi,
Université Pierre et Marie Curie,
France

*Correspondence:

Eliana Vassena
e.vassena@donders.ru.nl

Specialty section:

This article was submitted to
Decision Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 15 March 2017

Accepted: 18 May 2017

Published: 06 June 2017

Citation:

Vassena E, Holroyd CB and
Alexander WH (2017) Computational
Models of Anterior Cingulate Cortex:
At the Crossroads between Prediction
and Effort. *Front. Neurosci.* 11:316.
doi: 10.3389/fnins.2017.00316

INTRODUCTION

Humans and other animals continually adapt their behavior in response to a rapidly changing environment, which requires speed and flexibility in evaluating environmental feedback. Research over the past two decades has identified anterior cingulate cortex (ACC) as a major neural hub for these computations (Rushworth et al., 2011), but the empirical evidence spans a wide variety of cognitive, affective and social functions (Bush et al., 2000; Nee et al., 2011; Gasquoine, 2013).

ACC is implicated in a lengthy list of processes (Shackman et al., 2011), including error detection (Gehring et al., 1993), conflict monitoring (Barch et al., 2001; van Veen and Carter, 2002), response selection (Holroyd and Coles, 2002), error likelihood (Brown and Braver, 2005), attention and task preparation (Luks et al., 2002; Aarts et al., 2008; Aarts and Roelofs, 2010), integration of outcome uncertainty and action values (Khamassi et al., 2015), reward prediction and prediction errors (Jessup et al., 2010; Silvetti et al., 2013; Vassena et al., 2014a), reward prediction errors experienced by others (Apps and Ramnani, 2014; Apps et al., 2016), prediction of effort required by the task (Vassena et al., 2014b; Chong et al., 2017), and perception of pain (Vogt, 2005; Fuchs et al., 2014). An automated machine-learning technique applied to a large data base also found

ACC involvement in most tasks (Yarkoni et al., 2011). The diversity of contexts in which ACC activity has been observed has led to the ironic conclusion that ACC is involved in everything (Ebitz and Hayden, 2016), and that the search for a “unimodal” characterization may be unobtainable (Bush, 2009), a position that favors the possibility of a multitude of separate signals mapped to different areas in ACC (Bush et al., 2000; Kolling et al., 2016a,b). This ubiquitous ACC activation has stimulated a search for an overarching theoretical framework that can account for all of the data while complying with principles of parsimony, falsifiability, and neurobiological plausibility (Alexander and Brown, 2010b).

This paper reviews the history of computational models of ACC function, highlights the current state-of-the-art, and discusses future directions. First, we summarize early seminal models that addressed single phenomena or relatively circumscribed sets of findings. These attempts mainly aimed at explaining fMRI and EEG data related to conflict and prediction errors. Second, we describe more recent models that account for growing evidence that ACC is implicated in effortful control (Walton et al., 2003, 2007; Vassena et al., 2014b; Holroyd and Umemoto, 2016; Klein-Flügge et al., 2016). Third, we describe models that broaden the explanatory scope to include lesion data and single-cell recordings under a shared underlying computational principle, none of which are comprehensive (see **Table 1** for a schematic comparison across all models). Finally, we discuss current and future attempts to bridge the remaining gaps.

EARLY SEMINAL MODELS

The first computational accounts of ACC function underscored its involvement in different task settings and cognitive processes. Perhaps the most influential of these is the *conflict-monitoring model* (Botvinick et al., 2001), which identified ACC as a conflict monitor that increases in activation as a function of conflict between available response options. On this account, stimuli that are incompatible on two (or more) stimulus dimensions (such as word meaning and ink color in the Stroop task) can activate competing response channels (e.g., left and right button presses); conflict is defined as the multiple of the activity of these channels, signaling a need for increased top-down control. Although conflict-related activity has reliably been measured in ACC with fMRI and EEG (Botvinick et al., 1999; Yeung et al., 2004; Carter and van Veen, 2007; Roberts and Hall, 2008), findings in patients and non-human animal literature are controversial (Yeung, 2013). In particular, ACC lesions do not consistently impair the cognitive control adjustments that, according to the theory, should follow conflict detection (Swick and Jovanovic, 2002; Fellows and Farah, 2005; di Pellegrino et al., 2007; Sheth et al., 2012), and scant neurophysiological evidence from monkey single-cell recordings is highly debated (Nakamura et al., 2005; Cole et al., 2009; Ebitz and Platt, 2015). Subsequently, several groups reported neurophysiological and neuroimaging findings inconsistent with the conflict monitoring proposal (Amiez et al., 2006; Burle et al., 2008;

Woodward et al., 2008; Hyafil et al., 2009; Kouneiher et al., 2009).

Brown and Braver (2005) later proposed the *error likelihood model*. According to this account, ACC associates errors with the stimulus context in which they occur, providing a means to predict the context-dependent likelihood of error commission. This model provided an early overarching attempt that explained both error and conflict activity. Subsequent experiments verified critical aspects of the model: as predicted, stimulus features associated with higher likelihood of errors, but not with conflict, elicited greater ACC activity (Brown, 2009). Failures to replicate the error likelihood effect (Nieuwenhuis et al., 2007) have been attributed to individual differences in risk-aversion (with more risk-averse subjects showing larger error-likelihood effects, Brown and Braver, 2007, 2008). One important limitation of this proposal was the inability to simulate the effect of unexpected errors (errors committed in contexts with low error likelihood elicit greater activity than errors in high error likelihood contexts, Brown and Braver, 2005).

An account by Holroyd and Coles (2002, 2008) proposed that ACC acts as a “motor control filter” that decides which action policy should be selected for a particular task. On this view, the value of action policies are learned via reward prediction error signals carried to ACC by midbrain dopamine system. These signals are proposed to encode discrepancies between expected and actual rewards that underlie the production of the error-related negativity (ERN) component of the event-related potential (ERP). Notably, this reinforcement learning model of the ERN (*RL-ERN theory*) shifts the role of ACC from the evaluative domain (i.e., detecting response conflict or error likelihood) to the action selection domain, explaining how ACC signals affect behavior. Aspects of the RL-ERN theory have received strong empirical support (Walsh and Anderson, 2012; Sambrook and Goslin, 2015; Holroyd and Umemoto, 2016) and are compatible with evidence of ACC encoding action values in uncertain environments (e.g., during foraging, Kolling et al., 2012, 2016b). However, the explanatory scope of this proposal was mainly limited to EEG data, and is not easily translated to fMRI (cf. Holroyd et al., 2004; Nieuwenhuis et al., 2005; Becker et al., 2014; Ferdinand and Opitz, 2014). Furthermore, this model proposed a general role for ACC as a motor control filter, describing a high-level hierarchical mechanism for action selection, but did not make specific predictions about how reward and error signals regulate behavior.

Finally, Behrens et al. (2007) proposed that ACC is sensitive to the *volatility* of environmental outcomes. This proposal holds ACC responsible for detecting how rapidly reward contingencies change over time. The model provides a mechanism by which organisms can flexibly adapt their learning rate (i.e., the speed at which current knowledge of the world is updated with new information). The volatility measure computed by ACC is used to adjust this learning rate in order to optimize subsequent decision-making. Furthermore, according to the authors the volatility signal is dissociable from prediction errors signals, thus implicitly postulating co-existence of difference signals within ACC. One limitation of this proposal is that while the volatility signal is proposed to influence learning rate at the time of feedback,

TABLE 1 | Schematic comparison of all discussed models.

Model	Publication	Model type	Effects	Data type	Species
EARLY MODELS					
Conflict monitoring	Botvinick et al., 2001; Yeung et al., 2004	Connectionist	Conflict, errors	fMRI, EEG	Humans
Error likelihood	Brown and Braver, 2005	Rate-coded neurons	Conflict, errors	fMRI	Humans
Motor control filter	Holroyd and Coles, 2002, 2008	Reinforcement learning	Errors, prediction, reward prediction error	EEG	Humans
Volatility	Behrens et al., 2007	Bayesian	Volatility	fMRI	Humans
RECENT EFFORT MODELS					
Choice difficulty	Botvinick, 2007; Shenhav et al., 2014	Connectionist	Choice difficulty in decision-making	fMRI	Humans
Adaptive effort allocation	Verguts et al., 2015	Reinforcement learning	Physical and cognitive effort and cost-benefit trade off	fMRI, single-cell, lesion	Humans, rodents, monkeys
Expected value of control	Shenhav et al., 2013	Conceptual	Cognitive control and cost-benefit trade off in decision-making	fMRI	Humans
Synchronization by oscillations	Verguts, 2017a	Rate-coded neurons	Cognitive control driven by theta oscillations	(Intra-cranial) EEG	Humans
RECENT UNIFYING MODELS					
PRO	Alexander and Brown, 2011	Rate-coded neurons; reinforcement learning	Prediction and prediction error, conflict, error, pain	fMRI, EEG, single-cell	Humans, rodents, monkeys
PRO-Effort	Vassena et al., in press	= PRO	= PRO + effort	= PRO	Humans
PRO-Control	Brown and Alexander, 2017	= PRO	= PRO + foraging, choice difficulty	= PRO + lesion	Humans, monkeys
RVPM	Silvetti et al., 2011	Rate-coded neurons; reinforcement learning	Reward prediction and prediction error, conflict, error, volatility	fMRI, EEG, single-cell, lesion	Humans, monkeys
HRL-ACC	Holroyd and McClure, 2015	Reinforcement learning	Effort, task switching, hierarchical behaviors	Lesion	Rodents
RNN-ACC	Shahnazian and Holroyd, 2017	Connectionist	Distributed coding of extended action sequences, conflict, prediction errors	Single-cell, fMRI, EEG	Rodents, humans
HER	Alexander and Brown, 2015	Predictive coding	as PRO, + dlPFC	fMRI, EEG, lesion, single-cell	Humans, monkeys
ACC-LPFC	Khamassi et al., 2011	Rate-coded neurons; reinforcement learning	Reward prediction error, salience, exploration-exploitation trade-off	Single cell, fMRI	Humans, monkeys

For every model (early models, recent models effort models, and recent unifying models), the table provides first publication reference, model type (implementation), data type (data that the model was conceived to explain), and species to which this data belongs. Within model type, connectionist refers to the Parallel Distributed Processing approach (McClelland et al., 1987); reinforcement-learning refers to the approach described in Sutton and Barto (1998); rate-coded neurons refers to the approach described in Dayan and Abbot (2001).

this model does not address how ACC contributes to action selection.

Although these computational models provided the first steps toward a mechanistic understanding of ACC function, they share a limitation in having been mainly conceived to explain one type of experimental data. This aspect is perhaps particularly problematic when based on fMRI data: BOLD measurements provide an indirect and possibly biased means for assessing neuronal activity (Logothetis, 2002, 2008), and further, increases in activity in ACC may reflect synaptic activity from projecting regions rather than firing by local neurons in ACC.

RECENT MODELS RELATED TO EFFORT AND DIFFICULTY

Recent findings have drawn attention to the central role of ACC in control processes requiring effort. Generally, ACC seems to

be more active when subjects prepare for difficult or effortful tasks, even in absence of error, conflict, and choice (Mulert et al., 2005; Aarts et al., 2008; Vassena et al., 2014b). ACC lesions impair decisions that evaluate trade-offs between effort expenditure and reward value in non-human animals (Walton et al., 2002, 2003, 2007), and are associated with motivational impairments and apathy in humans (e.g., Devinsky et al., 1995; Holroyd and Umemoto, 2016).

Botvinick (2007) anticipated this line of research with a simple model proposing that the conflict signal may drive effort avoidance, thus linking the conflict monitoring theory with decision-making. This idea was later extended to the proposal that ACC codes for *choice difficulty* (i.e., conflict between choice options), based on the observation that BOLD-fMRI ACC activity during decision-making negatively correlates with value differences between available options (Pochon et al., 2008; Shenhav et al., 2014). While not explicitly modeling effort, this

proposal is one of the first to point to a role of ACC in coding difficulty.

The *adaptive effort allocation model* by Verguts et al. (2015) addresses the role of ACC in effortful control explicitly, accounting for the empirical finding that expectation of effort in absence of choice or conflict is associated with increased ACC activity (Vassena et al., 2014b). On this account, ACC units implement a “boosting” mechanism, biasing behavior toward more effortful options when it is worth it (i.e., when they are predicted to procure a large enough reward). The model predicts that boosting increases the signal-to-noise ratio in task-related brain areas, thereby ensuring successful task completion. Although carrying a cost that increases linearly as a function of task difficulty, the boosting mechanism ensures sufficient cognitive or physical effort is deployed to obtain the reward at stake. This model effectively implements an “effort module” that can influence other cortical regions as appropriate to the task at hand.

In line with the adaptive effort allocation model, the *expected value of control* framework (EVC) proposes that ACC computes the value of exerting cognitive control (Shenhav et al., 2013), integrating a variety of signals (including conflict, reward, costs, effort, choice difficulty, and so on) in order to determine the degree of control worth applying over task performance. The EVC theory hypothesizes a role for ACC in calculating the “value of control” based on a combination of multiple different signals, some of which have also been ascribed to ACC; the EVC framework thus postulates calculations of the expected value of control as an additional role of ACC rather than a single mechanism explaining the variety of signals observed within the region. This framework has recently contributed to a lively and ongoing debate on the neural mechanisms of *foraging*, with a series of experiments inspired by the ecology literature pointing to a critical role for ACC: according to this proposal, ACC activity reflects the relative value of foraging, i.e., of leaving a known environment in order to explore a new environmental patch, which is associated with higher uncertainty but also potentially higher rewards (Kolling et al., 2012, 2016b). Whereas Shenhav and colleagues suggested that such a foraging signal reflects choice difficulty in a foraging context (Shenhav et al., 2014), Kolling and colleagues proposed that choice difficulty and foraging are dissociable and coded in segregated sub-regions of ACC. Overall, this proposal would appear to be consistent with Shenhav and colleague’s EVC account, assuming an additional signal in ACC that codes for foraging and inputs to the calculation of the value of exerting control. While providing a theoretical framework with potentially wide explanatory scope, the EVC theory has yet to be translated to a detailed computational framework with testable (viz. falsifiable) predictions. Although developed by a different group of investigators, one could consider the Adaptive Effort Allocation model (Verguts et al., 2015) as a possible computational instantiation of the EVC framework.

Another recent model has specified a role for ACC in synchronizing neural oscillations across brain areas (Verguts, 2017a). This complementary perspective suggests that ACC

exerts top-down control with bursts of activity in the theta frequency band that synchronize task-related areas throughout cortex, resulting in more efficient cortical communication. This proposal aligns with evidence that theta oscillations originating in ACC reflect effortful control (Holroyd, 2016; Holroyd and Umemoto, 2016), and is unique in describing the role of ACC in cognitive control in terms of synchronizing neural oscillations.

Overall, these recent proposals complement their predecessors in so far as they account for effort and control effects, which were only partially explained by previous models. However, the explanatory scope remains limited to the domain of effort-based behavior, and mostly neglects to account for data across different experimental models within a single framework.

RECENT UNIFYING TRENDS

In parallel with the development of ACC models of effort processing, several computational models have tried to account for a wider array of empirical findings within a single, overarching theoretical framework. In particular, whereas many previous theories focused on explaining functional neuroimaging data, recent models have widened their scope to include lesion and neurophysiological data. An early step in this direction was the *Predicted Response Outcome* (PRO) model (Alexander and Brown, 2011), which assigns to ACC the role of a stimulus-action-outcome predictor. In this account, ACC predicts the likelihood of upcoming actions and outcomes based on stimulus input from the environment; the predictions are then compared with actually experienced outcomes to produce a prediction error when an expected outcome and an actual outcome do not match. This error signal informs the prediction units, updating the predictions for future reference. Therefore, ACC is mainly sensitive to the (un)predictability of outcomes, regardless of their affective valence, as well as the unpredicted non-occurrence of these outcomes. Under this simple computational principle, the PRO model is able to simulate a wide variety of empirical findings, including sensitivity of ACC to conflict, errors, reward prediction errors and pain, for both neuroimaging and single-cell data (Alexander and Brown, 2014; Jahn et al., 2014, 2016). Simply put, the proposed mechanism for monitoring the (un)predictability of outcomes and their deviations from expectation provides a unifying framework for understanding such a diverse array of empirical findings.

A similar framework, the *Reward Value and Prediction Model* (RVPM), independently proposed by Silvetti et al. (2011), implements a comparable prediction mechanism in ACC, with one major difference: according to the RVPM, ACC predicts the value of future outcomes, but only when reward is at stake. By contrast, according to the PRO model, ACC predicts the likelihood of any outcomes (even for events that have no intrinsic value; Alexander and Brown, 2014). Like the PRO model, the RVPM successfully explains a wide range of data based on the principle of prediction and prediction errors (Silvetti et al., 2013). However, while both models have the potential to be extended to

the domain of effortful control, a translation to the effort domain has been recently proposed only for the PRO model (Vassena et al., in press).

While neuroimaging studies have suggested a critical role for ACC in performance monitoring and control, neuropsychological reports on patients with ACC lesions do not show the dramatic impairments one would expect based on these findings (Yeung, 2013). Partly for this reason, Holroyd and Yeung (2012) proposed that ACC selects and maintains extended, goal-directed action sequences, rather than instigate moment-to-moment changes in behavior following conflicts and errors. On this view, ACC impairments would impact complex high-level goal-pursuit rather than individual, low-level actions. These ideas were implemented by Holroyd and McClure (2015) in a 3-level, hierarchical reinforcement learning (HRL) model of rodent behavior (*HRL-ACC model*). Here a mid-level module associated with a caudal region of ACC selects tasks for execution and, in line with the literature on ACC involvement in effortful control, applies a control signal that attenuates effort-related costs incurred by a low-level action selection mechanism, ensuring that the task is completed successfully. Likewise, a high-level module located in rostral ACC selects the “meta-task” for execution and applies a control signal over caudal ACC that attenuates effortful costs incurred in task switching, facilitating shifts between different task strategies. The control levels are regulated according to tonic dopamine levels in cortex, which are assumed to code for average reward rate. On this view, the role of ACC in foraging relates to increased control by caudal ACC for exploiting a current patch, vs. increased control by rostral ACC for switching to alternative patches. Crucially, the model accounts for the effects of ACC lesions on rodent behavior and is broadly compatible with comparable observations in humans with ACC lesions.

Although the HRL-ACC model implements some of the key aspects of the proposal that ACC is responsible for selecting and motivating the execution of extended behaviors (Holroyd and Yeung, 2012), it does not directly account for neuroimaging and single-cell evidence in ACC. To address this gap, Shahnazian and Holroyd (2017) simulated the role of ACC in the production of goal-directed action sequences using a recurrent neural network (RNN) approach that had been previously used to simulate the production of hierarchical action sequences (Botvinick and Plaut, 2004). This *RNN-ACC model* predicts each successive event in the sequence and, like the PRO (Alexander and Brown, 2011) and RVPM (Silvetti et al., 2011) models, generates prediction errors when the events are unexpected, as observed in functional neuroimaging and EEG data (Botvinick et al., 2001; Alexander and Brown, 2010a; Wessel et al., 2012). Further, unique to this model, the predictions are hypothesized to be encoded as highly distributed representations across ensembles of neurons in ACC, as observed in studies of non-human animals (e.g., Ma et al., 2014a,b). Nevertheless, although this model is inspired by the broader theoretical framework of the HRL-ACC theory (Holroyd and Yeung, 2012), it has yet to be integrated with the HRL-ACC model (Holroyd and McClure, 2015).

DISCUSSION

Modeling ACC function faces the challenge of accounting for a multitude of empirical findings within a single coherent framework. Although progress has been made over the last two decades in explaining a wider array of empirical findings associated with different experimental techniques (Verguts, 2017b), the debate about underlying computational principles remains lively. We see several outstanding conceptual issues that remain to be resolved, many of which lie at the crossroads between predictive mechanisms and effortful control.

A first issue is that effortful control has been addressed with dedicated mechanisms that are constrained in scope. For example, the *adaptive effort allocation model* (Verguts et al., 2015) does not account for ACC activity related to predictions and prediction errors. Although the proposed mechanism could be implemented with separate ACC units, with an RVPM or PRO-like module computing prediction error signals and a boosting module driving adaptive effort exertion, this potential integration remains speculative. Conversely, we have recently proposed a possible translation of the PRO model to the effort domain (Vassena et al., in press). This model explains effort-related effects in terms of outcome prediction and error signaling, where required effort is considered as an outcome of the choice to engage in the task. As such, the PRO model explains increased ACC activity following a choice to engage in an effortful task as deriving from the “surprise” of choosing a high-effort trial. The model reconciles effort-related neuroimaging data in ACC with the variety of findings already explained by the PRO model under the unifying principles of prediction and prediction error (Jahn et al., 2014, 2016).

The ACC-HRL theory is characterized by similar tensions between different sources of evidence (Holroyd and Yeung, 2012). On the one hand, the ACC-HRL model accounts for the effects of ACC damage on rodent behavior (Holroyd and McClure, 2015) and is compatible with both neuroimaging data related to effort and control, and electrophysiological evidence of ACC reward prediction errors (Holroyd and Umemoto, 2016; Umemoto et al., 2017). On the other hand, it does not explicitly address the conflict and surprise signals commonly observed in ACC, nor any ACC single-cell data in non-human animals. Conversely, the ACC-RNN model accounts for surprise and error signals while simultaneously describing single-cell activity as arising from distributed representations of ensembles of ACC cells while animals execute goal-directed action sequences (Shahnazian and Holroyd, 2017). However, this architecture does not fully exploit hierarchical representations, nor utilizes reward signals for regulating control levels. A natural next step would be to integrate these approaches in line with recent examples (e.g., Cooper et al., 2014), using a more biologically-realistic network that incorporates finer temporal dynamics into the unit activity (Sussillo, 2014).

A related issue concerns the neurobiological plausibility of the proposed accounts across species. Some proposals are based on (or at least compatible with) neurophysiological and lesion findings in non-human primates and/or rodents, while others are mainly based on human EEG or fMRI findings (see **Table 1**).

A comprehensive account of ACC function should bridge apparent inconsistencies across experimental modalities, linking the indirect results of neuroimaging with single-neuron activity (as attempted for example by the PRO, RVPM, and ACC-RNN models).

Second, an important challenge in modeling ACC function is to account not only for the monitoring processes in which ACC is involved (such as conflict and error detection), but also how these computations modulate subsequent behavior. A recent proposal extended the PRO model in this direction (*PRO-Control*, Brown and Alexander, 2017), suggesting how prediction and error signals computed in ACC can serve as the basis for proactive and reactive control. While the PRO-Control assigns the same computations to ACC as in previous iterations of the model (Alexander and Brown, 2011, 2014; Alexander et al., 2015), it is able to account for additional behavioral and imaging effects related to deploying control, including effects of foraging value and choice difficulty (cf. the controversy mentioned above, Kolling et al., 2016b; Shenhav et al., 2016). Another modeling approach that may explain how ACC signals drive behavior is the meta-learning perspective. For example, Verguts and Notebaert (2008, 2009) proposed that control may emerge as a result of Hebbian learning, while Khamassi et al. (2013) proposed that ACC monitoring function determines learning rate, thus informing exploration-exploitation trade-offs in other brain regions.

A further challenge is to account not only for the function of ACC during normal behavior, but also in circumstances in which behavior and ACC function is impaired, as in the case of lesions to ACC (e.g., Devinsky et al., 1995; Fellows and Farah, 2005; Kennerley et al., 2006; Walton and Mars, 2007; Camille et al., 2011; Tsuchida and Fellows, 2013) or in clinical disorders such as, e.g., substance dependence, depression, or obsessive-compulsive disorder (Rive et al., 2013; Gowin et al., 2014; Barch et al., 2016). Initial work has attempted to link existing computational accounts of ACC to behavioral dysfunction (Alexander et al., 2015; Holroyd and Umemoto, 2016; Vassena et al., in press), and additional efforts in this direction may further refine understanding of the role of ACC during both normal and abnormal behavior.

Finally, a larger objective is to develop theories that integrate ACC function into the broader network of brain areas involved in control and decision-making. While the interaction of ACC with additional brain regions, including dorsolateral prefrontal cortex (dlPFC) and basal ganglia, was incorporated into early models of ACC (Botvinick et al., 2001; Holroyd and Coles, 2002; Kerns et al., 2004), the nature of this interaction has remained mostly an issue of secondary concern. Although the basal ganglia are recognized as a hub for interaction of motor, cognitive and motivational processes, modeling efforts to describe interactions between these areas and ACC are surprisingly scarce (but see, e.g., Hikosaka and Isoda, 2010; Cockburn and Frank, 2011). The ACC-HRL theory takes a step in remediating this deficit by proposing how ACC interacts with dorsolateral prefrontal cortex, orbitofrontal

cortex, the striatum, and other brain areas (Holroyd and Yeung, 2012; Holroyd and McClure, 2015; Holroyd and Umemoto, 2016). Likewise, the *Hierarchical Error Representation model* (HER, Alexander and Brown, 2015) examines the interaction of ACC and dlPFC through a hierarchical predictive coding framework, which replicates the PRO model at hierarchical levels that map onto a putative rostrocaudal gradient of abstraction in prefrontal cortex (Badre and D'Esposito, 2007; Taren et al., 2011; Nee and D'Esposito, 2016), and that sub-serve different higher-order cognitive functions. The HER model explains the function of large regions of PFC as primarily concerning the computation, representation, and manipulation of quantities derived from prediction error. Khamassi et al. (2011) have also proposed a computational account that simulates cellular activity in both ACC and LPFC, predicting that feedback-related signals in ACC modulate exploration-exploitation trade-off in LPFC during decision-making. This proposal implements dopamine input to ACC not only during feedback-related reward prediction errors, but also at the occurrence of any salient event, reconciling previous proposals based on dopaminergic input (cf. Holroyd and Coles, 2002, 2008) with neurophysiological evidence that dopamine also responds to salient but non-rewarding events (Horvitz, 2000).

These modeling attempts highlight important objectives for future models of ACC. As a first goal, existing ACC models with relatively wide explanatory power should be extended to other domains, especially effortful control, that heretofore have been the target of more constrained models. As a second goal, models of ACC should be integrated into more comprehensive accounts that explain how the ACC interacts with other brain regions. The benefits of such an approach are two-fold. First, the wider scope of empirical data predicted by the models would provide means for their falsification. In much the same way that the failure to observe single neurons encoding conflict signals paved the way for new models that could account for single-unit activity as well as for the activity of neural ensembles, the possible failure of current models to account for effortful decision-making may point the direction toward even more comprehensive accounts. Second, in the event that existing models can be extended to account for motivational effects in ACC, they provide a basis for understanding interactions with additional brain regions, providing insights into the function of the brain beyond cingulate cortex alone.

AUTHOR CONTRIBUTIONS

EV, CBH, and WHA reviewed the literature, drafted the manuscript and provided critical comments.

FUNDING

This work was funded by H2020 Marie Skłodowska-Curie Actions, project PreMotive number 705630.

REFERENCES

- Aarts, E., and Roelofs, A. (2010). Attentional control in anterior cingulate cortex based on probabilistic cueing. *J. Cogn. Neurosci.* 23, 716–727. doi: 10.1162/jocn.2010.21435
- Aarts, E., Roelofs, A., and van Turenout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *J. Neurosci.* 28, 4671–4678. doi: 10.1523/JNEUROSCI.4400-07.2008
- Alexander, W. H., and Brown, J. W. (2010a). Competition between learned reward and error outcome predictions in anterior cingulate cortex. *Neuroimage* 49, 3210–3218. doi: 10.1016/j.neuroimage.2009.11.065
- Alexander, W. H., and Brown, J. W. (2010b). Computational models of performance monitoring and cognitive control. *Top. Cogn. Sci.* 2, 658–677. doi: 10.1111/j.1756-8765.2010.01085.x
- Alexander, W. H., and Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344. doi: 10.1038/nn.2921
- Alexander, W. H., and Brown, J. W. (2014). A general role for medial prefrontal cortex in event prediction. *Front. Comput. Neurosci.* 8:69. doi: 10.3389/fncom.2014.00069
- Alexander, W. H., and Brown, J. W. (2015). Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Comput.* 27, 2354–2410. doi: 10.1162/NECO_a_00779
- Alexander, W. H., Fukunaga, R., Finn, P., and Brown, J. W. (2015). Reward salience and risk aversion underlie differential ACC activity in substance dependence. *Neuroimage Clin.* 8, 59–71. doi: 10.1016/j.nicl.2015.02.025
- Amiez, C., Joseph, J. P., and Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex* 16, 1040–1055. doi: 10.1093/cercor/bhj046
- Apps, M. A., and Ramnani, N. (2014). The anterior cingulate gyrus signals the net value of others' rewards. *J. Neurosci.* 34, 6190–6200. doi: 10.1523/JNEUROSCI.2701-13.2014
- Apps, M. A., Rushworth, M. F., and Chang, S. W. (2016). The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron* 90, 692–707. doi: 10.1016/j.neuron.2016.04.018
- Badre, D., and D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J. Cogn. Neurosci.* 19, 2082–2099. doi: 10.1162/jocn.2007.19.12.2082
- Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., and Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb. Cortex* 11, 837–848. doi: 10.1093/cercor/11.9.837
- Barch, D. M., Pagliaccio, D., and Luking, K. (2016). Mechanisms underlying motivational deficits in psychopathology: similarities and differences in depression and schizophrenia. *Curr. Top. Behav. Neurosci.* 27, 411–449. doi: 10.1007/7854_2015_376
- Becker, M. P., Nitsch, A. M., Miltner, W. H., and Straube, T. (2014). A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *J. Neurosci.* 34, 3005–3012. doi: 10.1523/JNEUROSCI.3684-13.2014
- Behrens, T. E., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221. doi: 10.1038/nn1954
- Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7, 356–366. doi: 10.3758/CABN.7.4.356
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624
- Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., and Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402, 179–181. doi: 10.1038/46035
- Botvinick, M., and Plaut, D. C. (2004). Doing without schema hierarchies: a recurrent connectionist approach to normal and impaired routine sequential action. *Psychol. Rev.* 111, 395–429. doi: 10.1037/0033-295X.111.2.395
- Brown, J. W. (2009). Conflict effects without conflict in anterior cingulate cortex: multiple response effects and context specific representations. *Neuroimage* 47, 334–341. doi: 10.1016/j.neuroimage.2009.04.034
- Brown, J. W., and Alexander, W. H. (2017). Foraging value, risk avoidance, and multiple control signals: how the ACC controls value-based decision-making. *J. Cogn. Neurosci.* doi: 10.1162/jocn_a_01140. [Epub ahead of print].
- Brown, J. W., and Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 307, 1118–1121. doi: 10.1126/science.1105783
- Brown, J. W., and Braver, T. S. (2007). Risk prediction and aversion by anterior cingulate cortex. *Cogn. Affect. Behav. Neurosci.* 7, 266–277. doi: 10.3758/CABN.7.4.266
- Brown, J. W., and Braver, T. S. (2008). A computational model of risk, conflict, and individual difference effects in the anterior cingulate cortex. *Brain Res.* 1202, 99–108. doi: 10.1016/j.brainres.2007.06.080
- Burle, B., Roger, C., Allain, S., Vidal, F., and Hasbroucq, T. (2008). Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. *J. Cogn. Neurosci.* 20, 1637–1655. doi: 10.1162/jocn.2008.20.1110
- Bush, G. (2009). *Dorsal Anterior Midcingulate Cortex: Roles in Normal Cognition and Disruption in Attention-Deficit/Hyperactivity Disorder*. New York, NY: Oxford University Press.
- Bush, G., Luu, P., and Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222. doi: 10.1016/S1364-6613(00)01483-2
- Camille, N., Tsuchida, A., and Fellows, L. K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J. Neurosci.* 31, 15048–15052. doi: 10.1523/JNEUROSCI.3164-11.2011
- Carter, C. S., and van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cogn. Affect. Behav. Neurosci.* 7, 367–379. doi: 10.3758/CABN.7.4.367
- Chong, T. T., Apps, M., Giehl, K., Silence, A., Grima, L. L., and Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol.* 15:e1002598. doi: 10.1371/journal.pbio.1002598
- Cockburn, J., and Frank, M. J. (2011). “Reinforcement learning, conflict monitoring, and cognitive control: an integrative model of cingulate-striatal interactions and the ERN,” in *Neural Basis of Motivational and Cognitive Control*, ed R. B. Mars (Cambridge, MA: MIT Press), 311–331.
- Cole, M. W., Yeung, N., Freiwald, W. A., and Botvinick, M. (2009). Cingulate cortex: diverging data from humans and monkeys. *Trends Neurosci.* 32, 566–574. doi: 10.1016/j.tins.2009.07.001
- Cooper, R. P., Ruh, N., and Mareschal, D. (2014). The goal circuit model: a hierarchical multi-route model of the acquisition and control of routine sequential action in humans. *Cogn. Sci.* 38, 244–274. doi: 10.1111/cogs.12067
- Dayan, P., and Abbot, L. F., (2001). *Theoretical Neuroscience: Computational and Mathematical Model of Neural Systems*. Cambridge MA: MIT Press.
- Devinsky, O., Morrell, M. J., and Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain J. Neurol.* 118(Pt 1), 279–306. doi: 10.1093/brain/118.1.279
- di Pellegrino, G., Ciaramelli, E., and Ládavas, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *J. Cogn. Neurosci.* 19, 275–286. doi: 10.1162/jocn.2007.19.2.275
- Ebitz, R. B., and Hayden, B. Y. (2016). Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. *Nat. Neurosci.* 19, 1278–1279. doi: 10.1038/nn.4387
- Ebitz, R. B., and Platt, M. L. (2015). Neuronal activity in primate dorsal anterior cingulate cortex signals task conflict and predicts adjustments in pupil-linked arousal. *Neuron* 85, 628–640. doi: 10.1016/j.neuron.2014.12.053
- Fellows, L. K., and Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control? *Brain* 128, 788–796. doi: 10.1093/brain/awh405
- Ferdinand, N. K., and Opitz, B. (2014). Different aspects of performance feedback engage different brain areas: disentangling valence and expectancy in feedback processing. *Sci. Rep.* 4:5986. doi: 10.1038/srep05986
- Fuchs, P. N., Peng, Y. B., Boyette-Davis, J. A., and Uhelski, M. L. (2014). The anterior cingulate cortex and pain processing. *Front. Integr. Neurosci.* 8:35. doi: 10.3389/fmint.2014.00035
- Gasquoine, P. G. (2013). Localization of function in anterior cingulate cortex: from psychosurgery to functional neuroimaging. *Neurosci. Biobehav. Rev.* 37, 340–348. doi: 10.1016/j.neubiorev.2013.01.002

- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390. doi: 10.1111/j.1467-9280.1993.tb00586.x
- Gowin, J. L., Stewart, J. L., May, A. C., Ball, T. M., Wittmann, M., Tapert, S. F., et al. (2014). Altered cingulate and insular cortex activation during risk-taking in methamphetamine dependence: losses lose impact. *Addiction* 109, 237–247. doi: 10.1111/add.12354
- Hikosaka, O., and Isoda, M. (2010). Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends Cogn. Sci.* 14, 154–161. doi: 10.1016/j.tics.2010.01.006
- Holroyd, C. B. (2016). “The waste disposal problem of effortful control,” in *Motivation and Cognitive Control*, ed T. Braver (New York, NY: Psychology Press), 235–260.
- Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709. doi: 10.1037/0033-295X.109.4.679
- Holroyd, C. B., and Coles, M. G. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex* 44, 548–559. doi: 10.1016/j.cortex.2007.08.013
- Holroyd, C. B., and McClure, S. M. (2015). Hierarchical control over effortful behavior by rodent medial frontal cortex: a computational model. *Psychol. Rev.* 122, 54–83. doi: 10.1037/a0038339
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., et al. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat. Neurosci.* 7, 497–498. doi: 10.1038/nn1238
- Holroyd, C. B., and Umemoto, A. (2016). The research domain criteria framework: the case for anterior cingulate cortex. *Neurosci. Biobehav. Rev.* 71, 418–443. doi: 10.1016/j.neubiorev.2016.09.021
- Holroyd, C. B., and Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn. Sci.* 16, 122–128. doi: 10.1016/j.tics.2011.12.008
- Horvitz, J. C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656. doi: 10.1016/S0306-4522(00)00019-1
- Hyafil, A., Summerfield, C., and Koehlin, E. (2009). Two Mechanisms for task switching in the prefrontal cortex. *J. Neurosci.* 29, 5135–5142. doi: 10.1523/JNEUROSCI.2828-08.2009
- Jahn, A., Nee, D. E., Alexander, W. H., and Brown, J. W. (2014). Distinct regions of anterior cingulate cortex signal prediction and outcome evaluation. *Neuroimage* 95, 80–89. doi: 10.1016/j.neuroimage.2014.03.050
- Jahn, A., Nee, D. E., Alexander, W. H., and Brown, J. W. (2016). Distinct regions within medial prefrontal cortex process pain and cognition. *J. Neurosci.* 36, 12385–12392. doi: 10.1523/JNEUROSCI.2180-16.2016
- Jessup, R. K., Bussemeyer, J. R., and Brown, J. W. (2010). Error effects in anterior cingulate cortex reverse when error likelihood is high. *J. Neurosci.* 30, 3467–3472. doi: 10.1523/JNEUROSCI.4130-09.2010
- Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., and Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940–947. doi: 10.1038/nn1724
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., and Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026. doi: 10.1126/science.1089910
- Khamassi, M., Enel, P., Dominey, P. F., and Procyk, E. (2013). Medial prefrontal cortex and the adaptive regulation of reinforcement learning parameters. *Prog. Brain Res.* 202, 441–464. doi: 10.1016/B978-0-444-62604-2.00022-8
- Khamassi, M., Lallée, S., Enel, P., Procyk, E., and Dominey, P. F. (2011). Robot cognitive control with a neurophysiologically inspired reinforcement learning model. *Front. Neurobot.* 5:1. doi: 10.3389/fnbot.2011.00001
- Khamassi, M., Quilodran, R., Enel, P., Dominey, P. F., and Procyk, E. (2015). Behavioral regulation and the modulation of information coding in the lateral prefrontal and cingulate cortex. *Cereb. Cortex* 25, 3197–3218. doi: 10.1093/cercor/bhu114
- Klein-Flügge, M. C., Kennerley, S. W., Friston, K., and Bestmann, S. (2016). Neural signatures of value comparison in human Cingulate Cortex during decisions requiring an effort-reward trade-off. *J. Neurosci.* 36, 10002–10015. doi: 10.1523/JNEUROSCI.0292-16.2016
- Kolling, N., Behrens, T. E., Mars, R. B., and Rushworth, M. F. (2012). Neural mechanisms of foraging. *Science* 336, 95–98. doi: 10.1126/science.1216930
- Kolling, N., Behrens, T., Wittmann, M. K., and Rushworth, M. (2016a). Multiple signals in anterior cingulate cortex. *Curr. Opin. Neurobiol.* 37, 36–43. doi: 10.1016/j.conb.2015.12.007
- Kolling, N., Wittmann, M. K., Behrens, T. E., Boorman, E. D., Mars, R. B., and Rushworth, M. F. S. (2016b). Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.* 19, 1280–1285. doi: 10.1038/nn.4382
- Kouneiher, F., Charron, S., and Koehlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nat. Neurosci.* 12, 939–945. doi: 10.1038/nn.2321
- Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 357, 1003–1037. doi: 10.1098/rstb.2002.1114
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878. doi: 10.1038/nature06976
- Luks, T. L., Simpson, G. V., Feiwell, R. J., and Miller, W. L. (2002). Evidence for anterior cingulate cortex involvement in monitoring preparatory attentional set. *Neuroimage* 17, 792–802. doi: 10.1006/nimg.2002.1210
- Ma, L., Hyman, J. M., Lindsay, A. J., Phillips, A. G., and Seamans, J. K. (2014a). Differences in the emergent coding properties of cortical and striatal ensembles. *Nat. Neurosci.* 17, 1100–1106. doi: 10.1038/nn.3753
- Ma, L., Hyman, J. M., Phillips, A. G., and Seamans, J. K. (2014b). Tracking progress toward a goal in corticostriatal ensembles. *J. Neurosci.* 34, 2244–2253. doi: 10.1523/JNEUROSCI.3834-13.2014
- McClelland, J. L., Rumelhart, D. E., Group, P. R., and others (1987). *Parallel Distributed Processing*. Cambridge: MIT Press.
- Mulert, C., Menzinger, E., Leicht, G., Pogarell, O., and Hegerl, U. (2005). Evidence for a close relationship between conscious effort and anterior cingulate cortex activity. *Int. J. Psychophysiol.* 56, 65–80. doi: 10.1016/j.ijpsycho.2004.10.002
- Nakamura, K., Roesch, M. R., and Olson, C. R. (2005). Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *J. Neurophysiol.* 93, 884–908. doi: 10.1152/jn.00305.2004
- Nee, D. E., and D’Esposito, M. (2016). The hierarchical organization of the lateral prefrontal cortex. *ELife* 5:e12112. doi: 10.7554/eLife.12112
- Nee, D. E., Kastner, S., and Brown, J. W. (2011). Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex. *Neuroimage* 54, 528–540. doi: 10.1016/j.neuroimage.2010.08.027
- Nieuwenhuis, S., Schweizer, T. S., Mars, R. B., Botvinick, M. M., and Hajcak, G. (2007). Error-likelihood prediction in the medial frontal cortex: a critical evaluation. *Cereb. Cortex* 17, 1570–1581. doi: 10.1093/cercor/bhl068
- Nieuwenhuis, S., Slagter, H. A., von Geusau, N. J., Helsenfeld, D. J., and Holroyd, C. B. (2005). Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *Eur. J. Neurosci.* 21, 3161–3168. doi: 10.1111/j.1460-9568.2005.04152.x
- Pochon, J. B., Riis, J., Sanfey, A. G., Nystrom, L. E., and Cohen, J. D. (2008). Functional imaging of decision conflict. *J. Neurosci.* 28, 3468–3473. doi: 10.1523/JNEUROSCI.4195-07.2008
- Rive, M. M., van Rooijen, G., Veltman, D. J., Phillips, M. L., Schene, A. H., and Ruhé, H. G. (2013). Neural correlates of dysfunctional emotion regulation in major depressive disorder. A systematic review of neuroimaging studies. *Neurosci. Biobehav. Rev.* 37, 2529–2553. doi: 10.1016/j.neubiorev.2013.07.018
- Roberts, K. L., and Hall, D. A. (2008). Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory stroop tasks. *J. Cogn. Neurosci.* 20, 1063–1078. doi: 10.1162/jocn.2008.20074
- Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E., and Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron* 70, 1054–1069. doi: 10.1016/j.neuron.2011.05.014
- Sambrook, T. D., and Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychol. Bull.* 141, 213–235. doi: 10.1037/bul0000006
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., and Davidson, R. J. (2011). The integration of negative affect, pain, and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* 12, 154–167. doi: 10.1038/nrn2994
- Shahnazian, D., and Holroyd, C. B. (2017). Distributed representations of action sequences in anterior cingulate cortex: a recurrent neural network approach. *Psychon. Bull. Rev.* doi: 10.3758/s13423-017-1280-1. [Epub ahead of print].

- Shenhav, A., Botvinick, M. M., and Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79, 217–240. doi: 10.1016/j.neuron.2013.07.007
- Shenhav, A., Cohen, J. D., and Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* 19, 1286–1291. doi: 10.1038/nn.4384
- Shenhav, A., Straccia, M. A., Cohen, J. D., and Botvinick, M. M. (2014). Anterior Cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat. Neurosci.* 17, 1249–1254. doi: 10.1038/nn.3771
- Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., et al. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* 488, 218–221. doi: 10.1038/nature11239
- Silvetti, M., Seurinck, R., and Verguts, T. (2011). Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* 5:75. doi: 10.3389/fnhum.2011.00075
- Silvetti, M., Seurinck, R., and Verguts, T. (2013). Value and prediction error estimation account for volatility effects in ACC: a model-based fMRI study. *Cortex* 49, 1627–1635. doi: 10.1016/j.cortex.2012.05.008
- Sussillo, D. (2014). Neural circuits as computational dynamical systems. *Curr. Opin. Neurobiol.* 25, 156–163. doi: 10.1016/j.conb.2014.01.008
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press: Cambridge.
- Swick, D., and Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: neuropsychological evidence for topographic specificity. *Neuropsychologia* 40, 1240–1253. doi: 10.1016/S0028-3932(01)00226-3
- Taren, A. A., Venkatraman, V., and Huettel, S. A. (2011). A parallel functional topography between medial and lateral prefrontal cortex: evidence and implications for cognitive control. *J. Neurosci.* 31, 5026–5031. doi: 10.1523/JNEUROSCI.5762-10.2011
- Tsuchida, A., and Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex* 49, 1790–1800. doi: 10.1016/j.cortex.2012.10.014
- Umehoto, A., HajiHosseini, A., Yates, M. E., and Holroyd, C. B. (2017). Reward-based contextual learning supported by anterior cingulate cortex. *Cogn. Affect. Behav. Neurosci.* 17, 642–651. doi: 10.3758/s13415-017-0502-3
- van Veen, V., and Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiol. Behav.* 77, 477–482. doi: 10.1016/S0031-9384(02)00930-7
- Vassena, E., Deraeve, J., and Alexander, W. H. (in press). Predicting motivation: computational models of PFC can explain neural coding of motivation and effort-based decision-making in health and disease. *J. Cogn. Neurosci.*
- Vassena, E., Krebs, R. M., Silvetti, M., Fias, W., and Verguts, T. (2014a). Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia* 59, 112–123. doi: 10.1016/j.neuropsychologia.2014.04.019
- Vassena, E., Silvetti, M., Boehler, C. N., Achten, E., Fias, W., and Verguts, T. (2014b). Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS ONE* 9:e91008. doi: 10.1371/journal.pone.0091008
- Verguts, T. (2017a). Binding by random bursts: a computational model of cognitive control. *J. Cogn. Neurosci.* 29, 1103–1118. doi: 10.1162/jocn_a_01117
- Verguts, T. (2017b). “Computational models of cognitive control,” in *The Wiley Handbook of Cognitive Control*, ed T. Egner (Chichester, UK: John Wiley & Sons, Ltd.). doi: 10.1002/9781118920497.ch8
- Verguts, T., and Notebaert, W. (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychol. Rev.* 115, 518–525. doi: 10.1037/0033-295X.115.2.518
- Verguts, T., and Notebaert, W. (2009). Adaptation by binding: a learning account of cognitive control. *Trends Cogn. Sci.* 13, 252–257. doi: 10.1016/j.tics.2009.02.007
- Verguts, T., Vassena, E., and Silvetti, M. (2015). Adaptive effort investment in cognitive and physical tasks: a neurocomputational model. *Front. Behav. Neurosci.* 9:57. doi: 10.3389/fnbeh.2015.00057
- Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* 6, 533–544. doi: 10.1038/nrn1704
- Walsh, M. M., and Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci. Biobehav. Rev.* 36, 1870–1884. doi: 10.1016/j.neubiorev.2012.05.008
- Walton, M. E., Bannerman, D. M., Alterescu, K., and Rushworth, M. F. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J. Neurosci.* 23, 6475–6479.
- Walton, M. E., Bannerman, D. M., and Rushworth, M. F. (2002). The role of rat medial frontal cortex in effort-based decision making. *J. Neurosci.* 22, 10996–11003.
- Walton, M. E., and Mars, R. B. (2007). Probing human and monkey anterior cingulate cortex in variable environments. *Cogn. Affect. Behav. Neurosci.* 7, 413–422. doi: 10.3758/CABN.7.4.413
- Walton, M. E., Rudebeck, P. H., Bannerman, D. M., and Rushworth, M. F. (2007). Calculating the cost of acting in frontal cortex. *Ann. N. Y. Acad. Sci.* 1104, 340–356. doi: 10.1196/annals.1390.009
- Wessel, J. R., Danielmeier, C., Morton, J. B., and Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *J. Neurosci.* 32, 7528–7537. doi: 10.1523/JNEUROSCI.6352-11.2012
- Woodward, T. S., Metzack, P. D., Meier, B., and Holroyd, C. B. (2008). Anterior cingulate cortex signals the requirement to break inertia when switching tasks: a study of the bivalency effect. *Neuroimage* 40, 1311–1318. doi: 10.1016/j.neuroimage.2007.12.049
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., and Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8, 665–670. doi: 10.1038/nmeth.1635
- Yeung, N. (2013). “Conflict monitoring and cognitive control,” in *Oxford Handbook of Cognitive Neuroscience, Vol. 13442*, eds K. Ochsner and S. Kosslyn (Oxford: Oxford University Press), 275–299.
- Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959. doi: 10.1037/0033-295X.111.4.931

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Vassena, Holroyd and Alexander. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.