

Misdirected vigor: Differentiating the control of value from the value of control

Amitai Shenhav^{1,2}, Sebastian Musslick², Matthew M. Botvinick^{3,4}, Jonathan D. Cohen^{2,5}

¹ *Department of Cognitive, Linguistic, and Psychological Sciences, Carney Institute for Brain Science, Brown University, Providence, RI 02912*

² *Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540*

³ *DeepMind, London, UK NIC 4AG*

⁴ *Gatsby Computational Neuroscience Unit, University College London, WIT 4JG*

⁵ *Department of Psychology, Princeton University, Princeton, NJ 08540*

Correspondence should be addressed to:

Amitai Shenhav
190 Thayer St., Box 1821
Brown University
Providence, RI 02912

Email: amitai_shenhav@brown.edu
Telephone: (401) 863-5198
Fax: (401) 863-2255

Theories of dACC function have to contend with an increasingly long and diverse list of signals that have been tied to this region^{1,2}. To account for this apparent heterogeneity, we recently proposed a theory of dACC function that embraces that heterogeneity and offers a unifying function focused on the evaluation, motivation and allocation of cognitive control. Our theory proposes that dACC integrates information relevant to determining the Expected Value of Control (EVC) – including signals pertaining to demands for control (i.e., how much and what kind of control is needed) and those pertaining to the potential incentives for control (i.e., what are the consequences of exerting control) – and uses this to determine how best to allocate cognitive control^{1,2}. We have shown that this theory accounts for variability in behavior and control states across a variety of experiments and task conditions^{3,4}, and that it can account for the confluence of seemingly unrelated task-related variables within this region, including those related to reward, decision conflict, and the unexpectedness of a stimulus or outcome (i.e., surprise) (Fig. 1).

Vassena and colleagues⁵ recently sought to test predictions of the EVC model, and concluded that it was unable to account for patterns of task-related dACC activity that they observed. Specifically, they show that dACC activity increased with levels of conflict and surprise during a choice task, and argue that the EVC model predicts the opposite pattern of results. Here, we identify misconstruals in analysis and interpretation that led to this conclusion.

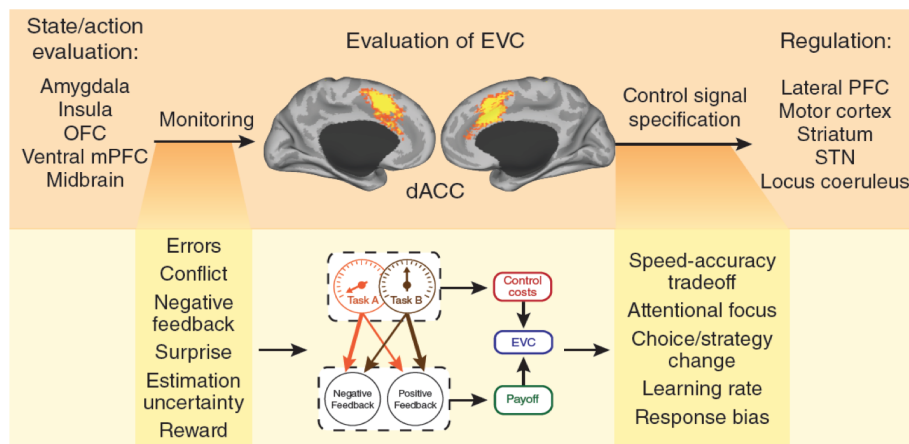


Figure 1. The EVC theory proposes that dACC integrates information about incentives and demands for control, including signals of processing conflict and unexpected changes in the environment (i.e., surprise), to determine how best to adjust control allocation with and across tasks. Vassena et al. join three previous experiments^{6,7} in demonstrating that dACC tracks both conflict and surprise during value-based choice, but diverges from those studies in suggesting that these effects are contrary to the predictions of the EVC theory. Figure adapted from Shenhav et al. (2016).

Participants in Vassena et al.’s study performed a task that required them to quickly determine which of two sets of options would yield more reward. To apply the EVC model to their task, the authors make several critical assumptions. First, they assume that the scope of control during this task is limited to a single consideration: *how vigorously should I respond to a given option on this trial?* Second, they assume that the sole incentive determining the value of control is the monetary reward for this task, which is weighed against the cost of exerting control. Thus, they argue that their control signal (vigor) should intensify as the reward for the associated option increases, such that an option worth 50 cents will elicit a more ‘vigorous’ response than a 20-

cent option, increasing its chance of being chosen. Applying this to EVC, Vassena et al. assume that dACC will, in their task, determine the overall value of *vigor* for each response. As we describe below, this inappropriately restricted version of our model, which we will refer to as the *Expected Value of Vigor (EVV)* model, not only fails to predict the patterns of dACC activity they observe, it also fails as a descriptive and mechanistically plausible model of control for their task (Fig. 2).

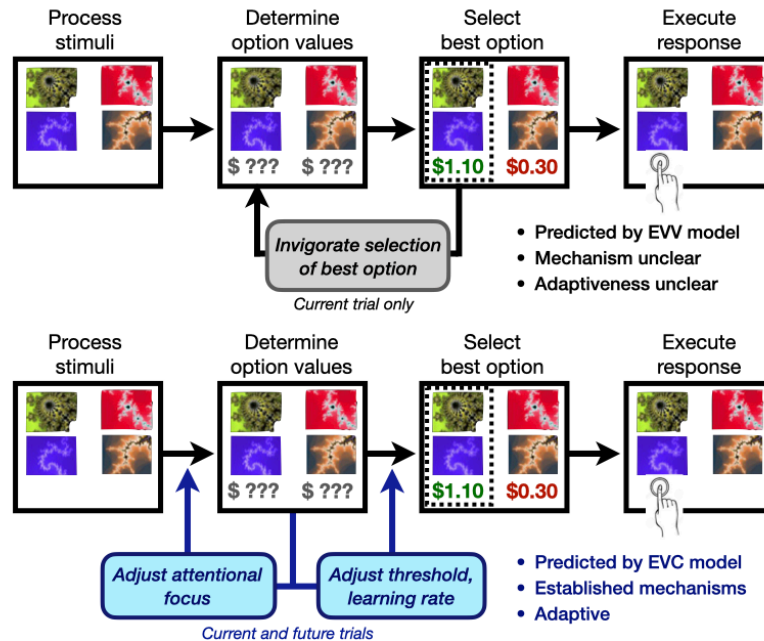


Figure 2. To refute the EVC theory, Vassena et al. simulate dACC activity during their task based on a reinterpretation of the model, EVV (**top**). Participants were tasked with quickly selecting which pair of previously learned options promised greater reward. Under the EVV model, dACC has advance information about which option is more rewarding (i.e. it effectively knows the correct response), and uses this to promote the choice of that more rewarding option (see their Supplementary Methods). By contrast, EVC (**bottom**) predicts that dACC uses information available while choosing, such as conflict and surprise, to adjust control directed at the task (e.g., attention to options, threshold) and broader goals (e.g., learning, attention to other tasks).

EVV is neither a necessary nor sufficient account of control

EVV proposes that the role of control in this task is to invigorate responses that will yield greater reward (Fig. 2, top). While intuitive at first blush, on closer scrutiny it becomes clear that this proposal is mechanistically and normatively problematic. To make their choices, participants performing this task need to determine the value of their options. The EVV model aims to intervene during this evaluation process, but to do so it *relies* on information about those option values (the very thing they are trying to determine). EVV can be calculated once these values are determined, but at that point it is likely that the participant already has enough information to make a choice.

The authors claim that vigor is necessary for hastening this choice, to ensure that participants respond before their deadline. As evidence of this, they show that a participant's threshold for making a decision varied based on the relative values of the options on a given trial: easy choices (those with large value differences) were associated with higher thresholds and hard choices

(small value differences) were associated with lower thresholds. But there is a much simpler explanation for this pattern of results. If deadlines were being enforced by a threshold that decreases (collapses) over the course of a decision – as has been shown to be normative and evident in similar tasks⁸ – then it would *appear* as though thresholds were being controlled as a function of difficulty level, when in fact these threshold differences would actually reflect differences in how much the threshold had collapsed organically by the time an easy versus hard choice was made.

EVV is thus designed to solve a problem that has already been addressed (guaranteeing that choices are made before a deadline), using a solution that lacks clear theoretical or empirical grounding (e.g., relying on information needed to make the choice). Furthermore, this solution fails to predict relevant features of their behavioral findings. As noted above, the authors found that thresholds were lowest for the hardest choices. While vigor may not be necessary to account for this finding (e.g., if a collapsing bound were in place), from a vigor perspective this finding suggests that these hard choices engendered high levels of vigor (control) to ensure a timely response. Their EVV calculations, however, make the opposite prediction: that control should be at its *lowest* when choices are hardest.

In the context of these more fundamental concerns about the model's validity, it is much less remarkable that Vassena and colleagues' EVV model *also* cannot account for patterns of dACC activity during this task.

EVC provides a more complete account of control allocation

The EVC theory differs from EVV in at least two important respects. First, drawing on decades of theoretical and empirical research, EVC theory proposes that the dACC monitors information about incentives and task demands that can bear on control allocation, not only for the current trial of the current task, but also for subsequent trials and other possible tasks. When someone encounters an unexpected event, or information that conflicts with their current goal, they mobilize control processes to optimize both short and long-term goals, including adjusting performance on the next trial of the current task (e.g., increased attention, adjusting response threshold), focusing greater attention on other relevant tasks, and/or adjusting the rate of learning about one's environment⁹⁻¹¹ (Fig. 2, bottom). EVC can accommodate this range of conflict- and surprise-triggered control adjustments (Fig. 1), some of which were evident in the between-trial performance adjustments participants made while performing Vassena et al.'s task.

Second, while both models share the critical assumption that control is allocated based on a cost-benefit analysis, EVV takes a very narrow view of what might constitute a benefit: namely, the monetary reward that could accrue from the current trial. Here, again, EVC draws insights from the extant literature, which suggests that inputs to this cost-benefit analysis are broader in type (e.g., incorporating intrinsic rewards such as a bias towards being accurate) and scope (e.g., other timescales and tasks). For instance, participants will have likely learned over their lifetimes that surprise and conflict signals are generally associated with an increased value for certain control states (e.g., increased attentional engagement, evaluation, and/or caution), whether or not those control adjustments accrue reward for the task immediately at hand^{3,9,12,13}.

By accounting for the array of inputs to and outputs of decisions to allocate control, EVC can explain why Vassena and colleagues, like others^{6,7}, found that dACC activity scales with both conflict and surprise during a choice task.

Towards stronger tests of motivated control models

Vassena and colleagues sought to rule out an EVC account of their dACC findings, but their data rule out only an inappropriately restricted version of that model. As we explain above, a more appropriate implementation of EVC should, under plausible assumptions, be able to generate patterns of control allocation that qualitatively mirror patterns of neural activity they observed in dACC. However, translating these into quantitative estimates of fMRI activity requires overcoming two major obstacles.

First, it requires understanding the normative and empirical basis of control adjustments during this particular kind of task, and the incentives that drive those adjustments. This obstacle has largely been overcome by the majority of research into motivated control, by focusing on tasks which – unlike the current one – involve control processes that have been normatively motivated and empirically well-characterized (e.g., Stroop, flanker, task-switching)¹³⁻¹⁵. These properties have enabled the development of process models that can be used in conjunction with empirical data to distinguish between different types and intensities of control^{3,4}.

Second, predicting population-level neural activity (e.g., fMRI data) will require a better understanding of the representational content of neural activity associated with motivated control, and how this translates into BOLD activity. For instance, as we have previously noted^{1,6}, conflict and surprise can elicit increases in BOLD activity because they signal a number of relevant inputs and outputs for a motivated control decision, including increases in the demand for control, the costs of the demanded control, and the intensity of the control output.

Vassena and colleagues' preferred model of dACC (PRO⁵) provides a normative, empirically-grounded, and testable account of behavior and neural activity, and their latest paper offers further evidence in favor of this model. These strengths are obscured when their only reference points for model comparison are (1) accounts of dACC that have already been roundly refuted (as with the “choice difficulty” model; e.g., Refs. ^{2,6,7,9,10}) and (2) implementations of an existing model that are incomplete, implausible, and inconsistent with empirical evidence (as in the case of EVV vs. EVC). By providing additional clarity on the nature of control and its evaluation, we hope that we have provided a path towards further analysis and comparison of each of these respective models.

Acknowledgements

We are grateful to Romy Frömer and Harrison Ritz for helpful discussions and feedback on earlier versions of this manuscript.

References

- 1 Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217-240 (2013).
- 2 Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286-1291 (2016).
- 3 Lieder, F., Shenhav, A., Musslick, S. & Griffiths, T. L. Rational metareasoning and the plasticity of cognitive control. *PLOS Computational Biology* **14**, e1006043 (2018).
- 4 Musslick, S., Shenhav, A., Botvinick, M. M. & Cohen, J. D. in *Proceeding of the 2nd Multidisciplinary Conference on Reinforcement Learning & Decision Making (2015)*.
- 5 Vassena, E., Deraeve, J. & Alexander, W. H. Surprise, value and control in anterior cingulate cortex during speeded decision-making. *Nature Human Behaviour*, 1-11 (2020).
- 6 Shenhav, A., Straccia, M. A., Cohen, J. D. & Botvinick, M. M. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat. Neurosci.* **16**, 1127-1139, doi:10.1038/nn.3771 (2014).
- 7 Lin, H., Saunders, B., Hutcherson, C. A. & Inzlicht, M. Midfrontal theta and pupil dilation parametrically track subjective conflict (but also surprise) during intertemporal choice. *NeuroImage* **172**, 838-852 (2018).
- 8 Bhui, R. Testing Optimal Timing in Value-Linked Decision Making. *Computational Brain & Behavior* **2**, 85-94 (2019).
- 9 Wessel, J. & Aron, A. On the Globality of Motor Suppression: Unexpected Events and Their Influence on Behavior and Cognition. *Neuron* **93**, 259 - 280 (2017).
- 10 McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally Dissociable Influences on Learning Rate in a Dynamic Environment. *Neuron* **84**, 870-881, doi:10.1016/j.neuron.2014.10.013 (2014).
- 11 Cavanagh, J. F. & Frank, M. J. Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* **18**, 414-421, doi:10.1016/j.tics.2014.04.012 (2014).
- 12 Bejjani, C., Zhang, Z. & Egner, T. Control by association: Transfer of implicitly primed attentional states across linked stimuli. *Psychonomic Bulletin and Review* **25**, 617-626 (2018).
- 13 Bustamante, L., Lieder, F., Musslick, S., Shenhav, A. & Cohen, J. D. Learning to Overexert Cognitive Control in a Stroop Task. PsyArxiv 3rynj [Preprint]. February 20, 2020. Available from: <https://doi.org/10.31234/osf.io/3rynj>. (2020).
- 14 Parro, C., Dixon, M. L. & Christoff, K. The neural basis of motivational influences on cognitive control. *Hum. Brain Mapp.* **39**, 5097-5111 (2018).
- 15 Frömer, R., Lin, H., Dean Wolf, C. K., Inzlicht, M. & Shenhav, A. When effort matters: Expectations of reward and efficacy guide cognitive control allocation. BioRxiv 095935 [Preprint]. May 14, 2020. Available from: <https://doi.org/10.1101/2020.05.14.095935>. (2020).