

# Predicting Motivation: Computational Models of PFC Can Explain Neural Coding of Motivation and Effort-based Decision-making in Health and Disease

Eliana Vassena<sup>1,2</sup>, James Deraeve<sup>2</sup>, and William H. Alexander<sup>2</sup>

## Abstract

■ Human behavior is strongly driven by the pursuit of rewards. In daily life, however, benefits mostly come at a cost, often requiring that effort be exerted to obtain potential benefits. Medial PFC (MPFC) and dorsolateral PFC (DLPFC) are frequently implicated in the expectation of effortful control, showing increased activity as a function of predicted task difficulty. Such activity partially overlaps with expectation of reward and has been observed both during decision-making and during task preparation. Recently, novel computational frameworks have been developed to explain activity in these regions during cognitive control, based on the principle of prediction and prediction error (predicted response–outcome [PRO] model [Alexander, W. H., & Brown, J. W. Medial prefrontal cortex as an action–outcome predictor. *Nature Neuroscience*, 14, 1338–1344, 2011], hierarchical error representation [HER] model [Alexander, W. H., & Brown, J. W. Hierarchical error representation: A computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation*, 27, 2354–2410, 2015]). Despite the broad explanatory power of these models, it is not clear whether they

can also accommodate effects related to the expectation of effort observed in MPFC and DLPFC. Here, we propose a translation of these computational frameworks to the domain of effort-based behavior. First, we discuss how the PRO model, based on prediction error, can explain effort-related activity in MPFC, by reframing effort-based behavior in a predictive context. We propose that MPFC activity reflects monitoring of motivationally relevant variables (such as effort and reward), by coding expectations and discrepancies from such expectations. Moreover, we derive behavioral and neural model-based predictions for healthy controls and clinical populations with impairments of motivation. Second, we illustrate the possible translation to effort-based behavior of the HER model, an extended version of PRO model based on hierarchical error prediction, developed to explain MPFC–DLPFC interactions. We derive behavioral predictions that describe how effort and reward information is coded in PFC and how changing the configuration of such environmental information might affect decision-making and task performance involving motivation. ■

## INTRODUCTION

Attaining a goal often requires commitment, implementation of a precise course of actions, and deployment of sufficient resources to reach successful completion. How humans fulfill this process is largely investigated in the field of cognitive neuroscience under the term of goal-directed behavior. A crucial underlying cognitive mechanism is prediction: the ability to evaluate the environment, formulate expectations about future events based on previous experiences, and, finally, compare such expectations with subsequent outcomes to update one's knowledge about the current state of the world. Furthermore, adaptive interaction with the environment entails predicting the impact of one's action and evaluating outcomes. The human PFC is the neural machinery that supports crucial mechanisms involved in goal-directed behavior (Miller & Cohen, 2001). In particular, the medial portion of PFC (MPFC, including dorsal ACC [dACC]) is implicated in

prediction and outcome evaluation (Jahn, Nee, Alexander, & Brown, 2014; Alexander & Brown, 2011), including situations in which the outcome carries value for the agent, such as a reward (Vassena, Krebs, Silvetti, Fias, & Verguts, 2014; Silvetti, Seurinck, & Verguts, 2011, 2013; Rushworth, Walton, Kennerley, & Bannerman, 2004). Notably, a variety of other functions have been attributed to MPFC (Vassena, Holroyd, & Alexander, 2017), including error monitoring (Holroyd, et al., 2004; van Veen, Holroyd, Cohen, Stenger, & Carter, 2004), conflict detection (Botvinick, Braver, Barch, Carter, & Cohen, 2001), pain and affect processing (Nee, Kastner, & Brown, 2011; Bush, Luu, & Posner, 2000), and value-based decision-making (Rushworth, Kolling, Sallet, & Mars, 2012; Rangel & Hare, 2010; Rushworth & Behrens, 2008). Recent computational work has explained this array of empirical findings under the unifying principle of prediction and prediction error. In this framework, MPFC formulates predictions tracking stimuli, actions, and outcomes and computes a signal termed prediction error, which scales with the discrepancy between predicted and actual outcomes (predicted

<sup>1</sup>Radboud University Nijmegen, <sup>2</sup>Ghent University

response–outcome [PRO] model; Alexander & Brown, 2011). This mechanism allows rapid prediction updating according to environmental feedback, be it an error, a painful stimulus, or a reward. An extended version of the same model, the hierarchical error representation (HER) model (Alexander & Brown 2015), expands the same computational principle in a hierarchical architecture, capturing more complex high-level cognitive processes involving the interaction of MPFC and dorsolateral PFC (DLPFC), typically associated with higher-level cognitive functions such as working memory and goal maintenance (Miller & Cohen, 2001).

The goal of this article is to explore the power of these computational accounts, in terms of generating novel neural and behavioral predictions for untested contexts and populations. These frameworks have proven useful across several fields of cognition, yet they have not been put to test in the field of effortful behavior and motivation. Goal-directed behavior generally involves competing factors, including the value of the prospective goals, how much effort one is willing to exert to attain the desired goal, and preparation for the necessary effortful performance (Botvinick & Braver, 2015; Westbrook & Braver, 2013, 2015). First, we will describe MPFC involvement in effort-based behavior. Then, we illustrate how the PRO model can be generalized to the domain of motivation. We propose that MPFC activity reflects monitoring of motivationally relevant variables such as reward and required effort, instead of coding an explicit cost-benefit or choice signal per se. We illustrate novel model-based simulations as well as theoretical predictions, which can be used to guide further empirical enquiry. We discuss how the PRO framework makes neural and behavioral predictions for clinical conditions in which motivation is impaired, such as depression and other psychiatric disorders (Treadway, Bossaller, Shelton, & Zald, 2012). Subsequently, we discuss the future directions in translating the HER model to the domain of motivation, extrapolating behavioral predictions. From such predictions, we derive implications for measuring and potentially training motivation-related cognitive mechanisms in clinical populations.

## **EFFORT-BASED DECISION-MAKING AND PERFORMANCE IN MPFC**

Experimental manipulation of effort in behavioral and neuroimaging experiments has yielded a wealth of findings in the past decade. Typically, effort is perceived as aversive (Kool, McGuire, Rosen, & Botvinick, 2010), yet humans decide to engage in it when doing so leads to a benefit, such as a reward. In the framework of decision-making and neuroeconomics, the net value of a potential reward is discounted (i.e., decreased) by the amount of effort required to obtain the reward (Apps, Grima, Manohar, & Husain, 2015; Nishiyama, 2014; Hartmann, Hager, Tobler, & Kaiser, 2013). This seems to hold across different types of

effort (Nishiyama, 2016) and guides decisions to engage one's resources in the task at hand (Kurzban, Duckworth, Kable, & Myers, 2013). Several studies in animals described the neural mechanisms underlying this cost-benefit evaluation, showing a pivotal role of MPFC in interaction with striatal and subcortical nuclei (Hosokawa, Kennerley, Sloan, & Wallis, 2013; Kennerley, Dahmubed, Lara, & Wallis, 2009; Walton et al., 2009; Rushworth & Behrens, 2008; Walton, Rudebeck, Bannerman, & Rushworth, 2007; Walton, Kennerley, Bannerman, Phillips, & Rushworth, 2006; Walton, Bannerman, Alterescu, & Rushworth, 2003; Walton, Bannerman, & Rushworth, 2002). In the last few years, a similar network has been characterized in humans. Lesions of MPFC may result in a condition known as akinetic mutism (Devinsky, Morrell, & Vogt, 1995), whereby patients show difficulties in initiating speech and movement, due not to an impairment of related systems but rather to the inability or "lack of will" to start it. Electrical stimulation of the same region seems to induce a general feeling of being more motivated and more willing to persevere in effortful endeavors (Parvizi, Rangarajan, Shirer, Desai, & Greicius, 2013, although this single-case study presents some methodological caveats). More recently, neuroimaging studies have shown involvement of the striatum and MPFC in effort–reward trade-off computations (Massar, Libedinsky, Weiyan, Huettel, & Chee, 2015; Engstrom, Landtblom, & Karlsson, 2013; Botvinick, Huffstetler, & McGuire, 2009; Crosson, Walton, O'Reilly, Behrens, & Rushworth, 2009; Mulert et al., 2008). Furthermore, expecting to perform a more cognitively challenging task is associated with increased activity in the striatum and MPFC, overlapping with activity induced by the prospect of a higher reward (Vassena, Silvetti, et al., 2014; Krebs, Boehler, Roberts, Song, & Woldorff, 2012).

These results suggest a crucial contribution of MPFC to effort-based behavior, hypothesized to compute the willingness to engage in the task at hand, given that, upon completion, a reward will be received. This principle has been defined in recent theories (Shenhav, Botvinick, & Cohen, 2013; Holroyd & Yeung, 2012) and formalized in a computational model wherein MPFC calculates the value of boosting certain actions over others, accordingly guiding behavior in cognitive and physical tasks (Verguts, Vassena, & Silvetti, 2015). Such computations are thought to influence decision-making (Treadway, Buckholz, et al., 2012), resource allocation, task preparation (Vassena, Silvetti, et al., 2014; Kurniawan, Guitart-Masip, Dayan, & Dolan, 2013), and response vigor (Kurniawan, Guitart-Masip, & Dolan, 2011), even at the lowest layers of the motor system (Vassena, Cobbaert, Andres, Fias, & Verguts, 2015).

In summary, growing evidence supports a pivotal role of MPFC in effort-based behavior. However, such empirical effects and theorizing efforts have so far failed to provide a precise computational characterization able to account for this line of evidence within other existing computational frameworks of MPFC function.

## The PRO Model

According to the PRO model, MPFC implements two core mechanisms: (1) learning to predict the outcome of responses generated in response to environmental stimuli and (2) signaling discrepancies between predictions and observations. Using these two primary signals as an index of MPFC activity, the PRO model has been shown to account for a range of effects observed in MPFC related to cognitive control and decision-making, including effects of error, conflict, and error likelihood. Critically, the PRO model explains these effects without reference to the underlying affective import: Feedback related to behavioral error is equivalent to feedback indicating correct behavior in the sense that both forms of feedback constitute an outcome that can be predicted on the basis of task-related stimuli. An open question, therefore, is how the PRO model might be extended to account for effects in which behavior is influenced not only by the likelihood of an event occurring but also by the value of that event (Figure 1).

### Translating the PRO Model to Effort-based Motivation

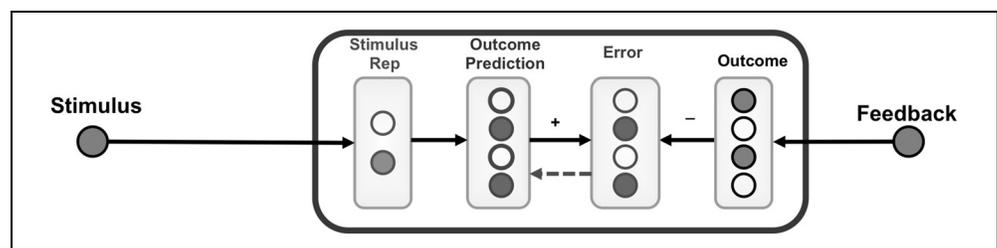
According to the PRO framework, MPFC activity encodes prediction error, resulting in increased activity for more unexpected (surprising) events. However, several studies investigating effort-based behavior report increased activity in the same region of MPFC when more effort needs to be invested (i.e., in the presence of a more demanding task; Vassena, Silvetti, et al., 2014; Krebs et al., 2012). To reconcile this apparent inconsistency, we hypothesize that MPFC contribution to effort-based decision-making parallels its role in cognitive control—MPFC predicts the amount of effort (as well as reward) associated with certain environmental cues and the likelihood of the choice to engage or not in the required behavior. In other words, we propose that MPFC monitors effort cues and decisions, with the same mechanisms used to monitor the occurrence of any other stimulus and response outcome.

Decisions regarding whether to engage in an effortful task carry multiple consequences. First, the choice to

perform an effortful task entails exerting actual effort to perform the task (regardless whether the task is performed successfully). In addition, performing a task carries with it the possibility of success, in which case the participant receives positive feedback, often in the form of monetary reward. Alternately, the participant may fail to perform the task successfully, in which case, negative feedback is provided indicating failure. In the simulations, such failure corresponds to not realizing the monetary reward, rather than losing money (although a loss condition could also be simulated as easily). In the framework of the PRO model, then, the outcomes predicted during choices regarding whether to engage in an effortful task are (1) the level of effort to be exerted and (2) the potential expected payoff. Furthermore, our implementation relies on two assumptions. First, greater effort is considered an aversive outcome, which generally tends to be avoided if possible (Kool et al., 2010). Second, as in the original model implementation (Alexander & Brown, 2011), outcomes can be more or less salient: Increasing levels of reward and effort correspond to increasing salience in the model. This assumption is based on the observation that effort is frequently perceived as aversive, plausibly generating increased arousal level.

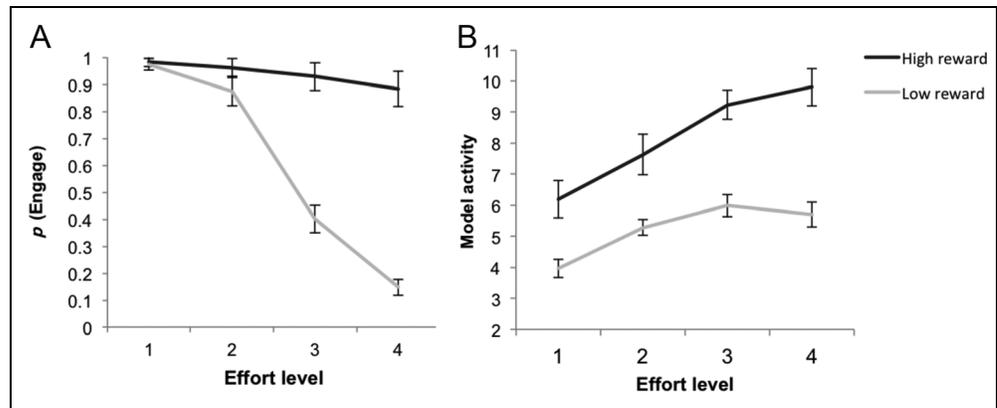
Under these assumptions, we simulated effort-based decision-making with the PRO model. The parameter set used here was the same used in simulations reported in earlier work (Alexander & Brown, 2011, 2014), with no additions to the architecture of the model, and therefore not specifically tailored to the current context (the code is available at [github.com/modelbrains/PRO\\_Effort](https://github.com/modelbrains/PRO_Effort)). One should note that, in this case, the PRO model is not performing the task itself but rather monitoring the choice of engaging in more or less effortful and rewarding trials (i.e., updating its predictions as a function of the experienced effort and reward, as if the task had been performed), as opposed to accepting a default option with a low reward value and no effort. In this formulation, MPFC activity reflects a monitoring signal, tracking the (un)predictability of motivationally relevant variables, instead of explicitly computing a cost-benefit trade-off or driving choice. Related work (cf. Brown & Alexander, 2017) suggests how signals generated by the PRO model

**Figure 1.** PRO model architecture (adapted from Alexander & Brown, 2011). The circles outside the box represent environmental input (stimulus and feedback). The circles inside the box represent units coding neural activity. Stimulus representations code environmental stimuli.



Depending on previous occurrence, certain stimuli predict certain outcomes, as coded in outcome prediction units. Outcome units code real environmental outcome (feedback). A comparison between outcome prediction units and outcome units results in an error signal (discrepancy between predicted and actual outcomes). This error signal feeds back into the outcome prediction unit, to update such predictions.

**Figure 2.** Model predictions. (A) Behavioral predictions. The y axis shows the probability of choosing to engage in a task. The x axis shows four different effort levels (varying parametrically from easy [Level 1] to hard [Level 4]). The gray line indicates a low reward upon successful completion. The black line indicates a high reward upon successful completion. The plot shows that, with a low reward, increasing task difficulty reduces the probability of engaging in the task, whereas, with a high reward, the model engages in the increasingly effortful task anyway. (B) Neural predictions. The y axis shows MPFC activity at the time of cue. The x axis shows the four effort levels. The gray line indicates low reward. The black line indicates high reward. The plot shows that model activity is overall higher when reward is high. Moreover, when reward is high, activity linearly increases as a function of increasing effort. When reward is low, model activity only increases up to Effort level 3.



may be deployed elsewhere in the brain to drive choice behavior. In addition, the adaptation of the PRO model to the context of effort-based decision-making suggests that the role of MPFC is primarily in monitoring the level of prospective reward and effort and does not necessarily drive decisions to engage in a proposed task or, once engaged, to maintain performance levels sufficient to realize successful completion of a task. Rather, according to additional applications of the PRO model in this issue (cf. Brown & Alexander, 2017), signals generated by MPFC are incorporated into decision processes occurring beyond cingulate itself. This interpretation of MPFC function is at odds with other models of MPFC function (Shenhav et al., 2013; Holroyd & Yeung, 2012) and provides a novel view of the role of MPFC in motivated behavior that may be the target of future research.

For simulations of the effort-based decision-making task, the model was presented a compound cue indicating the level of prospective reward (four levels) and level of prospective effort (four levels). Each reward level was modeled as a single input unit, as was each effort level, for a total of 16 unique compound stimuli reflecting combinations of effort and reward information. After a decision to perform the task, the model received feedback related to the level of reward received and the level of effort expended. The strength of the feedback signal for both effort and reward was set to the level indicated by the corresponding model input (1–4) multiplied by a constant (0.48 for reward, 0.55 for effort). The constant was selected by hand to reproduce the qualitative pattern of behavioral effects reported in the literature (Klein-Flügge, Kennerley, Saraiva, Penny, & Bestmann, 2015). After a decision not to engage in the effort task, feedback was set to one fourth of the lowest reward level.

Figure 2 shows the results of the simulations. The model behavior replicates qualitatively effort avoidance tendencies of human participants (see Figure 2A; for simplicity, only low and high reward conditions are dis-

played): As the required effort (task difficulty) increases, the probability of engaging in the task decreases (i.e., the prediction that one will choose to engage). Plausibly, the prospect of a high reward changes this pattern: When a higher reward is expected, the probability of engaging in more effortful tasks decreases only slightly relative to low-reward conditions. These behavioral predictions are consistent with empirical findings of previous studies (Klein-Flügge, Kennerley, Friston, & Bestmann, 2016; Apps et al., 2015). By looking at activity of the prediction units in the PRO model (Figure 2B), one can extrapolate quantitative predictions about expected activity in MPFC across different effort and reward conditions. According to the simulation, MPFC activity monotonically and linearly increases as a function of increased required effort (task difficulty) when reward prospect is high. However, when reward prospect is low, MPFC activity increases less steeply and only up to a certain degree of required effort, subsequently decreasing as the probability of engaging in trials with high demand for low reward drops. To our knowledge, this neural prediction is yet to be tested and could be investigated by recording MPFC activity during effort-based decision-making when difficulty is manipulated parametrically.

### Alternative Models of Effort-based Behavior

Other theoretical and computational models have been developed to account for MPFC contribution to effort-based behavior (Verguts et al., 2015; Shenhav et al., 2013). These models present one major difference with respect to the PRO framework: They explicitly operationalize effort as a cost to be computed in MPFC. As a result, although these models work well in predicting effort-based decisions and task performance, they do not provide an explicit computational characterization of how MPFC contributes to other empirical effects.

Verguts and colleagues (2015) assign MPFC a role in calculating the benefit of deploying effort in addition to signaling potential rewarding outcomes. Their adaptive effort investment model operationalizes effort explicitly by implementing what the authors call “boosting.” In this model, units representing MPFC activity compute the value of boosting, namely, exerting the effort needed to energize a more difficult action (be it a physical action or a cognitive task). Boosting, as in exerting effort, entails a cost. If the value of boosting outweighs the cost, the more effortful action will be selected. This results in the following predicted pattern of activity: Overall activity in MPFC should be higher for larger rewards, increase with increasing task difficulty as long as the reward is worth the effort, and drop for tasks too difficult to be solved. To our knowledge, this prediction still requires empirical testing. In line with this model, Shenhav and colleagues proposed the “expected value of control theory” (EVC; Shenhav et al., 2013). This theoretical framework assigns MPFC the role of computing the value of exerting control, by combining “component computations” estimating costs, benefits, and consequences associated with control signals (Shenhav, Cohen, & Botvinick, 2016). Input signals to such computations may include error, conflict, difficulty, and prediction error signals, which may originate outside MPFC.

From the implementation point of view, one should consider that the adaptive effort allocation and EVC frameworks rely on very different assumptions as compared with the PRO model. The first two place computation of the value of boosting (for cognitive or physical action in Verguts et al.) or exerting control (cognitive tasks in Shenhav et al.) in MPFC, whereas the PRO model places prediction and prediction error computations in MPFC. Moreover, whereas the PRO model postulates a shared underlying computational principle, adaptive effort allocation and EVC imply the coexistence of different computations (cost and value of boosting and prediction error for the first; separate cost, benefit, and consequence estimation for the second). However, further modeling work is required to extrapolate predictions, which may disentangle the models based on available empirical evidence.

The main advantage of the PRO model is parsimony: The same architecture explains effort-related effects as well as a wide variety of empirical effects previously measured in MPFC (ranging from prediction error, cognitive control, conflict, and etc.; Alexander & Brown, 2011). This is not the case for the adaptive effort investment model, which is specifically tailored for effort-based behavior and is therefore not applicable in other contexts, at least in its current implementation.

One limitation of the PRO model is that it does not perform the task and is not responsible for action selection: MPFC units compute predictions and compare them with outcomes. This assumes that the reward and cost trade-off computations and the choice to engage or

not in the task at hand are implemented elsewhere. Candidate areas could potentially be other subregions of PFC or possibly the BG and especially the striatum, shown in several studies in both humans and animals to contribute to effort-based decisions and task preparation (Bailey, Simpson, & Balsam, 2016; Vassena, Silveti, et al., 2014; Prévost, Pessiglione, Météreau, Cléry-Melin, & Dreher, 2010; Botvinick et al., 2009; Salamone, Correa, Farrar, & Mingote, 2007). One shortcoming common to both PRO and EVC/adaptive effort allocation frameworks is that they are agnostic about cost computation. Effort is plausibly defined as a function of task difficulty and higher effort equals higher cost. However, the nature and source of such cost signal is a topic of ongoing empirical and theoretical work (Holroyd, 2016; Kurzban et al., 2013).

### **Predictions and Implications for Clinical Populations**

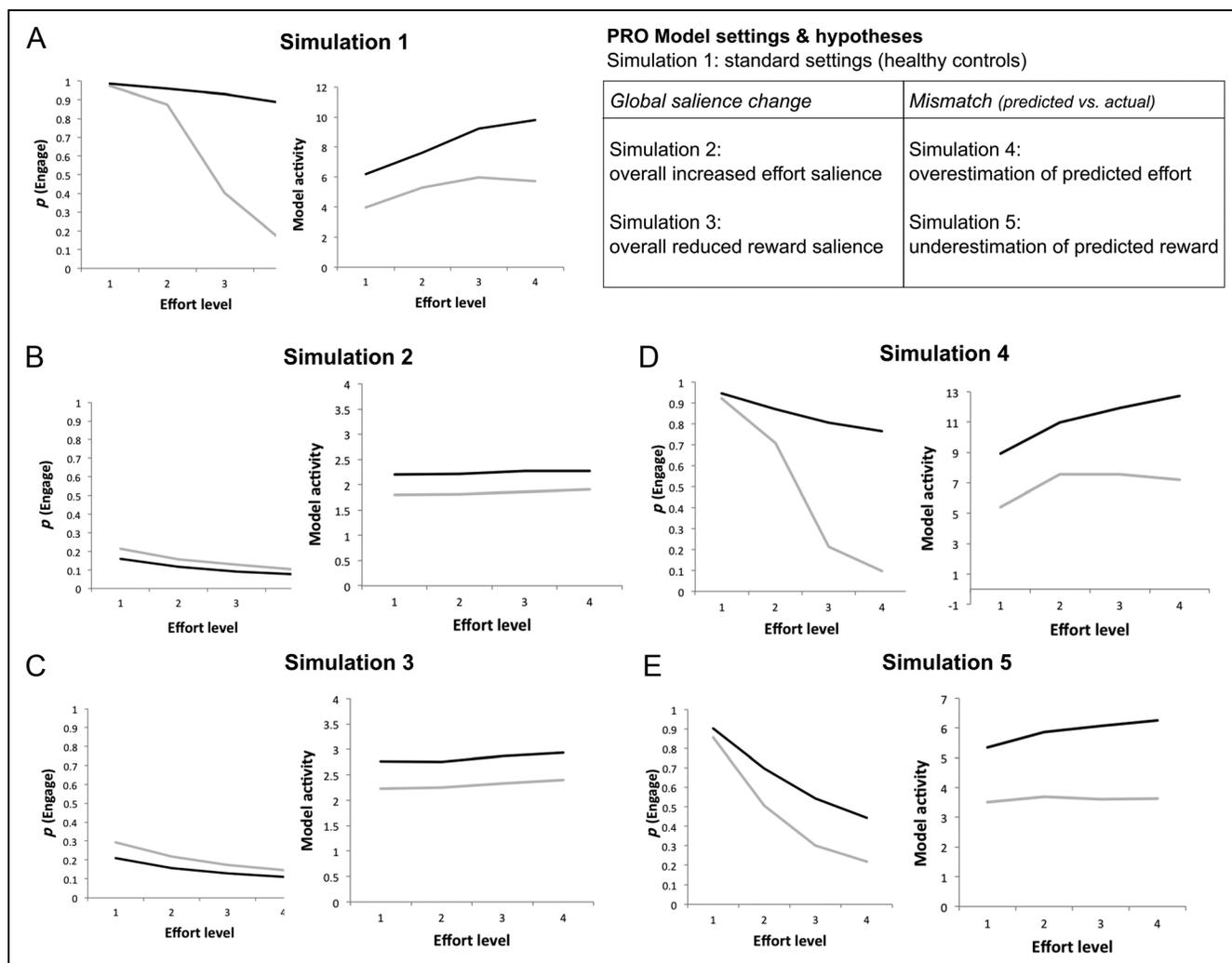
Adaptive decision-making and energization of behavior pose a challenge in several daily life situations. In a number of psychiatric conditions, these mechanisms are impaired. Recent studies showed that decision-making regarding whether to undertake an effortful task is altered in depression, bipolar disorder, and schizophrenia (Culbreth, Westbrook, & Barch, 2016; Hershenberg et al., 2016; McCarthy, Treadway, Bennett, & Blanchard, 2016; Silvia et al., 2016; Treadway, 2016; Barch, Treadway, & Schoen, 2014; Silvia, Nusbaum, Eddington, Beaty, & Kwapil, 2014; Yang et al., 2014; Treadway, Bossaller, et al., 2012). Symptoms often include reduced willingness to exert effort, although data across different pathologies or effort types do not always align. For example, both schizophrenic and depressed patients show reduced allocation of physical effort for higher rewards as compared with controls, whereas evidence concerning cognitive effort is mixed (Barch, Pagliaccio, & Luking, 2016). The same authors suggest a different underlying deficit for the two conditions: Depressed patients seem to show impaired hedonic processing, whereas schizophrenic patients tend to show impaired reinforcement learning and action selection. Moreover, effort-related deficits in schizophrenia point to an effort allocation deficit, rather than reduced effort expenditure per se (McCarthy et al., 2016; Treadway, Peterman, Zald, & Park, 2015), with patients performing less optimal decisions. Such a complex picture confirms alteration of effort-based decision-making in such clinical populations and calls for more precise quantitative frameworks, able to identify the mechanisms underlying different impairments.

Here, we use the PRO model, adapted as described above for modeling effort-related dynamics in healthy participants, to simulate the possible neuroetiology underlying clinical disorders, which could explain the behavioral symptoms measured in clinical samples. In the PRO model, outcome representation units may be modulated by salience (Alexander & Brown, 2011) suggesting that

compromised function in clinical populations may be a result of altered perception of salient events (Alexander, Fukunaga, Finn, & Brown, 2015). Model simulations and theoretical predictions are described in Figure 3.

These simulations use the basic architecture of the PRO model without modification as in Simulation 1. To simulate altered function during effort-based decision-making, we assume that clinical disorders entail alterations in the processing of information related to either reward or effort information. One possible alteration driving impairment in decision-making could be attributed to a global salience change: In some populations, the global salience of decision variables might be affected. Patients may be overly sensitive to the costs of engaging in a task (such as required effort, Simulation 2, Figure 3B) or have

reduced sensitivity to potential reward (Simulation 3, Figure 3C). To simulate these hypotheses, we multiply the effort level from simulation by a factor of 2 (Simulation 2) or the reward information by a factor of 0.5 (Simulation 3) to reflect increased effort salience or decreased reward salience. The results of these simulations show that increasing the salience of effort and reducing the salience of reward have similar effects in the model: The probability of engaging in a task is decreased over all levels of reward and effort. The pattern of MPFC activity predicted by the model is also severely attenuated relative to control simulation: Activity is slightly higher in the high-reward as compared with low-reward condition but does not seem to track effort as it did in the control simulation.



**Figure 3.** PRO model simulations of impaired motivation. In all plots, the y axis shows the probability of engaging in the task (left) and the model activity (right). The x axis shows four possible effort levels, parametrically increasing from easy (Level 1) to hard (Level 4). Gray lines indicate low reward upon task completion. Black lines indicate high reward upon task completion. (A) Simulation 1. Behavioral and neural predictions for healthy controls. The table on the right illustrates the hypotheses of possible impairments as modeled with the PRO model and relative explanation. We hypothesize two core possible mechanisms driving impairments in patients. The first is altered global salience, with either an overall increased effort salience (Simulation 2) or an overall increased reward salience (Simulation 3). The second is mismatch between predicted and actual outcomes, with either a possible overestimation of predicted effort (as compared with actual, Simulation 4) or a possible underestimation of the predicted reward (as opposed to actual, Simulation 5). (B–E) Simulations 2–5. Behavioral (choices) and neural (model activity) predictions under the different hypotheses.

Another possible alteration underlying the impairment in clinical populations might be a mismatch: Predictions made by the model regarding effort and reward levels might not correspond to veridical experience. The model may overestimate the level of effort required (Simulation 4) or underestimate the value of the reward on offer (Simulation 5). The inability to accurately estimate required effort and potential reward would generate a mismatch between prediction and outcome: Predicted effort could be overestimated, leading to abnormal effort avoidance, whereas mismatches between predicted and experienced rewards could lead to decreased motivation in performing the task. To simulate these hypotheses, effort-related feedback to the model was multiplied by a factor of 2 (Simulation 4), whereas the valence information used for updating top-down control weights (Alexander & Brown, 2011; Supplementary Figure 1) remained unchanged. The net effect is that the model's prediction of effort level exceeds the effort experienced by the model after choices to engage in an effortful task. In Simulation 5, reward-related feedback to the model was multiplied by 0.5 (whereas valence information was unchanged), with the interpretation that the level of predicted reward did not match the experienced level. Simulation results for effort mismatch (Figure 3D) and reward mismatch (Figure 3E) show that such mismatches in effort and reward prediction yield qualitatively different predictions regarding behavior: Overestimation of effort level leads to increased discounting of low reward offers, whereas behavior in high-reward conditions is relatively unaffected compared with control simulations. Conversely, underestimation of reward produces a general increase in discounting—both high- and low-reward conditions are discounted more heavily compared with control simulations.

To our knowledge, the hypothesized mechanisms (global salience impairment vs. predicted/actual mismatch) cannot be disentangled on the basis of existing data. Future experimental work to test this may incorporate a model-based fMRI experiment with patients performing an effort-based decision-making task. One could simulate model-based predictors of MPFC activity for each hypothesized mechanism on the basis of participants' actual performance. This would show which mechanisms better explain brain activity measured in MPFC (i.e., the one giving better fit between model activity and neural data). Empirical verification in clinical populations showing impairments of effort-based behavior would shed light on potential mechanisms underlying symptom origin and provide (in)validation for the PRO as plausible neurofunctional account of MPFC contribution to motivated behavior.

### Limitations and Critical Aspects

Translating the PRO framework to a motivational context allows explaining effort-based behavior under a working computational model of MPFC functioning without pos-

tulating an MPFC function dedicated to explicit cost computations. However, this translation leaves some critical aspects unanswered, open for future research. First, in our conceptualization, we do not distinguish between different types of effort costs, such as physical versus cognitive effort. Here, we only assume higher effort to be more salient and aversive, irrespective of its specific nature. Previous research comparing neural circuits involved in different effortful tasks (Schmidt, Lebreton, Cléry-Melin, Daunizeau, & Pessiglione, 2012) suggests that the type of effort determined the network involved in task execution, with motor regions implicated in a physical task as opposed to parietal regions implicated in a cognitive task. In both cases, the relevant network was more active in the high-effort condition. Moreover, a shared motivational hub was identified in the striatum, showing increased activity irrespective of effort type. In both animal and human research, the striatum has been implicated in cost-benefit trade-offs (Salamone et al., 2016; Westbrook & Braver, 2016; Vassena, Silvetti, et al., 2014; Botvinick et al., 2009) and is often coactive with MPFC. An intriguing possibility is that striatal dopamine-driven trade-off computations provide MPFC with the necessary cost signal regulating subsequent behavior, irrespective of effort type. These speculations should be investigated in further research.

Second, we do not include a mechanistic explanation of the aversive nature of effort. The neural origin of this computation is still debated in the literature. It has been proposed that perception of effort cost derives from its opportunity cost (i.e., engaging resources that could be utilized differently; Kurzban et al., 2013). A recent account hypothesizes effort cost to derive from accumulation of waste product at the neural level, resulting from using up neural resources (Holroyd, 2016). The model is currently agnostic to the origin of this signal, which we consider an avenue for future modeling and experimental work.

Third, we formulated effort-based behavior as a decision-making problem, where effort and reward are considered outcomes of the decision to engage in the task at hand. However, this does not account for monitoring ongoing effort exertion. Maintaining a certain level of vigor throughout a period (e.g., holding a grip) could be seen as the result of a series of decisions to keep engaging throughout the entire period, depending on (presumably striatal) cost and reward signals fed into MPFC. This intriguing idea should be addressed in future modeling and experimental work.

Fourth, we do not simulate MPFC activity variations within a trial. Theoretically, the PRO model states that MPFC continuously predicts stimulus-outcome associations (Alexander & Brown, 2014). This means that, at the beginning of a trial, before effort- or reward-related information being presented, the model would predict average outcomes (in the context of effort-based decision-making, these predictions would converge on the mean

reward and effort for the overall task). After cue presentation, this prediction would be updated when experiencing the actual effort, suggesting that MPFC activity should reflect the degree by which task-related cues on a specific trial diverge from the average experimental value. Preliminary evidence for such a computation is reported in a TMS study measuring motor-evoked potentials (Vassena et al., 2015), which showed that motor cortex excitability during cue presentation was related to prediction error in expected value (discrepancy between average expected value and value of the actual cue on the current trial, integrating a certain degree of required effort and potential reward). However, how this result speaks to MPFC contribution in the process remains to be investigated. Conversely, activity after the choice regarding whether to engage with an effortful task should vary inversely with the tendency of the participant to engage: Participants with a lower overall tendency to engage in effortful tasks should show increased MPFC activity after choices to engage, whereas participants with a strong tendency for engaging should show increased MPFC activity after choices not to engage. These theoretical predictions require empirical testing and, possibly, additional modeling work to specify them quantitatively. From the methodological point of view, one would need to collect fMRI data at a timescale with sufficient resolution to contrast MPFC activity at both cue and outcome or to use EEG–fMRI simultaneous recordings to localize MPFC electrophysiological signature.

## **EFFORT-BASED DECISION-MAKING AND PERFORMANCE IN DLPFC**

In the existing literature, the link between DLPFC and effort-based behavior is more implicit, although it clearly emerges from the number of high-level functions implicating this region. Traditionally, DLPFC is assigned a pivotal role in supporting working memory updating and maintenance (Curtis & D'Esposito, 2003; Miller & Cohen, 2001). In recent years, evidence has accumulated showing a crucial contribution of this region to executive functions, including goal maintenance and task set representation (Ridderinkhof, van den Wildenberg, Segalowitz, & Carter, 2004). Activity in DLPFC is associated with maintaining stimulus representations and strategies for optimal task performance. Although several frameworks have been proposed to explain DLPFC function (Koechlin, 2014; Badre, 2008), a mechanistic account of how such representations and strategies are formed and manipulated to guide goal-directed behavior is still lacking. How DLPFC interacts with MPFC prediction signals remains unclear. Several studies investigating motivation and task preparation report coactivation of DLPFC and MPFC (Chong et al., 2017; Botvinick & Braver, 2015; Engström, Karlsson, Landtblom, & Craig, 2014; Vassena, Silvetti, et al., 2014; Engstrom et al., 2013; Krebs et al., 2012; Rypma, Berger, & D'Esposito, 2002). Across these studies, DLPFC activity increases as a function of expected effort,

task load, and working memory demands. Recently, starting from the principles outlined in the PRO model, it has been proposed that the underlying computational mechanism of DLPFC might also rely on the prediction and prediction error (Alexander & Brown, 2015). These authors proposed an updated version of the PRO model, extended in a hierarchical architecture: the HER model (Figure 4).

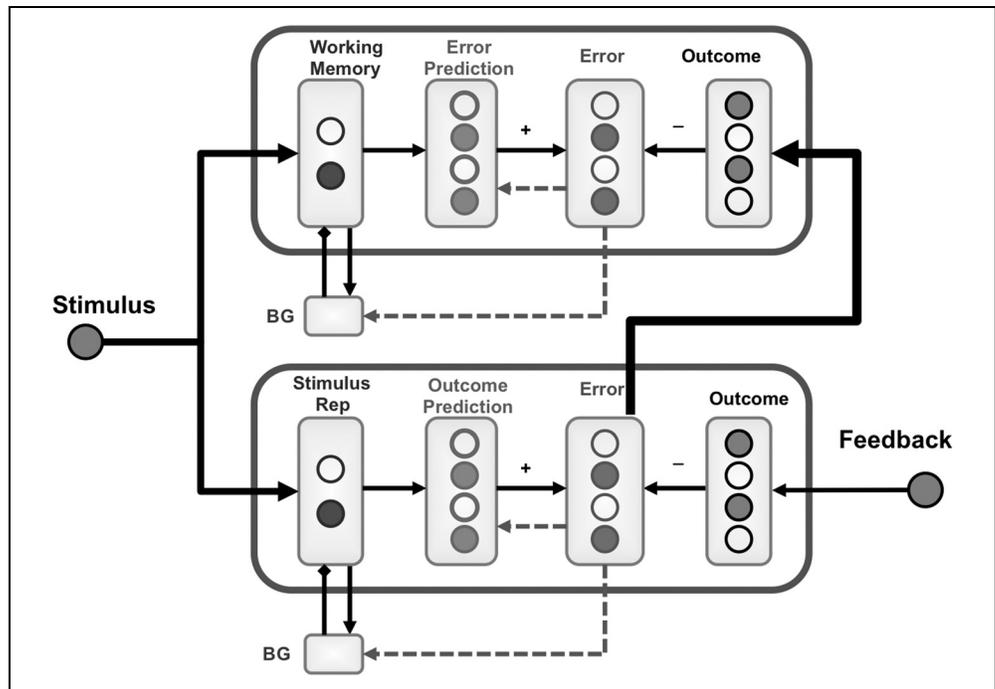
The HER model is composed of two or more hierarchical layers, and each layer replicates the functional form of the PRO model. The lowest layer receives input and feedback from the environment, updating predictions via prediction error, computed as the discrepancy between predicted and actual outcomes. In addition, the error signal also provides input to the layer above, where it is treated as a feedback signal; in other words, this higher layer learns predictions of the expected error of the lower layer, compares such prediction with the actual error signal, and updates the error prediction accordingly. This simple architecture provides a mechanistic account of how MPFC and DLPFC might interact, congruent with available empirical evidence (Alexander & Brown, 2015, 2016). The prediction error signal generated in MPFC not only results in an updated error prediction at the highest layer: This prediction is also linked to the environmental stimulus (or context), which was associated with the error. This results in a representation linking the error signal to the stimulus (or context) that preceded the error. In agreement with a substantial body of evidence, this model accounts for the primary role of MPFC in performance monitoring and error detection and for the role of DLPFC in maintaining task set representations providing context for MPFC function.

## **Future Directions: Translating the HER Model to Effort-based Behavior**

Despite its wide explanatory power, the HER framework has, to date, not been translated to the domain of motivation to accommodate for the aforementioned effort and task load effects observed in DLPFC. In the previous sections, we showed the potential of the PRO model to explain effort-related effects. Fundamentally, the HER model is an extension of the PRO model, which suggests that it might be well suited for a comparable translation to the effort domain.

The aim of the current section is twofold. First, we propose a theoretical explanation of how DLPFC–MPFC interaction in the context of the HER model could account for motivational effects observed in both regions. Second, we derive informal behavioral predictions from the HER model in its current formulation, which can be tested in both healthy and clinical populations to further challenge the validity of the model. One should note that such interpretations and predictions are highly speculative at this stage. The purpose of this section is to provide a series of directions and predictions to drive empirical investigation of DLPFC–MPFC contribution to effort-based behavior.

**Figure 4.** HER model architecture (adapted from Alexander & Brown, 2015). The circles outside the box represent environmental input (stimulus and feedback). The circles inside the box represent units coding neural activity. This figure shows a two-layer version of the HER model. Each layer replicates the architecture of the PRO model (cf. Figure 1): stimulus representation unit code for environmental stimuli, leading to prediction of a certain outcome (outcome prediction units); outcome units code for actual outcome; comparison between outcome prediction and actual outcome produces an error signal, coded in the error units. In the HER model, the activity of error units at the lowest layer scales with the discrepancy between predicted and actual outcomes (as in the PRO model); error signals update predictions in this layer but are also fed to the upper layer, where actual error signal is compared with the predicted error signal. The resulting second-level error signal is used to update predictions of the future error signal.



The HER model is built on the principle that error signals in MPFC are equivalent to other environmental feedback signals and are therefore subject to the same prediction and error processes. When an error signal is unexpected, the error prediction is updated. This error history is stored in DLPFC as error representations linked to stimuli or environmental contexts. This implies that, when the same stimulus or environmental context reoccurs, the corresponding DLPFC error representation is also reactivated. We hypothesize that this representation will in turn upregulate MPFC activity, reinstating the signal experienced at the time of error, but this time with the purpose of exerting control to prevent the error from happening again (thus leading to a better prediction or a successful behavioral outcome). In this formulation, the translation to a motivational context becomes evident: A performance error, for example, due to task difficulty, would be signaled by increased MPFC activity, tagging that particular behavioral instance as requiring extra effort. Next time the same instance reoccurs, the reactivated error representation can provide information necessary to inform top-down control and resource allocation to result in successful task performance. Noteworthy, this speculative explanation does not require an explicit operationalization of effort or other motivational factors: Thus, the HER model in its current architecture could be able to account for both prediction- and effort-related signals in MPFC and DLPFC. The empirical validity of this explanatory framework is to be tested in future research, which should provide neurobiological evidence for the type of MPFC-DLPFC dynamics described above.

Besides the theoretical implications for understanding PFC circuitry, the model relies on assumptions that require empirical testing. The hierarchical structure of the HER model is consistent with other accounts of PFC function, postulating the existence of a cortical rostrocaudal hierarchical gradient (Koechlin, 2016; Badre, 2008). According to these theories, caudal regions of PFC encode more concrete representations (action related or more recent in time), whereas more rostral regions encode more abstract representations (task sets, rules, context, or information further in the past to be maintained). This is implemented in the HER model, where, in a typical simulation of a working memory task, different items to be stored in working memory are encoded at different levels of the hierarchy (depending on order of processing; see for example the 12AX task simulations in Alexander & Brown, 2015). An underlying assumption is serial processing, not only for a series of stimuli but also for complex stimuli composed by different stimulus features. When placed in the context of motivation and decision-making literature, this assumption is quite relevant: Most of the experiments referenced above combine motivationally salient information of different types, such as required effort and available rewards, presenting it simultaneously. The empirical question remains open as to whether such simultaneous presentation results in simultaneous or serial processing of the presented information sources, and to date, this question has not been addressed. The HER framework hypothesizes that such features would be processed serially in a specific and preferred order. Simulations showed that altering this order, by imposing

a nonpreferred order, can impact performance (Alexander & Brown, 2015).

Presenting motivationally salient information before task performance typically influences accuracy, RT, and task preparation in several tasks requiring cognitive control (Janssens, De Loof, Pourtois, & Verguts, 2016; Boehler, Schevernels, Hopf, Stoppel, & Krebs, 2014; van den Berg, Krebs, Lorist, & Woldorff, 2014; Vassena, Silvetti, et al., 2014; Aarts & Roelofs, 2010). When applied to the domain of effort-based behavior, the order hypothesis predicts that altering order of processing of reward and effort information might result in a shift in perceived subjective value and consequently affect (improve or deteriorate) performance.

### **Predictions and Implications for Clinical Populations**

These theoretical predictions naturally stem from the HER model, and empirical testing of their validity carries relevant implications. First, testing these predictions will (dis-)prove the validity of the assumptions underlying the model. Second, if altering order of processing can alter decision-making, one could test the potential of such manipulation to improve dysfunctional decision-making, for example, concerning health-related behavior such as physical exercise and eating habits. Third, if altering order of processing can alter performance, one could devise optimal ways to reconfigure available motivational information to improve cognitive performance, for example, in educational and school settings. Finally, all of the above have important implications for translational research and potential applications in clinical populations affected by disorders of motivation.

To date, the predictions listed above have not been empirically tested. It is however useful to speculate on the mechanisms, which could underlie such effects. One plausible explanation involves salience. If effort and reward information is processed serially, the order of processing when presentation is simultaneous may be influenced by the respective salience of informative cues. Patients with depression typically show reduced willingness to exert effort to obtain a reward: In other words, they are more effort avoidant as compared with controls (Yang et al., 2014; Treadway, Buckholz, Schwartzman, Lambert, & Zald, 2009). One possible reason could be the overestimation of the amount of the required effort, which would result in an unfavorable overall value, leading to the decision of not engaging in the task. Similarly, reward information could be underestimated, thus reducing the worth of the final value. Note that these hypotheses are in line with what was formulated and simulated with the PRO model in the previous section of this article, where we hypothesized impairment in perceived salience of effort and reward. What is particularly relevant with respect to the order effects is the possibility of intervention: Motivational impairment could derive from altered

perception of saliency; manipulating order of presentation may enforce a specific order of processing, artificially increasing or decreasing saliency of effort and/or reward information, and by tuning this manipulation, one might be able to determine the optimal configuration restoring normal perception of salience. Ideally, this process would result in increasing the willingness to exert effort in exchange for reward, thus counteracting the typical behavioral pattern of anhedonia, a core symptom of depression (Silvia et al., 2014; Treadway, Bossaller, et al., 2012). Critically, alterations in effort- and reward-based decision-making have also been reported in other psychiatric conditions such as bipolar disorder and schizophrenia (Hershenberg et al., 2016; Gold, Waltz, & Frank, 2015; Barch et al., 2014; Fervaha et al., 2013) and preclinical traits of apathy (Bonnelle et al., 2015), although showing different patterns of impairment. Testing the predictions derived from the HER model across different clinical samples could provide insights on shared and dissociable underlying etiopathogenetic mechanisms; moreover, such deeper theoretical understanding could foster development of behavioral treatments aimed at improving decision-making and behavioral outcomes for these patients in daily life.

### **GENERAL DISCUSSION**

This article reviews the theoretical frameworks provided by the PRO and HER models, modeling the neurofunctional architecture of MPFC and DLPFC. Such models have originally been developed based on the core principles of prediction and prediction error to explain empirical effects found in these regions. Here, we discussed how these models may generalize to the domain of motivation, focusing on effort-based behavior. We show that effort effects in MPFC can be successfully accounted for by the PRO model, which provides further predictions regarding behavior and neural activity in both healthy and clinical populations. Furthermore, we discuss the potential translation of the HER model to the domain of effort-based behavior, which accounts for empirical effects measured in DLPFC and provides interesting empirical predictions regarding the effect of order of processing on decision-making and task performance: If these predictions are borne out, such effects could lead to the development of useful interventions to influence altered perception of salience of effort and reward information in clinical populations, potentially improving abnormal behavior.

One primary goal of this article is to emphasize the importance of exploiting precise theoretical frameworks to derive predictions to test experimentally. The first advantage of such mathematically precise frameworks resides in the ability to explain several behavioral and neural effects observed in a brain region under the same computational principle. The second advantage is the possibility to generate new predictions based on the

same model, which can translate to contexts to date untested or different populations. This feature is particularly useful to guide further theory-driven empirical inquiry. In a scientific age where empirical tools proliferate, basing experimental research on strong a priori hypotheses has become a necessary condition to allow drawing statistically meaningful and generalizable conclusions. Finally, such theoretical rigor and quantitative predictive precision provide a great tool to test potential translational applications, with broad explanatory power for understanding the neurobiology of disease.

## Acknowledgments

W. H. A. and J. D. were supported by FWO-Flanders Odysseus II Award G.OC44.13N. E. V. was supported by the Marie Skłodowska-Curie action with a standard IF-EF fellowship, within the H2020 framework (H2020-MSCA-IF2015, grant number 705630).

Reprint requests should be sent to Eliana Vassena, Donders Institute for Brain Cognition and Behavior, Radboud University Nijmegen, Kapittelweg 29, 6525 EN, Nijmegen, The Netherlands, or via e-mail: e.vassena@donders.ru.nl.

## REFERENCES

- Aarts, E., & Roelofs, A. (2010). Attentional control in anterior cingulate cortex based on probabilistic cueing. *Journal of Cognitive Neuroscience*, *23*, 716–727.
- 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. (2014). A general role for medial prefrontal cortex in event prediction. *Frontiers in Computational Neuroscience*, *8*, 69.
- Alexander, W. H., & Brown, J. W. (2015). Hierarchical error representation: A computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation*, *27*, 2354–2410.
- Alexander, W. H., & Brown, J. W. (2016). Frontal cortex function derives from hierarchical predictive coding. *bioRxiv*, 76505. <https://doi.org/10.1101/076505>.
- Alexander, W. H., Fukunaga, R., Finn, P., & Brown, J. W. (2015). Reward salience and risk aversion underlie differential ACC activity in substance dependence. *Neuroimage*, *8*, 59–71.
- Apps, M. A. J., Grima, L. L., Manohar, S., & Husain, M. (2015). The role of cognitive effort in subjective reward devaluation and risky decision-making. *Scientific Reports*, *5*, 16880.
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*, 193–200.
- Bailey, M. R., Simpson, E. H., & Balsam, P. D. (2016). Neural substrates underlying effort, time, and risk-based decision making in motivated behavior. *Neurobiology of Learning and Memory*, *133*, 233–256.
- Barch, D. M., Pagliaccio, D., & Luking, K. (2016). Mechanisms underlying motivational deficits in psychopathology: Similarities and differences in depression and schizophrenia. *Current Topics in Behavioral Neurosciences*, *27*, 411–449.
- Barch, D. M., Treadway, M. T., & Schoen, N. (2014). Effort, anhedonia, and function in schizophrenia: Reduced effort allocation predicts amotivation and functional impairment. *Journal of Abnormal Psychology*, *123*, 387–397.
- Boehler, C. N., Schevernels, H., Hopf, J.-M., Stoppel, C. M., & Krebs, R. M. (2014). Reward prospect rapidly speeds up response inhibition via reactive control. *Cognitive, Affective & Behavioral Neuroscience*, *14*, 593–609.
- Bonnelle, V., Veromann, K.-R., Burnett Heyes, S., Lo Sterzo, E., Manohar, S., & Husain, M. (2015). Characterization of reward and effort mechanisms in apathy. *Journal of Physiology-Paris*, *109*, 16–26.
- Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.
- Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective & Behavioral Neuroscience*, *9*, 16–27.
- 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.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–222.
- Chong, T. T.-J., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biology*, *15*, e1002598.
- Croxxon, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E., & Rushworth, M. F. (2009). Effort-based cost-benefit valuation and the human brain. *Journal of Neuroscience*, *29*, 4531–4541.
- Culbreth, A., Westbrook, A., & Barch, D. (2016). Negative symptoms are associated with an increased subjective cost of cognitive effort. *Journal of Abnormal Psychology*, *125*, 528–536.
- Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, *7*, 415–423.
- Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain: A Journal of Neurology*, *118*, 279–306.
- Engström, M., Karlsson, T., Landtblom, A.-M., & Craig, A. D. B. (2014). Evidence of conjoint activation of the anterior insular and cingulate cortices during effortful tasks. *Frontiers in Human Neuroscience*, *8*, 1071.
- Engstrom, M., Landtblom, A.-M., & Karlsson, T. (2013). Brain and effort: Brain activation and effort-related working memory in healthy participants and patients with working memory deficits. *Frontiers in Human Neuroscience*, *7*, 140.
- Fervaha, G., Graff-Guerrero, A., Zakzanis, K. K., Fousias, G., Agid, O., & Remington, G. (2013). Incentive motivation deficits in schizophrenia reflect effort computation impairments during cost-benefit decision-making. *Journal of Psychiatric Research*, *47*, 1590–1596.
- Gold, J. M., Waltz, J. A., & Frank, M. J. (2015). Effort cost computation in schizophrenia: A commentary on the recent literature. *Biological Psychiatry*, *78*, 747–753.
- Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes*, *100*, 192–196.
- Hershenberg, R., Satterthwaite, T. D., Daldal, A., Katchmar, N., Moore, T. M., Kable, J. W., et al. (2016). Diminished effort on a progressive ratio task in both unipolar and bipolar depression. *Journal of Affective Disorders*, *196*, 97–100.
- Holroyd, C. B. (2016). The waste disposal problem of effortful control. In T. Braver (Ed.), *Motivation and Cognitive Control* (pp. 235–260). New York, NY: Psychology Press.

- Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, *16*, 122–128.
- 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. *Nature Neuroscience*, *7*, 497–498.
- Hosokawa, T., Kennerley, S. W., Sloan, J., & Wallis, J. D. (2013). Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. *Journal of Neuroscience*, *33*, 17385–17397.
- 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.
- Janssens, C., De Loof, E., Pourtois, G., & Verguts, T. (2016). The time course of cognitive control implementation. *Psychonomic Bulletin & Review*, *23*, 1266–1272.
- Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience*, *21*, 1162–1178.
- Klein-Flügge, M. C., Kennerley, S. W., Friston, K., & Bestmann, S. (2016). Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. *Journal of Neuroscience*, *36*, 10002–10015.
- Klein-Flügge, M. C., Kennerley, S. W., Saraiva, A. C., Penny, W. D., & Bestmann, S. (2015). Behavioral modeling of human choices reveals dissociable effects of physical effort and temporal delay on reward devaluation. *PLoS Computational Biology*, *11*, e1004116.
- Koehlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *369*. <https://doi.org/10.1098/rstb.2013.0474>.
- Koehlin, E. (2016). Prefrontal executive function and adaptive behavior in complex environments. *Current Opinion in Neurobiology*, *37*, 1–6.
- Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General*, *139*, 665–682.
- Krebs, R. M., Boehler, C. N., Roberts, K. C., Song, A. W., & Woldorff, M. G. (2012). The involvement of the dopaminergic midbrain and cortico-striatal-thalamic circuits in the integration of reward prospect and attentional task demands. *Cerebral Cortex*, *22*, 607–615.
- Kurniawan, I. T., Guitart-Masip, M., Dayan, P., & Dolan, R. J. (2013). Effort and valuation in the brain: The effects of anticipation and execution. *Journal of Neuroscience*, *33*, 6160–6169.
- Kurniawan, I. T., Guitart-Masip, M., & Dolan, R. J. (2011). Dopamine and effort-based decision making. *Frontiers in Neuroscience*, *5*, 81.
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*, 661–679.
- Massar, S. A. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. L. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *Neuroimage*, *120*, 104–113.
- McCarthy, J. M., Treadway, M. T., Bennett, M. E., & Blanchard, J. J. (2016). Inefficient effort allocation and negative symptoms in individuals with schizophrenia. *Schizophrenia Research*, *170*, 278–284.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Mulert, C., Seifert, C., Leicht, G., Kirsch, V., Ertl, M., Karch, S., et al. (2008). Single-trial coupling of EEG and fMRI reveals the involvement of early anterior cingulate cortex activation in effortful decision making. *Neuroimage*, *42*, 158–168.
- Nee, D. E., Kastner, S., & Brown, J. W. (2011). Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex. *Neuroimage*, *54*, 528–540.
- Nishiyama, R. (2014). Response effort discounts the subjective value of rewards. *Behavioural Processes*, *107*, 175–177.
- Nishiyama, R. (2016). Physical, emotional, and cognitive effort discounting in gain and loss situations. *Behavioural Processes*, *125*, 72–75.
- Parvizi, J., Rangarajan, V., Shirer, W. R., Desai, N., & Greicius, M. D. (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. *Neuron*, *80*, 1359–1367.
- Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate valuation subsystems for delay and effort decision costs. *Journal of Neuroscience*, *30*, 14080–14090.
- Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, *20*, 262–270.
- Ridderinkhof, K. R., van den Wildenberg, W. P., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*, 129–140.
- Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*, 389–397.
- Rushworth, M. F. S., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: One or many serial or parallel systems? *Current Opinion in Neurobiology*, *22*, 946–955.
- Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, *8*, 410–417.
- Rypma, B., Berger, J. S., & D'Esposito, M. (2002). The influence of working-memory demand and subject performance on prefrontal cortical activity. *Journal of Cognitive Neuroscience*, *14*, 721–731.
- Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*, 461–482.
- Salamone, J. D., Correa, M., Yohn, S., Lopez Cruz, L., San Miguel, N., & Alatorre, L. (2016). The pharmacology of effort-related choice behavior: Dopamine, depression, and individual differences. *Behavioural Processes*, *127*, 3–17.
- Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort. *PLoS Biology*, *10*, e1001266.
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*, 217–240.
- Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*, 1286–1291.
- Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience*, *5*, 75.
- Silvetti, M., Seurinck, R., & Verguts, T. (2013). Value and prediction error estimation account for volatility effects in ACC: A model-based fMRI study. *Cortex*, *49*, 1627–1635.
- Silvia, P. J., Mironovová, Z., McHone, A. N., Sperry, S. H., Harper, K. L., Kwapil, T. R., et al. (2016). Do depressive

- symptoms “blunt” effort? An analysis of cardiac engagement and withdrawal for an increasingly difficult task. *Biological Psychology*, *118*, 52–60.
- Silvia, P. J., Nusbaum, E. C., Eddington, K. M., Beaty, R. E., & Kwapil, T. R. (2014). Effort deficits and depression: The influence of anhedonic depressive symptoms on cardiac autonomic activity during a mental challenge. *Motivation and Emotion*, *38*, 779–789.
- Treadway, M. T. (2016). The neurobiology of motivational deficits in depression—An update on candidate pathomechanisms. *Current Topics in Behavioral Neurosciences*, *27*, 337–355.
- Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: A translational model of motivational anhedonia. *Journal of Abnormal Psychology*, *121*, 553–558.
- Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., et al. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *Journal of Neuroscience*, *32*, 6170–6176.
- Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the “EEfRT”? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS One*, *4*, e6598.
- Treadway, M. T., Peterman, J. S., Zald, D. H., & Park, S. (2015). Impaired effort allocation in patients with schizophrenia. *Schizophrenia Research*, *161*, 382–385.
- van den Berg, B., Krebs, R. M., Lorist, M. M., & Woldorff, M. G. (2014). Utilization of reward-prospect enhances preparatory attention and reduces stimulus conflict. *Cognitive, Affective & Behavioral Neuroscience*, *14*, 561–577.
- van Veen, V., Holroyd, C. B., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2004). Errors without conflict: Implications for performance monitoring theories of anterior cingulate cortex. *Brain and Cognition*, *56*, 267–276.
- Vassena, E., Cobbaert, S., Andres, M., Fias, W., & Verguts, T. (2015). Unsigned value prediction-error modulates the motor system in absence of choice. *Neuroimage*, *122*, 73–79.
- 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*, 316.
- 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.
- Verguts, T., Vassena, E., & Silvetti, M. (2015). Adaptive effort investment in cognitive and physical tasks: A neurocomputational model. *Frontiers in Behavioral Neuroscience*, *9*, 57.
- Walton, M. E., Bannerman, D. M., Alterescu, K., & Rushworth, M. F. S. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *Journal of Neuroscience*, *23*, 6475–6479.
- Walton, M. E., Bannerman, D. M., & Rushworth, M. F. S. (2002). The role of rat medial frontal cortex in effort-based decision making. *Journal of Neuroscience*, *22*, 10996–11003.
- Walton, M. E., Groves, J., Jennings, K. A., Croxson, P. L., Sharp, T., Rushworth, M. F. S., et al. (2009). Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision making. *European Journal of Neuroscience*, *29*, 1678–1691.
- Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks*, *19*, 1302–1314.
- Walton, M. E., Rudebeck, P. H., Bannerman, D. M., & Rushworth, M. F. S. (2007). Calculating the cost of acting in frontal cortex. *Annals of the New York Academy of Sciences*, *1104*, 340–356.
- Westbrook, A., & Braver, T. S. (2013). The economics of cognitive effort. *Behavioral and Brain Sciences*, *36*, 704–705.
- Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective & Behavioral Neuroscience*, *15*, 395–415.
- Westbrook, A., & Braver, T. S. (2016). Dopamine does double duty in motivating cognitive effort. *Neuron*, *89*, 695–710.
- Yang, X.-H., Huang, J., Zhu, C.-Y., Wang, Y.-F., Cheung, E. F. C., Chan, R. C. K., et al. (2014). Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry Research*, *220*, 874–882.

Copyright of Journal of Cognitive Neuroscience is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.