Neurocomputational mechanisms of motivational influences on mental effort

Debbie Yee (debbie_yee@brown.edu), Mahalia Prater Fahey (mahalia_prater_fahey@brown.edu)
Xiamin Leng (xiamin_leng@brown.edu), Ziwei Cheng (ziwei_cheng@brown.edu)
Maisy Tarlow (maisy_tarlow@brown.edu), Joonhwa Kim (joonhwa_kim@brown.edu)
Kaitlyn Mundy (kaitlyn_mundy@brown.edu), Samuel Nevins (samuel_nevins@brown.edu)
Amitai Shenhav (amitai_shenhav@brown.edu)
Cognitive Linguistic and Psychological Sciences, Carney Institute for Brain Science
190 Thayer Street, Providence, RI 02906 USA

Abstract

Human motivation is fundamentally shaped by one's expectations of their outcomes (e.g., reward, punishment), as well as the type of effort required to attain these outcomes (e.g., attention vs. caution). In our fMRI study (n=100), we observed a dissociation between how rewards promoted increased attentional control (drift rate) vs. how penalties promoted increased caution (decision threshold). We found that *a priori* brain regions were associated with faster RT or increased accuracy. Model-based fMRI analyses revealed caudal vs. rostral dorsal anterior cingulate cortex regions are associated with drift rate and threshold, respectively. Together, these data reveal that distinct dACC regions underlie how motivational incentives can drive attention-related and caution-related strategies for the adaptive allocation of cognitive control.

Keywords: motivation; cognitive control; reward; punishment; drift diffusion model; model-based fMRI

Introduction

Humans adaptively use differing cognitive control strategies to maximize expected reward (e.g., increased attentional control for rewards, increased response caution for penalties) (Leng et al., 2021; Prater Fahey et al., 2023). Yet, the neural mechanisms that underlie how incentives modulate these computational strategies to facilitate mental effort remain unclear.

A large network of brain regions has been implicated in motivation and cognitive control, including the dorsal anterior cingulate (dACC), ventral striatum (VS), anterior insula (AI), lateral prefrontal cortex (lPFC), and inferior frontal gyrus (IFG) (Parro et al., 2018). The Expected Value of Control Theory organizes this neural circuit around a core set of computations that distinguish between temporally distinct sub-processes related to 1) evaluating the incentives for a given task and 2) integrating those incentives to optimally allocate mental effort (Shenhav et al., 2017). The dACC has been implicated in representing integrated motivational value (Yee et al., 2021), yet precisely how this value signal is translated to dissociable strategic adjustments in control is unknown (Ritz et al., 2022).

Methods

We conducted an fMRI study (n=100), where participants performed our Multi-Incentive Control Task, through which participants could earn low vs. high monetary rewards for accurate responses and were penalized with low vs. high monetary losses for inaccurate responses. Our self-paced interval-based design allowed individuals to select when and how much effort to exert over a fixed time interval (8-12 s). We applied the drift diffusion model to quantify to what extent reward vs. punishment incentives biased dissociable strategies for control allocation (e.g., attention vs. caution).

We selected regions of interest (ROIs) based on prior meta-analyses of motivation and cognitive control, value-based decision-making, and effort-based decision-making (Parro et al., 2018; Lopez-Gamundi et al., 2021; Bartra et al., 2013; Hampshire et al., 2010). Given the functional heterogeneity of the dACC (Vega et al., 2016), we divided dACC into caudal vs. rostral subregions based on participation in salience vs. control networks (Schaefer et al., 2017; Kong et al., 2021).

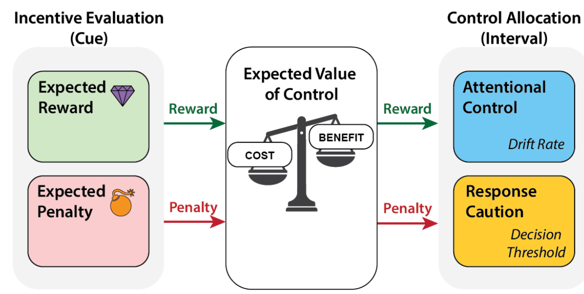

Figure 1: Expected Value of Control Model Predictions. The model predicts a dissociation between how reward biases increased attentional controls (via drift rate) and how penalty biases increased response caution (via decision threshold).

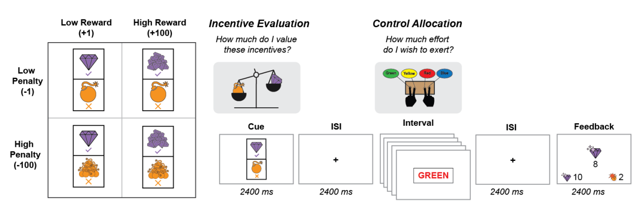

Figure 2: Multi-Incentive Control Task (MICT). At the start of each interval, a cue indicates potential rewards for correct responses and potential penalties for incorrect responses. During the interval, participants completed Stroop trials at their own pace (e.g., as fast or slow as they would like). At the end of each interval, they received feedback of their net earnings.

Results

Behavioral Task Performance

We parametrized task performance with the drift diffusion model (Ratcliff et al., 2016; Wiecki et al., 2013), a computational framework that combines accuracy and response time to estimate model parameters that reflect the processes that underlie the rate at which noisy evidence is accumulated (drift rate) until a response criterion is reached (decision threshold). Consistent with previous findings and normative model predictions, higher reward was associated with an increased drift rate (p=.009) and lower threshold ($p<.001$), and higher penalty was associated with an increased threshold ($p<.001$).

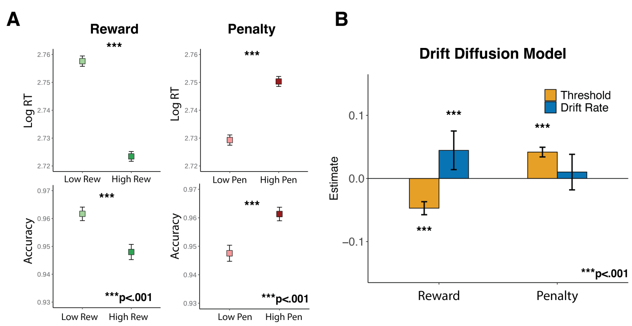

Figure 3: a) Task Performance. Participants were faster and less accurate for larger rewards (p's$<$.001), and were slower and more accurate for larger penalties (p's$<$.001). b) DDM Results. Higher reward is associated with increased attentional control (drift rate), and higher penalty is associated with increased response caution (decision threshold).

Distinct dACC regions for Attention and Caution

We extracted cue-related activity from our ROIs and used mixed-level models to examine their influence on RT and accuracy. First, we observed a cluster of ROIs associated with faster RT (VS, caudal dACC, lPFC-salience p's$<$.001). We observed a second ROI cluster associated with increased accuracy (AI, rostral dACC, lPFC-control, IFG, p's$<$.005). It is noteworthy that caudal dACC predicted faster RT ($p<.001$) and increased accuracy (p=.022), whereas rostral dACC predicted increased accuracy and a trend for slower RT (p=.28).

We performed model-based fMRI analyses that included ROI activity as regressors in the DDM. VS and caudal dACC were associated with increased drift rate ($p<.001$). Conversely, rostral dACC and IFG were associated with a higher threshold (p's$<$.06). To determine if caudal and rostral dACC activity contributed to shared variance, we included both ROIs in the DDM. We observed a dissociation in the control strategy, with caudal dACC predicting increased drift rate ($p<.001$) and rostral dACC predicting increased threshold ($p<.05$).

Neural modulators of motivational incentive effects on control allocation

Next, we tested if these ROIs encoded incentive effects on control allocation. We examined two-way interactions between ROIs and reward and penalty incentives in our mixed-level models. Greater activity in VS and caudal dACC was associated with a greater decrease in RT for high (relative to low) reward (p's$<$.003) and slower RT for high (relative to low) penalty (p's$<$.04). Conversely, greater activity in AI and rostral dACC was associated with increased accuracy for high (relative to low) penalty (p's$<$.07) and lower accuracy for high (relative to low) reward (p's$<$.03). In lPFC-control and IFG, we only observed interactions with penalty, with increased accuracy for high (relative to low) penalty p's$<$.099).

Model-based analyses revealed these incentive interactions primarily impacted the decision threshold. Two-way interactions revealed that higher caudal dACC activity was associated with a greater reduction in threshold for high reward (p=.02) and a greater increase in threshold for high penalty (p=.02). Similarly, greater rostral dACC activity was associated with a lower threshold for high reward (p=.02) and a trend towards a higher threshold for high penalty (p=.08).

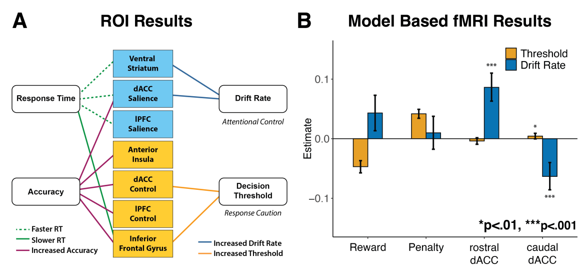

Figure 4: a) ROI Results. A subset of ROIs is associated with faster RT (VS, caudal dACC, lPFC-salience), and another subset is associated with increased accuracy (AI, rostral dACC, lPFC-control, IFG). VS and caudal dACC were associated with increased drift rate, whereas rostral dACC and IFG were associated with increased threshold. b) Model-based fMRI analyses revealed a dissociation between caudal dACC modulating drift rate and rostral dACC modulating threshold.

Discussion

Here, we identify neurocomputational mechanisms that underlie how motivational incentives influence dissociable control strategies for effort allocation. Specifically, we identified distinct caudal vs. rostral dACC regions associated with increased attentional control vs. response caution, respectively. These regions interacted with rewards and penalties to influence decision threshold, revealing how greater neural activity in these ROIs is linked to greater differentiation in incentive-modulated control adjustments. Future work aims to examine to what extent individual differences in these mechanisms are associated with environmental factors (e.g., socioeconomic status) and subjective measures (e.g., self-report ratings).

Acknowledgments

This work was supported by T32-MH126388 (DY), T32-MH115895 (XL), NSF Grad Research Fellowship (MPF), NSF CAREER Award 206611 and Carney Innovation Award (AS).

References

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412—427. doi: 10.1016/j.neuroimage.2013.02.063

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage*, *50*(3-3), 1313–1319. doi: 10.1016/j.neuroimage.2009.12.109

Kong, R., Yang, Q., Gordon, E., Xue, A., Yan, X., Orban, C., ... Yeo, B. T. T. (2021). Individual-Specific Areal-Level Parcellations Improve Functional Connectivity Prediction of Behavior. *Cerebral Cortex*, *31*(10), 4477–4500. doi: 10.1093/cercor/bhab101

Leng, X., Yee, D., Ritz, H., & Shenhav, A. (2021). Dissociable influences of reward and punishment on adaptive cognitive control. *PLOS Computational Biology*, *17*(12), e1009737. doi: 10.1371/journal.pcbi.1009737

Lopez-Gamundi, P., Yao, Y.-W., Chong, T. T.-J., Heekeren, H. R., Herrero, E. M., & Pallares, J. M. (2021). The neural basis of effort valuation: A meta-analysis of functional magnetic resonance imaging studies. *Neuroscience & Biobehavioral Reviews*. doi: 10.1016/j.neubiorev.2021.10.024

Parro, C., Dixon, M. L., & Christoff, K. (2018). The neural basis of motivational influences on cognitive control. *Human Brain Mapping*, *39*(12), 5097–5111. doi: 10.1002/hbm.24348

Prater Fahey, M., Yee, D. M., Leng, X., Tarlow, M., & Shenhav, A. (2023). Motivational context determines the impact of aversive outcomes on mental effort allocation. *bioRxiv*. doi: 10.1101/2023.10.27.564461

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260—281. doi: 10.1016/j.tics.2016.01.007

Ritz, H., Leng, X., & Shenhav, A. (2022). Cognitive Control as a Multivariate Optimization Problem. *Journal of Cognitive Neuroscience*, 1–23. doi: 10.1162/jocn_a_01822

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., ... Yeo, B. T. (2017). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*(September), 1—20. doi: 10.1093/cercor/bhx179

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, *40*(1), 99—124. doi: 10.1146/annurev-neuro-072116-031526

Vega, A. D., Chang, X. L. J., Banich, M. T., Wager, X. T. D., & Yarkoni, X. T. (2016). Large-Scale Meta-Analysis of Human Medial Frontal Cortex Reveals Tripartite Functional Organization. *Journal of Neuroscience*, *36*(24), 6553—6562. doi: 10.1523/jneurosci.4402-15.2016

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*(August), 1—10. doi: 10.3389/fninf.2013.00014

Yee, D. M., Crawford, J. L., Lamichhane, B., & Braver, T. S. (2021). Dorsal Anterior Cingulate Cortex Encodes the Integrated Incentive Motivational Value of Cognitive Task Performance. *J. Neurosci.*, *41*(16), 3707–3720. doi: 10.1523/jneurosci.2550-20.2021