Disentangling influences of aversive motivation on control allocation across distinct motivational contexts

Mahalia Prater Fahey*
Cognitive, Linguistic,
& Psychological Sciences
Brown University
Providence, RI 02906
mahalia_prater_fahey@brown.edu

Debbie M. Yee*
Cognitive, Linguistic,
& Psychological Sciences
Brown University
Providence, RI 02906
debbie_yee@brown.edu

Xiamin Leng
Cognitive, Linguistic,
& Psychological Sciences
Brown University
Providence, RI 02906
xiamin_leng@brown.edu

Maisy Tarlow
Cognitive, Linguistic,
& Psychological Sciences
Brown University
Providence, RI 02906
maisy_tarlow@brown.edu

Amitai Shenhav
Cognitive, Linguistic,
& Psychological Sciences
Brown University
Providence, RI 02906
amitai_shenhav@brown.edu

Abstract

Motivation and cognitive control are integral to adaptive goal-directed behavior. Prior research has shown how cognitive control allocation can be influenced by both positive (e.g., monetary bonuses) and negative incentives (e.g., monetary losses). However, it remains unclear to what extent decisions to allocate effort in cognitively demanding tasks are driven by the magnitude and valence of these incentives (e.g., reward vs. penalty), or if they are general or specific across different *motivational contexts*, i.e., whether a given incentive promotes action (*reinforcement*) or caution (*punishment*). Here, we use a combination of modeling and experimentation to characterize dissociable influences of incentives on control allocation across these distinct motivational contexts. We had participants perform a novel incentivized cognitive control task, which reinforced correct responses and penalized errors. Critically, in addition to varying reward motivation for accurate performance, we varied aversive motivation in two ways; by threatening monetary loss for either failing to perform well (negative reinforcement) or for performing poorly (punishment). Using a reward-rate-optimal model of control allocation, we generated predictions for how reinforcement and punishment should differentially influence control adjustments, such that higher levels of reinforcement should increase drift rate and decrease threshold, and higher levels of punishment should primarily increase threshold. We validated these predictions experimentally, demonstrating that normative patterns of control adjustment are found for varying levels of reinforcement independent of valence, and that these patterns are distinct from those predicted and observed for varying levels of punishment. By combining theoretical and empirical approaches to delineate the *motivational context* of incentives, this work provides novel insights into the multi-faceted and multivariate influences of motivation on cognitive control.

Keywords: punishment, negative reinforcement, cognitive control, drift diffusion model, reward rate optimization

Acknowledgements

This work was supported by National Science Foundation CAREER Grant 2046111 to A.S., and Brown University Office of the Vice President Research Seed Award to A.S. and D.M.Y. D.M.Y. was supported by National Institutes of Health Training Award T32-MH126388, and X.L. was supported by T32-MH115895. M.P.F. was supported by National Science Foundation Graduate Research Fellowship Program.

*M.P.F. and D.M.Y. contributed equally to this paper.

1 Introduction

The interaction between motivation and cognitive control is central to guiding adaptive behavior (Botvinick & Braver, 2015). Though prior research has examined the influence of *postive* (e.g., Frömer et al., 2021; Chiew & Braver, 2016) as well as *negative* incentives (e.g., Braem et al., 2013; Cubillo et al., 2019) on cognitive control, studies have rarely considered how the same incentive may drive different types of control signals (Danielmeier & Ullsperger, 2011; Ritz et al., 2022). For example, motivation to avoid negative outcomes can either promote more active engagement with a task (e.g., by increasing attention to task-relevant stimuli) or promote more cautious performance of the same task (e.g., by increasing one's threshold for responding). Here we argue that consideration of the *motivational context* of whether aversive incentives either reinforce or punish behavior may promote clearer understanding of the computational mechanisms underlying dissociable strategies for cognitive control allocation (Yee et al., 2022).

Recent work from our group has shown how a normative account can explain how positive and negative incentives differentially influence the manner and intensity through which individuals allocate control (Leng et al., 2021). The Expected Value of Control model posits that people weigh the expected benefits (e.g., net value of positive and negative outcomes) against the expected costs (e.g., the mental effort required to exert control) (Shenhav et al., 2017). The model predicts that cognitive control can be adjusted to modify specific parameters of the drift diffusion model (DDM) to upregulate increased attentional control (e.g., drift rate v) or changes in response threshold a. According to this framework, *Reward Rate* can be estimated as a function of task performance (e.g., error rate ER and response time RT), as well as reinforcement for a correct response R and punishment for incorrect response P.

$$RewardRate = \frac{R \times (1 - ER) - P \times ER}{DT + NDT} - E \times v^2 \tag{1}$$

Although our previous theoretical and experimental work has assumed that positive outcomes are tied to good performance and aversive outcomes with poor performance, a key feature of this normative account is that it can flexibly account for how negative incentives may promote good performance. By stipulating distinct parameters for positive reinforcement [R_{pos}] and negative reinforcement [R_{neg}] in our modified reward rate optimization model (Bogacz et al., 2006), we can investigate the degree to which negative reinforcement may produce similar patterns as positive reinforcement effects, as well as evaluate whether negative reinforcement versus punishment elicit distinct influences on cognitive control allocation. To test these hypotheses, we developed a novel task that varies positive reinforcement, negative reinforcement, and punishment. By distinguishing between the roles that avoidance of aversive outcomes play in boosting attention (e.g., negative reinforcement) versus response caution (e.g., punishment), our findings enable more precise characterization of the neural and computational mechanisms driving motivation and cognitive control.

2 Incentivized Cognitive Control Task

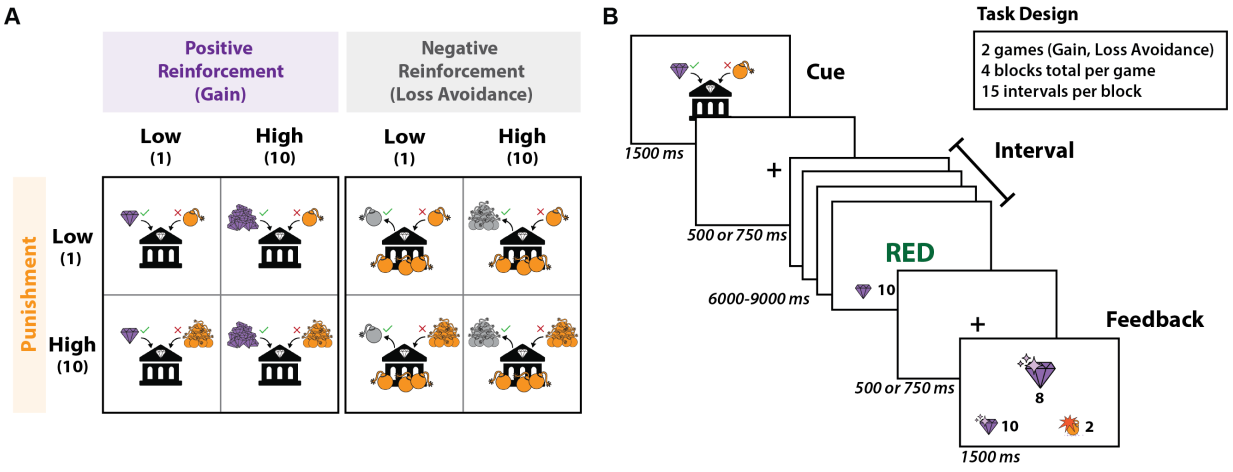

Figure 1: A) Participants completed a *Positive Reinforcement Task* and *Negative Reinforcement Task*, wherein they could earn or avoid losing money, respectively. We varied aversive motivation in two ways (e.g., threatening monetary loss for failing to perform well (negative reinforcement) or performing poorly (punishment) to evaluate the degree avoidance motivation may promote dissociable control adjustments. B) *Task Paradigm*. At the start of the interval, a cue indicates the amount of gain or loss avoided for each correct response and the penalty amount for each incorrect response. Participants completed varying Stroop trials within the interval, and receive feedback on their net earnings at the end of the interval.

Participants: 261 participants (18-55 years) completed the study via Prolific. Participants were born in and currently residing within the United States, currently enrolled in a Bachelor's program, and have normal/corrected vision with no colorblindness. 32 participants were excluded (31 participants failed comprehension quizzes, 1 subject for corrupted data). The final sample consisted of 229 participants (F: 144, M: 106, Other: 9; Mean Age: 22.63 yrs).

Procedure: Participants completed two "games" of Stroop tasks (counterbalanced) during which they could either earn money (Positive Reinforcement) or avoid losing money (Negative Reinforcement). They were initially endowed with 12.00 of bonus money which was converted into 1200 gems for the game and added to a bank. Each game consisted of intervals during which participants could complete as many Stroop trials as they wished within a fixed time interval (6-9s). Participants completed 4 blocks of 15 intervals in which we fixed the magnitude of one dimension while randomly varying the other dimension (e.g., if reward level fixed within a block, then punishment varied across intervals). At the end of each interval, any remaining bombs in the bank would destroy the equivalent number of gems.

Positive Reinforcement Task: Participants collected additional gems to add to their initial endowment. Each correct response added gems to participant's bank and each incorrect response added bombs to their bank. At the start of each interval, a cue indicated whether the participant would earn a small or large number of gems for each correct response, as well as whether they would add a small or large number of bombs for each error. At the end of the interval, participants received feedback on the net gems they earned or lost (gems added - bombs added).

Negative Reinforcement Task: This task was the same as above, except that now participants protected the gems in their intial endowment, which were threatened by 300 bombs that were added to their bank at the start of each interval. Each correct response removed bombs from the participant's bank and each incorrect response added additional bombs to their bank. At the start of each interval, a cue indicated the whether the participant would remove a small or large number of bombs for each correct response, as well as whether they would add a small or large number of bombs for each error. Participants received feedback on the net gems they saved or lost (bombs removed - bombs added).

2.1 Task Performance

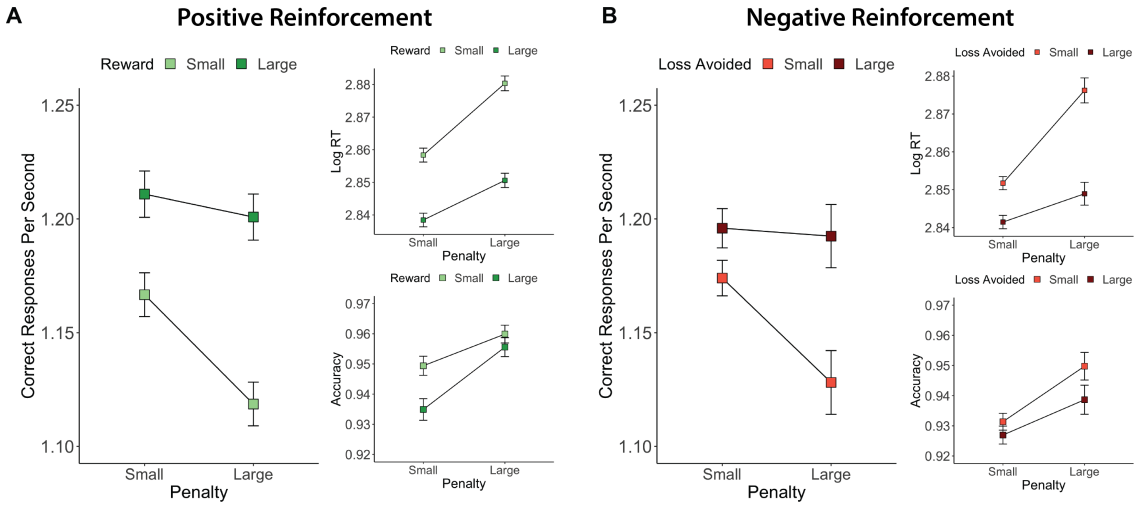

Figure 2: A) Positive Reinforcement Task Performance. B) Negative Reinforcement Task Performance.

In the *Positive Reinforcement Task*, participants completed more correct trials when expecting a large relative to a small reward (b = 9.20, p<.001). Conversely, participants completed fewer correct trials when expecting a large relative to a small penalty (b = -5.65, p<.001). When decomposing these effects by speed and accuracy, we find that larger expected rewards induced faster RTs (b = -11.94, p<.001) and moderately less accurate responses (b = -4.04, p<.001), while larger expected penalties induced slower (b = 10.15, p<.001) and more accurate (b = 8.82, p<.001) responses. Finally, we observed a significant interaction in response rate (b = -4.34, p<.001), in that during higher levels of punishment, the influence of reward level on RT increased (b = -5.67, p<.001) and accuracy decreased (b = 9.02, p<.001) (Fig 2A). These behavioral patterns are consistent with our previous study (Leng et al., 2021).

The behavioral patterns observed in the *Negative Reinforcement Task* (Fig 2B) largely mirrored those for positive reinforcement. Participants completed more correct trials when avoiding a large loss (i.e., more bombs removed) relative to avoiding a small loss (i.e., fewer bombs removed) (b = 6.69, p<.001). Participants completed fewer correct trials when expecting a large relative to a small penalty (b = -4.19, p<.001). Similar to above, larger avoided lossses induced faster response times (b = -10.10, p<.001) and moderately less accurate responses (b = -3.54, p<.001), whereas larger penalties

induced slower (b = 8.37, p<.001) and more accurate (b = 5.98, p<.001) responses. We also observed a signficant interaction in response rate (b = 3.99, p<.001), in that during higher levels of punishment, the influence of loss avoidance level on RT increased (b = -7.67, p<.001) as well as on accuracy increased (b = -3.42, p<.001).

3 Model Predictions and Empirical Results

3.1 Drift Diffusion Model Posterior Parameter Estimates

Following past work, we characterized Stroop responding as a drift diffusion process in which evidence is accumulated toward one response until the accumulated evidence reaches a threshold (Musslick et al., 2015; Leng et al., 2021). We fit the accuracies and RTs for the different incentive conditions separately for each task with a Hierarchical Drift Diffusion Model (Wiecki et al., 2013) to derive estimates of how a participant's drift rate and threshold varied across different levels of reinforcement and punishment. This model controls for the effect of congruency, starting point bias, and variabilty.

Consistent with patterns previously observed, and predicted by our reward-rate-optimal model (see Section 3.2), we found that during our *Positive Reinforcement Task*, larger expected rewards promoted higher drift rates (p=.011) and lower thresholds (p<.001), whereas larger expected penalties promoted higher thresholds (p<.001) and significant increase in drift rate (p=.019) (Fig 3A). Extending these past results, we show that the same patterns of parameter adjustments occur when correct responses are reinforced by loss avoidance rather than reward gains (Fig 3B). Larger expected loss avoidance (i.e., higher negative reinforcement) promoted increased drift rate (p=.033) and lower thresholds (p<.001), and larger expected penalties again promoted increased threshold (p<.001) and moderate increases in drift rate (p=.036).

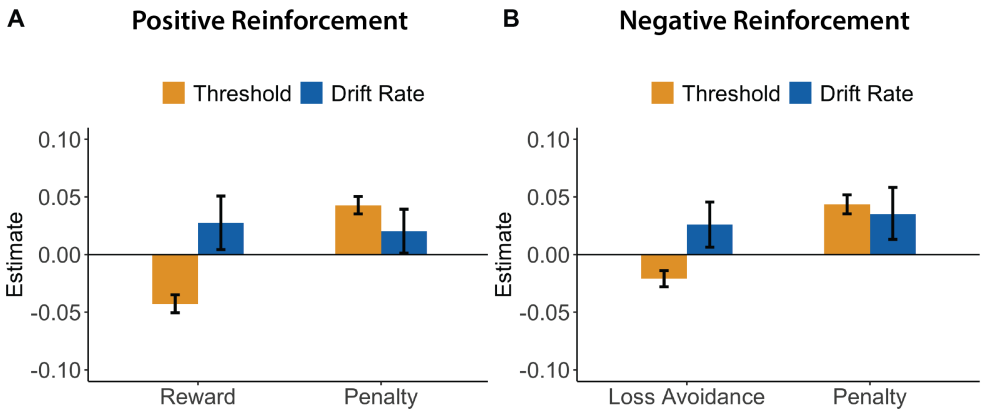

Figure 3: a) *Positive Reinforcement Task* posterior estimates. Higher reward is associated with higher drift rate and lower threshold, whereas higher penalty is associated with higher threshold. b) *Negative Reinforcement Task* posterior estimates. Higher loss avoidance is associated with higher drift rate and lower threshold, whereas higher penalty is associated with higher threshold and drift rate. In both plots, error bars indicate 95% credible intervals.

3.2 Reward Rate Model Predictions and Empirical Results

The patterns of incentive-related DDM parameter adjustments we observed are consistent with predictions of our reward-rate-based model of control allocation. The model predicts that people should respond to higher levels of reinforcement by increasing their drift rate and decreasing their threshold, whereas they should respond to higher levels of punishment by primarily increasing their threshold (Fig. 4A). The distribution of parameter estimates we observed across our task conditions (Fig. 3) matches these predictions, both when varying positive reinforcement (Fig. 4B) and negative reinforcement (Fig. 4C).

Finally, we sought to confirm that our model can not only predict qualitative patterns of drift rate and threshold adjustments, but can also use the joint configuration of these DDM parameters in a given task condition to estimate (infer) the levels of reinforcement (R) and punishment (P) that the participant was reacting to on those trials. Replicating Leng et al., we found that (log-transformed) R_{pos} estimates from the *Positive Reinforcement Task* were inferred to be significantly higher for high-reward relative to low-reward intervals (repeated-measures ANOVA $t_{(214)} = 5.41$, $p < 0.001$), and P was inferred to be significantly higher on high-punishment relative to low-punishment intervals ($t_{(214)} = 21.99$, $p < 0.001$). Critically, when performing the same analysis for the *Negative Reinforcement Task*, we found that our model similarly inferred that R_{neg} was higher for high-loss-avoidance relative to low-loss-avoidance intervals ($t_{(216)} = 8.91$, $p < 0.001$). P was again inferred to be higher for high-punishment relative to low-punishment intervals ($t_{(216)} = 26.84$, $p < 0.001$).

In both tasks, we observed a significant interaction between R and P, such that differences in P estimates across incentive levels were greater than estimates of the corresponding R estimates, both for positive reinforcement, R_{pos} ($F_{(1,214)}$ = 10534.37, $p<0.001$) and negative reinforcement, R_{neg} ($F_{(1,216)}$ = 11708.53, $p< 0.001$).

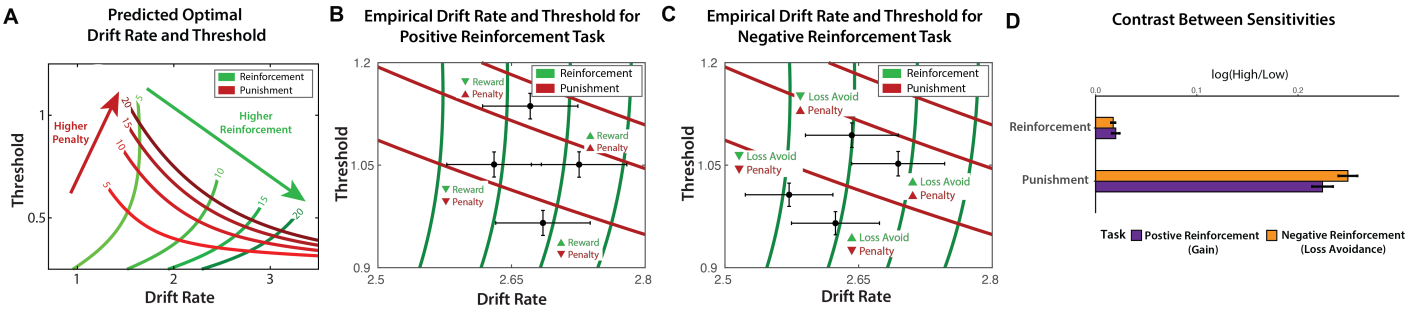

Figure 4: A) Predicted optimal drift rate and threshold estimates (adapted from Leng et al., 2021). B-C) Empirical drift rate and threshold values from our *Positive Reinforcement Task* (B) and *Negative Reinforcement Task* (C). D) Model-inferred differences in R for high vs. low reinforcement and P for high vs. low punishment, across tasks. Error bars reflect s.e.m.

4 Discussion

Our findings bolster prior work within the domain of positive reinforcement and provide evidence that this reward-rate-optimal model extends to characterize negative reinforcement influences on cognitive control allocation. Using a novel paradigm that explicitly delineates the *motivational context* of incentives, we observe a clear dissociation between how incentives promote increased attentional control (i.e., reinforcement) versus response caution (i.e., punishment) in an incentivized cognitive control task. Specifically, our model results suggest, and empirical findings confirm, that incentives that facilitate reinforcement vs. punishment elicit dissociable multivariate influences on control adjustments, revealing a richer understanding of computational mechanisms underlying the interplay between motivation and cognitive control.

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in Two-Alternative Forced-Choice Tasks. *Psychological Review*, *113*(4), 700–765. doi: 10.1037/0033-295x.113.4.700

Botvinick, M. M., & Braver, T. (2015). Motivation and Cognitive Control: From Behavior to Neural Mechanism. *Annual Review of Psychology*, *66*(1), 83–113. doi: 10.1146/annurev-psych-010814-015044

Braem, S., Duthoo, W., & Notebaert, W. (2013). Punishment sensitivity predicts the impact of punishment on cognitive control. *PloS one*, *8*(9), e74106. doi: 10.1371/journal.pone.0074106

Chiew, K. S., & Braver, T. S. (2016). Reward Favors the Prepared: Incentive and Task-Informative Cues Interact to Enhance Attentional Control. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(1), 52–66. doi: 10.1037/xhp0000129

Cubillo, A., Makwana, A. B., & Hare, T. A. (2019). Differential modulation of cognitive control networks by monetary reward and punishment. *Social Cognitive and Affective Neuroscience*, *14*(3), nsz006–. doi: 10.1093/scan/nsz006

Danielmeier, C., & Ullsperger, M. (2011). Post-Error Adjustments. *Frontiers in Psychology*, *2*, 233. doi: 10.3389/fpsyg.2011.00233

Frömer, R., Lin, H., Wolf, C. K. D., Inzlicht, M., & Shenhav, A. (2021). Expectations of reward and efficacy guide cognitive control allocation. *Nature Communications*, *12*(1), 1030. doi: 10.1038/s41467-021-21315-z

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

Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A Computational Model of Control Allocation based on the Expected Value of Control. In *Reinforcement learning and decision making*.

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

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

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., Leng, X., Shenhav, A., & Braver, T. S. (2022). Aversive motivation and cognitive control. *Neuroscience & Biobehavioral Reviews*, *133*, 104493. doi: 10.1016/j.neubiorev.2021.12.016