Title
Dissociable influences of reward and punishment on adaptive cognitive control

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors
Leng, Xiamin
Ritz, Harrison
Yee, Debbie
et al.

Publication Date
2020

Peer reviewed

Dissociable influences of reward and punishment on adaptive cognitive control

Xiamin Leng[1,*], Harrison Ritz[1], Debbie Yee[1], and Amitai Shenhav[1]

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

[*]Corresponding Author: xiamin_leng@brown.edu

Abstract

When deciding how to allocate cognitive control to a given task, people must consider both positive outcomes (e.g., praise) and negative outcomes (e.g., admonishment). However, it is unclear how these two forms of incentives differentially influence the amount and type of cognitive control a person chooses to allocate. To address this question, we had participants perform a self-paced incentivized cognitive control task, varying the magnitude of reward for a correct response and punishment for an incorrect response. Formalizing control allocation as a process of adjusting parameters of a drift diffusion model (DDM), we show that participants engaged in different strategies in response to variability in reward (adjusting drift rate) versus punishment (adjusting response threshold). We demonstrate that this divergent set of strategies is optimal for maximizing reward rate while minimizing effort costs. Finally, we show that these dissociable patterns of behavior enable us to estimate the motivational salience of positive versus negative incentives for a given individual.

Keywords: cognitive control; reward; punishment; decision-making; drift diffusion model

Introduction

When performing mentally demanding tasks, people need to decide how to deploy limited cognitive resources to achieve their goals. People are motivated to different degrees by the prospect of achieving a positive outcome versus avoiding a negative outcome (Lewin, 1935; Atkinson & Feather, 1966). For example, some students study hard to get praised by their parents while others do so to avoid embarrassment. The overall salience of these incentives will determine when and how a given person decides to invest cognitive control (Botvinick & Braver, 2015), including when they choose to disengage from effortful tasks (Wrosch, Scheier, Carver, & Schulz, 2003). While a great deal is known about how people adjust cognitive control in response to varying levels of potential reward (Yee & Braver, 2018), much less is known about how they do so in response to varying levels of potential punishment, nor what types of control allocation strategies are most adaptive under these two conditions.

Previous research has examined how control allocation varies as a function of the reward for performing well at a task, and demonstrated that participants generally perform better when offered greater reward (Braver et al., 2014; Krebs & Woldorff, 2017). For instance, when the reward for a cognitive control task (e.g., Stroop) is contingent on both speed and accuracy, participants are faster and/or more accurate as potential rewards increase (Krebs, Boehler, & Woldorff, 2010; Chiew & Braver, 2016; Froemer, Lin, Dean Wolf, Inzlich, & Shenhav, 2020). However, as this example demonstrates, different forms of control adjustments can produce different types of performance improvements (e.g., differentially prioritizing speed vs. accuracy). Past work has not tested whether the same types of control adjustment are favored when participants are incentivized to avoid poor performance versus achieve good performance.

To understand how people vary their control allocation across different forms of incentives, it is equally critical to understand why they do so. Recent theoretical work provides guidance in addressing this question. For instance, normative accounts of effort allocation propose that animals and humans vary the intensity of their effort to maximize their net reward per unit time (reward rate; Niv, Daw, Joel, & Dayan, 2007; Boureau & Dayan, 2011; Otto & Daw, 2019). Applying such theories to the specific domain of mental effort (i.e., cognitive control) allocation, the Expected Value of Control (EVC) model propose that people allocate the type and amount of cognitive control that maximizes the overall rate of expected rewards while minimizing expected effort costs (Shenhav, Botvinick, & Cohen, 2013; see also Manohar et al., 2015).

The EVC model has been successful at accounting for how people vary the intensity of a particular type of control to achieve greater rewards (Musslick, Shenhav, Botvinick, & Cohen, 2015; Lieder, Shenhav, Musslick, & Griffiths, 2018), but limitations in existing data have prevented it from describing how the type of control being allocated should depend on the type of incentive. Aside from the dearth of research on how people adjust control to positive versus negative incentives, a second critical limitation is that most existing studies examine how performance varies over a fixed set of trials (e.g., 200 total trials that must be completed over the course of an experiment). An appropriate test of normative predictions of the reward rate maximization inherent to EVC (and similar models) requires examining how performance varies when participants are allowed to perform as much or as little of the task as they would like over a fixed duration.

To address these open questions, we developed a novel task that measures cognitive control allocation over a fixed time interval. Participants performed a control-demanding task under different incentive types (reward vs. punishment) and incentive magnitudes (small vs. large), and we measured how participants adapted their cognitive control (e.g., prioritizing speed, accuracy, or both) to optimize their subjective reward rates. People demonstrated distinct patterns of cognitive control allocation for rewards versus punishments. With increasing reward, participants were faster while maintaining the same level of accuracy (completing more trials overall), whereas with increasing punishment they were overall more accurate but also slower (completing fewer trials overall). To provide a normative account of these interactions, we merged properties of existing models of reward rate optimization and control allocation (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Manohar et al., 2015; Musslick et al., 2015; Simen et al., 2009), modeling different types of control allocation as adjustments of different parameters in a Drift Diffusion Model (DDM; Ratcliff & McKoon, 2008). We show both normatively and empirically that evidence accumulation rate (a potential proxy for attentional focus) selectively increases with increasing potential reward, whereas response thresholds selectively increase with increasing potential punishment. Finally, we used this modified reward rate model to estimate the individual differences in sensitivity to reward and punishment based upon unique behavioral profiles, providing a compelling novel approach for inferring how people evaluate costs and benefits when deciding when and how much to allocate cognitive control.

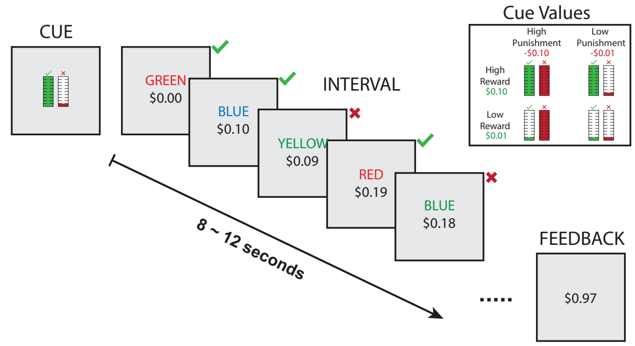


Figure 1: Task paradigm. At the start of each interval, a visual cue indicates the level of reward and punishment for that interval. Participants can complete as many Stroop trials as they want within that interval. The cumulative reward over a given interval is tracked at the bottom of the screen. Correct responses increase this value while incorrect responses decrease this value. Participants are told how much they earned at the end of each interval. The upper right inset shows the cues across the four conditions.

Incentivized Cognitive Control Task

We designed a new task to investigate cognitive control allocation in a self-paced environment (Figure 1). During this task, participants are given fixed time intervals (8-12s) to perform a cognitively demanding task (Stroop task), in which they have to name the ink color of a color word. Participants could perform as many Stroop trials as they wanted during each interval, with a new trial appearing immediately after each response. Since the duration of intervals was varied across the session, participants were discouraged from developing a trial-counting strategy (e.g., performing 10 responses per interval).

Participants were instructed that they would be rewarded for correct responses and penalized for incorrect responses. At the start of each interval, a visual cue indicated the level of reward and punishment associated with their responses in the subsequent interval. There are four distinct conditions in the experiment: high-reward/high-punishment (+10¢, -10¢), high-reward/low-punishment (+10¢, -1¢), low-reward/high-punishment (+1 ¢, -10¢), and low-reward/low-punishment (+1¢, -1¢) (Figure 1). During the interval, participants could complete as many Stroop trials as they would like. Below each Stroop stimulus, a tracker indicated the cumulative amount of monetary reward within that interval. After each interval, participants were informed how much they earned. The experiment was implemented within the PsiTurk framework (Gureckis et al., 2016) and the data was collected on the Mechanical Turk platform.

Behavioral Results

We collected data from 36 participants, but four participants were excluded due to poor performance (with mean accuracy below 60% or mean reaction time outside of 3 standard deviations of the mean). The final dataset consisted of 32 participants (10 F; Age: 35 ± 10 years).

Table 1: Mixed model results for normalized correct responses

Predictors	Estimates	t	df	p
Intercept	1.14	34.03	32	< **0.001**
Age	−0.04	−0.69	29	0.265
Female - Male	0.08	2.13	29	**0.044**
High - Low Punishment	−0.03	−4.73	29	< **0.001**
High - Low Reward	0.04	5.606	30	< **0.001**
Mean Congruency	−0.02	−3.586	36	< **0.001**
Reward× Punishment	0.01	2.168	41	**0.011**

The primary measurement of task performance is the number of correct responses within each interval divided by the duration of the interval (i.e., normalized correct responses). We fitted a linear mixed model (lme4 package in R; Bates, Mächler, Bolker, & Walker, 2015) to estimate the normalized correct responses as a function of contrast-encoded reward and punishment levels (High Reward - Low Reward, High Punishment - Low Punishment) as well as their interaction, controlling for age, gender and congruency effect, and using models with maximally specified random effects (Barr, Levy, Scheepers, & Tily, 2013) (Table 1).

(a)

(b)

(c)

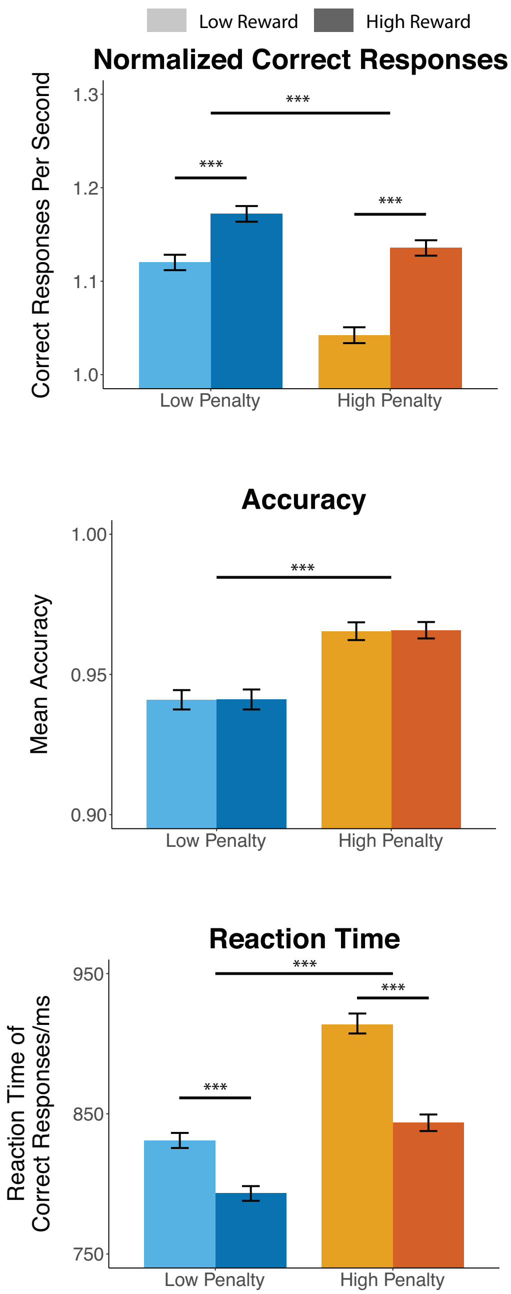


Figure 2: Behavioral results. Effect of reward and punishment on a) the number of correct responses completed in an interval (normalized by interval duration); b) trial-level accuracy and c) trial-level response time (correct responses only). ***: $p < 0.001$. Error bars reflect s.e.m..

Participants completed more trials with increasing potential reward, and fewer trials with increasing potential punishment ($p < 0.001$; Figure 2a). We observed a significant interaction between reward and punishment ($p = 0.011$) whereby the effect of reward level on performance was enhanced in high-punishment compared to low-punishment intervals. We did not find significant interaction between reward or penalty and mean congruency.

Additional analyses revealed that rewards versus punishments exerted distinct influences on speed versus accuracy. Accuracy was significantly higher with increasing potential punishment ($p < 0.001$), but did not vary with potential reward ($p = 0.932$; Figure 2b). Response time, on the other hand, did vary significantly with both types of incentives, but in opposite directions (Figure 2c). Consistent with the differences trial completion rate by condition (Figure 2a), participants were faster to respond with increasing reward ($p < 0.001$) but slower to respond with increasing punishment ($p < 0.001$). Together, these data reveal that participants applied distinct strategies under different incentive conditions.

Reward Rate Optimal Control Allocation: Normative Predictions

An existing reward-rate optimization model shows that, when deciding which strategy to apply in a given task, the normative estimate of reward rate will critically depend on how an individual weighzs the benefit of the reward for a correct response versus the cost of being punished for an error (Bogacz et al., 2006; Krueger et al., 2017). This subjective reward rate RR is expressed as

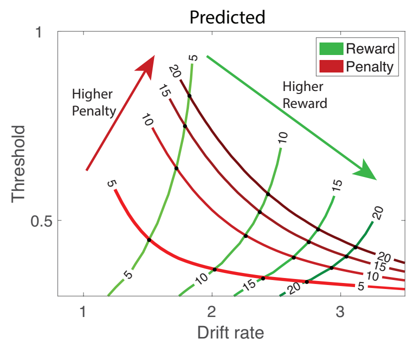
$$RR = \frac{R \times (1 - ER) - P \times ER}{DT + NDT} \tag{1}$$

where ER is error rate, DT is decision time, and NDT is non-decision time (e.g., time to execute a motor response). R and P indicate the weights for reward and punishment, respectively. The decision procedure in Stroop task can be characterized as a drift diffusion process in which evidence is accumulated toward one response until the accumulated evidence reaches a threshold (Musslick et al., 2015). The expectations of ER and DT depends on the drift rate (the speed of evidence accumulation) and threshold (Bogacz et al., 2006).

To correctly respond on a Stroop trial (i.e., name stimulus color), participants need to recruit cognitive control to overcome the automatic tendency to read the word. We first assume that participants performing our task choose between adjusting two strategies for achieving this goal: (1) increasing attentional focus on the Stroop stimuli (resulting in increased drift rate toward the correct response), and (2) increasing their threshold to require more evidence accumulation before responding. Second, we assume that participants seek to identify the combination of these two DDM parameters that maximizes reward rate. Third, we assume that increasing drift rate incurs a cost, which participants seek to minimize. The inclusion of this cost term is motivated by previous psychological and neuroscientific research (Shenhav et al., 2017) and by its sheer necessity for constraining the model from seeking implausibly high values of drift rate (i.e., as this cost approaches zero, the reward rate maximizing drift rate approaches infinity). A quadratic cost term was chosen based on previous work (v^2; Manohar et al., 2015; Musslick et al., 2015) and additional analyses showing that this outperforms a linear term (not shown here)

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

(a)



(b)

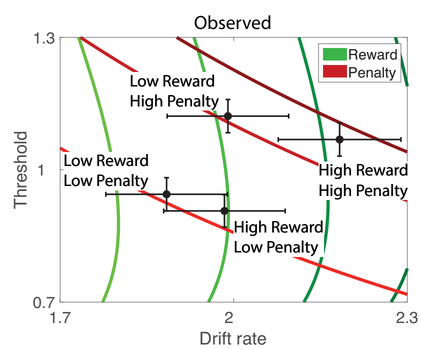


Figure 3: (a) Theoretical optimal threshold and drift rate under different pairs of normative values of reward (R) and punishment (P) weights. Each dot indicates the optimal combination of drift rate and threshold under a pair of R and P. Darker colors represents larger values of reward or penalty. (b) Observed combinations of average drift rate and threshold for the four experimental conditions (based on HDDM fits) follow the reward and punishment gradients predicted by our model. Each dot represents the combination of empirical values of R and P. Error bars reflect s.d..

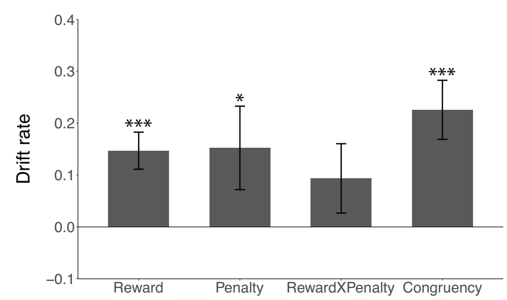
Under the assumption that participants implement adaptive cognitive strategies to maximize RR, we can generate predictions regarding the optimal settings for drift and threshold under different reward (R) and punishment (P) conditions. For different R and P values, we numerically identified the drift rate and threshold pair that would maximize reward rate (Figure 3a). In this model, the optimal drift rate and threshold is determined by the ratios between R, P and E. For these calculations, the magnitude of effort costs is held constant (E=1), putting reward and punishment into units of effort.

As R increases, the model suggests that the optimal strategy is to increase drift rate and reduce threshold. As P increases, the optimal strategy is to primarily increase threshold and slightly increase drift rate. These findings indicate that the weights for rewards and punishments jointly modulate the optimal strategy for allocating cognitive control, and that these two types of incentives focus on distinct aspects of the strategy. Specifically, they predict that that people will tend to increase drift rate the more they value receiving a reward for a correct response, whereas threshold will be modulated as a function of how much they value receiving a reward for a correct response (decreased threshold) and receiving a punishment for an incorrect response (increased threshold).

Reward Rate Optimal Control Allocation: Empirical Evidence

We next sought to test whether performance on our task was consistent with the predictions of our normative model. We fit accuracies and RTs across the different task conditions with a Hierarchical Drift Diffusion Model (HDDM) (Wiecki, Sofer, & Frank, 2013), which allowed us to derive estimates of how a participant's drift and threshold varied across different levels of reward and punishment.

(a)



(b)

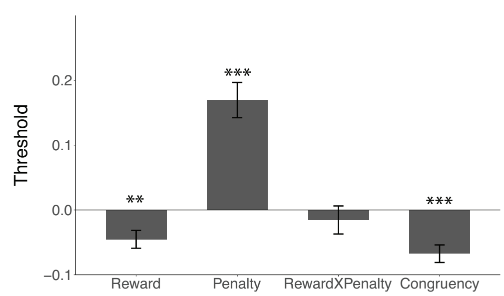


Figure 4: Fitted coefficients of reward, punishment, reward/punishment interaction and congruency on a) drift rate and b) threshold. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$. Error bars reflect s.d..

Consistent with predictions of our reward-rate-optimal DDM, drift rate increased with larger rewards as well as larger punishment (Figures 3b and 4). This suggests that participants were motivated to increase their overall attention allocation with increasing incentive magnitude. Also consistent with our normative predictions, we found that reward and punishment exhibited dissociable influences on threshold, with higher rewards promoting a lower threshold and higher punishment promoting a higher threshold. These findings control for the effect of congruency on DDM parameters (with incongruent trials being associated with lower drift rate and higher threshold). Thus, our empirical findings are consistent with the prediction that participants are optimizing reward rate, accounting for potential rewards, potential punishments, and effort costs.

Model-Based Estimates from Behavior Recover Original Values of Reward and Punishment

Participants' subjective valuation of incentives is a latent variable that must be inferred from task performance. Since we have a process model for mapping incentives onto control configuration (i.e., reward-rate optimization), we can 'reverse-engineer' participants' subjective valuation of reward and punishment based on their DDM parameters (e.g., drift rate and threshold; Figure 3b). Having identified condition-specific settings of drift rate and threshold for each participant, and moreover showing that they fit a qualitative pattern consistent with prediction of normative control adaptation, we can further use these DDM estimates to infer individualized subjective weights of reward value (R) and punishment value (P) for each of the four task conditions. Critically, this parameter recovery validates this approach for inferring individualized latent subjective valuation of reward and punishment incentive effects on adaptive cognitive control.

Here, we used performance-driven model estimates to 'reverse engineer' the individualized subjective weights of reward (R) and punishment (P) across the four task conditions. For each task condition, we first estimated the drift rate (v) and threshold (a) for each individual. We then calculated the partial derivatives of reward rate (RR) with respect to these condition-specific estimates of v and a. By setting these derivatives to 0 (i.e., optimizing the reward-rate equation), we can calculate the subjective weights of reward (R) and punishment (P) that make the estimated (v,a) the optimal strategy. This workflow can be summarized as follows:

$$DDM \rightarrow (v,a) \rightarrow \begin{cases} \frac{\partial RR}{\partial v} = 0 \\ \\ \frac{\partial RR}{\partial a} = 0 \end{cases} \rightarrow (R,P)_{optimal}$$

A repeated-measures ANOVA on our estimates of R and P (log-transformed) revealed a main effect of incentive magnitude ($F(1,251) = 15.96, p < 0.001$), with larger R on high-reward intervals ($t(63) = 7.59, p < 0.001$) and larger P on high-punishment intervals ($t(63) = 6.24, p < 0.001$). We also observed a main effect of valence, such that estimates of P were higher than estimates of R ($F(1,251) = 15.96, p < 0.001$). The ANOVA also revealed a significant interaction between valence and magnitude ($F(1,251) = 10.28, p = 0.0015$), such that P estimates differed more across punishment levels than R estimates differed across reward levels (Figure 5). These asymmetric effects of rewards and punishment on reward rate are consistent with research on loss aversion (Kahneman & Tversky, 1979) and error aversion (Hajcak & Foti, 2008).

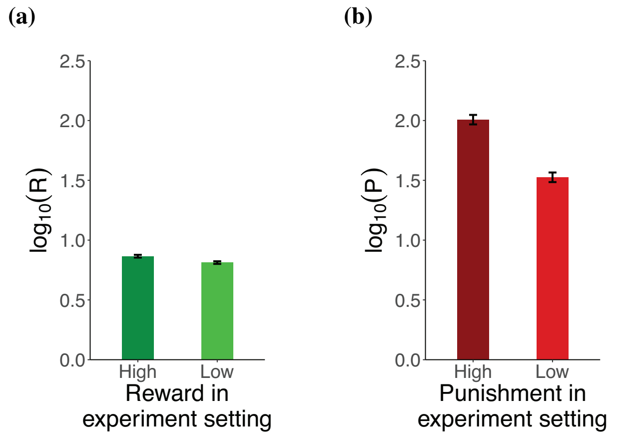


Figure 5: Log transformed weights for a) high vs low reward and b) high vs low punishment. We observed a significant interaction between valence and magnitude, such that the difference in the log weights between high and low punishment is significantly higher than the difference between high and low reward. Error bars reflect s.e.m..

Discussion

We investigated common and divergent influences of reward versus punishment on cognitive control allocation, and the normative basis for these incentive-related control adjustments. Participants performed a self-paced cognitive control task that offered the promise of monetary rewards for correct responses and monetary losses for errors. As reward increased, they responded faster and were therefore able to complete more trials as reward. As punishment increased, they responded slower and therefore completed fewer trials but were overall more accurate. We showed that these different patterns of incentive-related performance could be accounted for by a combination of two distinct strategies (adjustment of the strength of attention vs. response threshold), which are differentially optimal (i.e., reward rate maximizing) in response to these two types of incentives.

Our findings build on past research on reward rate maximization, which has shown that people alter their behavior and their cognitive strategies to maximize their subjective reward per unit time (Bogacz et al., 2006; Lieder et al., 2018; Otto & Daw, 2019). Our current experiment builds on this research in several important ways. First, we apply it to performance in a self-paced variant of a cognitive control task. Second, we model and experimentally manipulate the incentive value for a correct versus incorrect response. Third, to account for well-known costs of cognitive effort (Shenhav et al., 2017), we modeled such effort costs as scaling quadratically with changes in drift rate (Manohar et al., 2015). Finally, we used our model to perform reverse inference on our data, identifying the subjective weights of incentives that gave rise to performance on a given trial.

Our theoretical and empirical findings show that adjustments of threshold and drift can vary as a function of the task incentives, which then drive adaptive adjustments in cognitive control. Notably, achieving this result required us to build in the assumption that increases in drift rate incur a cost. Without this assumption – which is grounded in past research on mental effort (Manohar et al., 2015; Shenhav et al., 2017) – it would always be adaptive for an individual to maintain a maximal drift rate across conditions, as this would guarantee consistently fast and accurate responding. However, while it is clear that some form of cost function is necessary to constrain drift rate, follow-up work is needed to further characterize that function and the extent of its nonlinearity. We have also left open the question of whether and how a cost function applies to increases in response threshold. While there is reason to believe that threshold adjustments may incur analogous effort costs to attentional adjustments, in part given the control allocation mechanisms they share (Musslick et al., 2015), threshold adjustments carry an inherent cost in the form of a speed-accuracy tradeoff. It therefore wasn't strictly necessary to incorporate an additional effort cost for threshold in the current simulations, though it is possible such a cost would have further improved model predictions. Future work will investigate the boundary conditions of when a cost function can provide additional explanatory power to incentivized cognitive control.

Our combined theoretical and empirical approach enabled us to quantify the value participants placed on expected rewards and punishments, based only on their task performance. Our results showed that people weighed punishments more heavily than rewards, despite the currency being equivalent (i.e., amounts of monetary gain vs. loss). This finding is consistent with past work on loss aversion (Kahneman & Tversky, 1979), and more generally with the findings that distinct neural circuits are specialized for processing appetitive versus aversive outcomes (Bissonette, Gentry, Padmala, Pessoa, & Roesch, 2014; Pessiglione & Delgado, 2015). Critically, the current approach and findings hold promise for research into individual differences in sensitivity to rewards versus punishments. Not only can this method help to infer these sensitivity parameters for a given individual implicitly (i.e., based on task performance rather than self-report), it can also provide valuable insight into the cognitive and computational mechanisms that underpin adaptive control adjustments, and when and how they become maladaptive (e.g., in mood and anxiety disorders).

Overall, our task captured the influence of reward and punishment on self-paced cognitive control allocation. We investigated how individuals adjust their strategy for allocating cognitive control in a self-paced setting, as well as how monetary incentives are translated into subjective values to influence task performance. These results provide an important foundation for the computational mechanisms underpinning divergent strategies for optimizing reward rate. We present a novel adaptation to an existing reward-rate optimization model to account for the cost of cognitive control in motivated task performance. This is the first demonstration of a quantitative approach to account for reward, punishment, and effort cost on the adaptation of cognitive control, which has also been additionally validated by experimental data. These findings lend support to the Expected Value of Control model (Shenhav et al., 2013) and, critically, provide greater specificity to the computational and cognitive mechanisms underlying adaptive cognitive control.

References

Atkinson, J. W., & Feather, N. T. (1966). *A theory of achievement motivation* (Vol. 66). Wiley New York.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013, April). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.*, *68*(3).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, Articles*, *67*(1), 1–48. Retrieved from `https://www.jstatsoft.org/v067/i01` doi: 10.18637/jss.v067.i01

Bissonette, G. B., Gentry, R. N., Padmala, S., Pessoa, L., & Roesch, M. R. (2014). Impact of appetitive and aversive outcomes on brain responses: linking the animal and human literatures. *Front. Syst. Neurosci.*, *8*, 24.

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. *Psychol. Rev.*, *113*(4), 700–765.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annu. Rev. Psychol.*, *66*, 83–113.

Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, *36*(1), 74–97.

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., ... Others (2014). Mechanisms of motivation–cognition interaction: challenges and opportunities. *Cogn. Affect. Behav. Neurosci.*, *14*(2), 443–472.

Chiew, K. S., & Braver, T. S. (2016, January). Reward favors the prepared: Incentive and task-informative cues interact to enhance attentional control. *J. Exp. Psychol. Hum. Percept. Perform.*, *42*(1), 52–66.

Froemer, R., Lin, H., Dean Wolf, C. K., Inzlich, M., & Shenhav, A. (2020). When effort matters: Expectations of reward and efficacy guide cognitive control allocation. *biorXiv*.

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., … Chan, P. (2016). psiturk: An open-source framework for conducting replicable behavioral experiments online. *Behav. Res. Methods*, *48*(3), 829–842.

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological science*, *19*(2), 103–108.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the stroop task. *Cognition*, *117*(3), 341–347.

Krebs, R. M., & Woldorff, M. G. (2017). Cognitive control and reward.

Krueger, P. M., van Vugt, M. K., Simen, P., Nystrom, L., Holmes, P., & Cohen, J. D. (2017). Evidence accumulation detected in bold signal using slow perceptual decision making. *Journal of neuroscience methods*, *281*, 21–32.

Lewin, K. (1935). A dynamic theory of personality (dk adams & ke zaner, trans.). *NY: McGraw Hill*.

Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018). Rational metareasoning and the plasticity of cognitive control. *PLoS Comput. Biol.*, *14*(4), e1006043.

Manohar, S. G., Chong, T. T.-J., Apps, M. A. J., Batla, A., Stamelou, M., Jarman, P. R., … Husain, M. (2015). Reward pays the cost of noise reduction in motor and cognitive control. *Curr. Biol.*, *25*(13), 1707–1716.

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 *The 2nd multidisciplinary conference on reinforcement learning and decision making.*

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520.

Otto, A. R., & Daw, N. D. (2019). The opportunity cost of time modulates cognitive effort. *Neuropsychologia*, *123*, 92–105.

Pessiglione, M., & Delgado, M. R. (2015). The good, the bad and the brain: neural correlates of appetitive and aversive values underlying decision making. *Current opinion in behavioral sciences*, *5*, 78–84.

Ratcliff, R., & McKoon, G. (2008, April). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.*, *20*(4), 873–922.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240.

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. *Annu. Rev. Neurosci.*, *40*, 99–124.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *J. Exp. Psychol. Hum. Percept. Perform.*, *35*(6), 1865–1897.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical bayesian estimation of the Drift-Diffusion model in python. *Front. Neuroinform.*, *7*, 14.

Wrosch, C., Scheier, M. F., Carver, C. S., & Schulz, R. (2003, January). The importance of goal disengagement in adaptive Self-Regulation: When giving up is beneficial. *Self Identity*, *2*(1), 1–20.

Yee, D. M., & Braver, T. S. (2018). Interactions of motivation and cognitive control. *Current opinion in behavioral sciences*, *19*, 83–90.