Dissociable influences of reward and punishment on adaptive cognitive control

Xiamin Leng*, Debbie Yee, Harrison Ritz, Amitai Shenhav

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

*Correspondence should be addressed to Xiamin Leng (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 achieve task goals. 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 reward (primarily adjusting drift rate) versus punishment (primarily 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 infer the motivational salience of positive versus negative incentives for a given individual. These results provide a normative and mechanistic account for how reward and punishment differentially influence the adaptive allocation of cognitive control.

People must regularly decide how much mental effort to invest in a task, and for how long. When doing so, they weigh the costs of exerting this effort against the potential benefits that would accrue as a result (Shenhav et al., 2017). These benefits include not only the positive consequences (e.g., monetary gain) for achieving one's goal but also the negative consequences (e.g., monetary loss) that are avoided. Critically, people are likely to vary in the extent to which they are motivated by the prospect of achieving a positive outcome versus avoiding a negative outcome (Atkinson et al., 1966; Lewin, 1935; Yee & Braver, 2018). For example, some students study diligently to earn praise from 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 mental effort (i.e., engage relevant cognitive control processes; (Botvinick & Braver, 2015), including when they choose to disengage from effortful tasks (Wrosch et al., 2003). While much is known about how people adjust cognitive control in response to varying levels of potential reward (Botvinick & Braver, 2015; Dixon & Christoff, 2012; Yee & Braver, 2018), much less is known about how they similarly adjust to varying levels of potential punishment, nor the types of control allocation strategies that are most adaptive under these two incentive conditions.

Previous research has examined how control allocation varies as a function of the reward for performing well on a task, and demonstrated that participants generally perform better when offered a greater reward (Braver et al., 2014; Cools, 2015; Dixon & Christoff, 2012; Froböse et al., 2018; Krebs et al., 2010; Vassena et al., 2014). For instance, when earning rewards during 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 (Chiew & Braver, 2016; Frömer et al., 2020; Krebs et al., 2010). Thus, it is evident that different forms of control adjustments can differentially impact improvements in task performance (e.g., differentially prioritizing speed vs. accuracy). Past work has demonstrated that people vary in the types of control adjustments they make while performing a cognitive control task (Danielmeier & Ullsperger, 2011; Ritz, DeGutis, Frank, Esterman, & Shenhav, 2020; Ritz & Shenhav, 2019; van Veen et al., 2008). However, it remains unknown whether participants selectively deploy different forms of control adjustments when incentivized under distinct incentive regimes (i.e., to avoid poor performance versus achieve good performance).

Recent theoretical work provides guidance in addressing why people vary their control allocation across different forms of incentives (Musslick et al., 2015). 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; Boureau & Dayan, 2011; Niv et al., 2007; Otto & Daw, 2019). Building on this work, a recent normative model of cognitive control allocation proposes that people allocate the type and amount of cognitive control that maximizes the Expected Value of Control, i.e., the overall rate of expected rewards while minimizing expected effort costs, two quantities that together determine the Expected Value of Control (EVC; Manohar et al., 2015; Shenhav et al., 2013, 2016, 2017).

The EVC model has been successful at accounting for how people vary the intensity of a particular type of control (e.g., attention to a target stimulus/feature) to achieve greater rewards

(Lieder et al., 2018; Musslick et al., 2015). However, limitations in existing data have prevented EVC from describing how the type of control being allocated should depend on the type of incentive. One limitation, noted above, is 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 completed over the course of an experiment). A stronger test of reward rate maximization, and one that is arguably more analogous to real-world effort allocation, would allow participants to perform as much or as little of the task as they like over a fixed duration.

In the current study, we developed a novel task in which participants perform consecutive trials of a control-demanding task (the Stroop task) over a fixed time interval. We examined how the amount and type(s) of control allocated to this task varied under different incentive types (reward vs. punishment) and different magnitudes of those incentives (small vs. large). Participants demonstrated distinct patterns of task performance in the two incentive conditions: faster responses for increasing rewards, slower but more accurate responses for increasing punishment). We show that these patterns are consistent with normative predictions of a control allocation model that maximizes reward rate while minimizing effort costs. The model predicts, and data confirm, that rewards versus punishments favor divergent control strategies, selectively altering information processing versus response threshold. These divergent predictions further enabled us to "reverse infer" a given participant's sensitivity to reward and punishment based on their unique behavioral profiles. Collectively, this work demonstrates a compelling novel method for inferring variability in how people evaluate costs and benefits when deciding when and how much to allocate cognitive control.

Methods

Participants

We collected 32 participants (Gender: 31% Female; Age: 35±10 years) online through Amazon Mechanical Turk after excluding 4 participants with mean accuracy below 60% and mean reaction time outside of 3 standard deviations of the mean reaction time of all the participants.

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 (e.g., 10 seconds) 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. To discourage participants from developing a trial-counting strategy (e.g., aiming to complete 10 responses per interval), the duration of intervals varied across the session (i.e., ranging from 8 to 12 seconds).

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. We varied reward for correct responses (+1 cent or +10 cents) and punishment for incorrect responses (-1 cents or -10 cents) within each subject, which leads to four distinct conditions (Figure 1). Each participant performed 20 intervals per condition. 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).

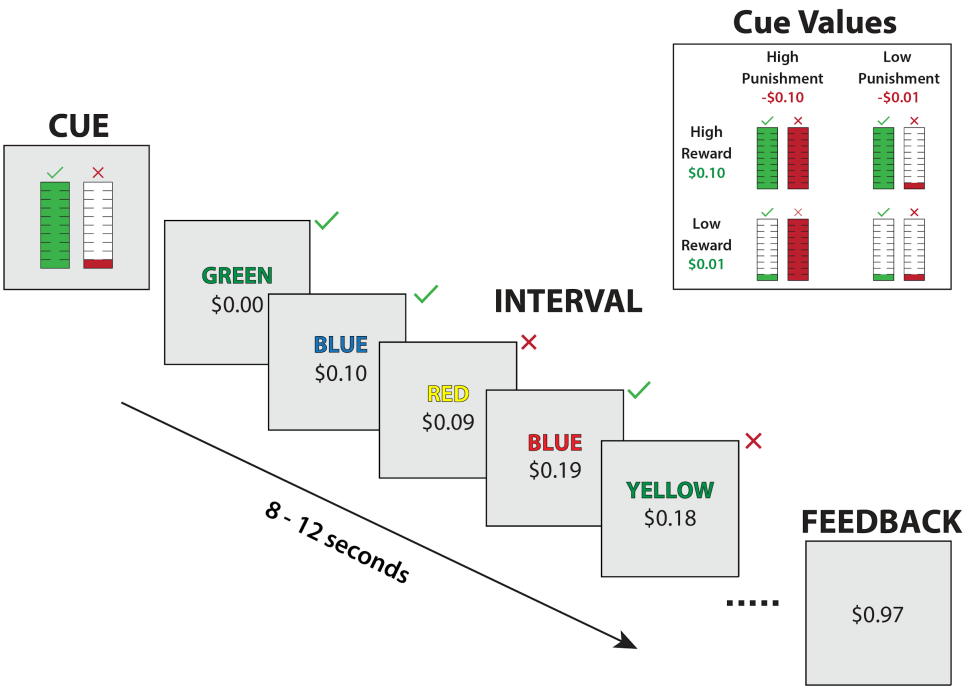

Figure 1. Incentivized Cognitive Control Task paradigm. At the start of each interval, a visual cue indicates the amount of reward (monetary gain) for correct responses and the penalty amount (monetary loss) for incorrect responses within that interval. Participants can complete as many Stroop trials 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. At the end of each interval, participants are told how much they earned. The upper right inset shows the cues across the four conditions.

Analyses

With this paradigm, we can analyze performance at the level of a given interval and at the level of responses to individual Stroop stimuli within that interval. We analyzed participants' interval-level performance by fitting a linear mixed model (lme4 package in R; (Bates et al., 2015) to estimate the correct responses per second as a function of contrast-coded reward and

punishment levels (High Reward= 1, Low Reward = -1, High Punishment=1, Low Punishment=-1) as well as their interaction. The models controlled for age, gender, and proportion of congruent stimuli, and using models with maximally specified random effects (Barr et al., 2013). To understand how the incentive effects on overall performance are composed of the influences on speed and accuracy, we separately fit linear mixed models to trial-wise reaction time (correct responses only) and accuracy, controlling for the stimuli congruency.

We parameterized participants' responses in the task as a process of noisy evidence accumulating towards one of two boundaries (correct vs. error) using the Drift Diffusion Model (DDM). The DDM is a mechanistic model of decision-making that decomposes choices into a set of constituent processes (e.g., evidence accumulation and response thresholding), allowing precise measurement of how different components of the choice process (e.g., RT and accuracy) are simultaneously optimized (Ratcliff & McKoon, 2008).

We characterized the optimal allocation of cognitive control as the maximization of the reward rate (Bogacz et al., 2006) with modification for effort cost. With the modified reward rate, we make predictions about the influences of incentives on control allocation by numerically identifying the optimal drift rate and threshold under varying reward and punishment. To validate our normative prediction, we fit accuracies and RTs across the different task conditions with a DDM (Wiecki et al., 2013), which allowed us to derive estimates of how a participant's drift rate and threshold varied across different levels of reward and punishment. We performed model comparison based on deviance information criterion (DIC; lower is better) to identify the best model for the behavioral data. Based on the assumption that participants' cognitive control allocation optimizes the reward rate, we inferred participants' subjective weights of reward and punishment from the estimated drift rate and threshold.

Results

Behavioral Performance

When participants were expecting a larger reward for each correct response, they completed more trials correctly in a given interval compared to when they were expecting smaller rewards ($p<0.001$; Figure 2A). Variability in punishment magnitude appeared to have the opposite influence on behavior. When participants were expecting a larger punishment for each incorrect response, they completed fewer correct trials in a given interval than when they were expecting smaller punishments ($p<0.001$; Figure 2B). We also observed a significant interaction between reward and punishment ($p=0.029$) whereby the reward-related improvements in interval-level performance were enhanced in high-punishment compared to low-punishment intervals.

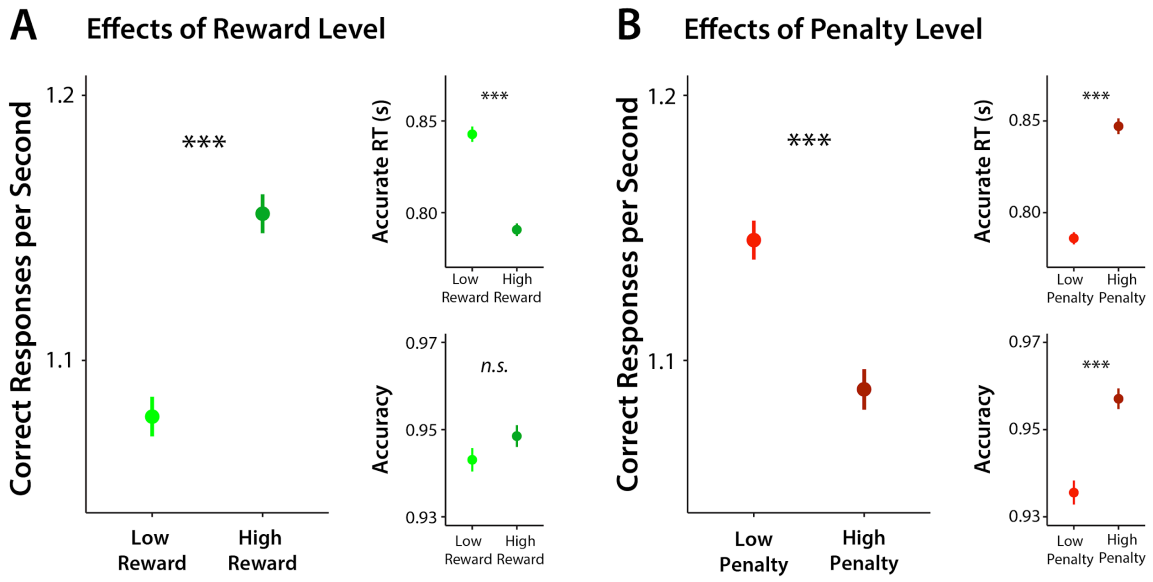

Figure 2. Effects of reward and punishment on overall task performance. A) With increasing expected reward, participants completed more correct responses per second within a given interval (**left**), which reflect faster responding on correct trials (**top right**) without any change in overall accuracy (**bottom right**). **B)** With increasing expected punishment, participants instead completed fewer trials per second over an interval, reflecting slower and more accurate responses. Error bars reflect 95% CI. n.s.: p>0.05; ***: p<0.001

When separately examining how incentives influenced speed and accuracy, we found an intriguing dissociation that helped account for the inverse effects of reward and punishment on the number of correct responses per interval. We found that larger potential rewards induced responses that were faster (p<0.001) but not more or less accurate (p=0.349; Figure 2A, Table 2). By contrast, larger potential punishment induced responses that were slower (p<0.001) but *also* more accurate (p<0.001; Figure 2B). These results control for trial-to-trial differences in congruency, which, as expected, revealed faster (p<0.001) and more accurate (p=0.017) responses for congruent stimuli compared to incongruent stimuli. Although there were no significant two-way interactions between incentives and congruency on performance, we observed a significant three-way interaction between reward, penalty, and congruency (p=0.022) specific to accuracy. Together, these data suggest that participants applied distinct strategies for engaging cognitive control across reward and punishment incentives.

Table 1. Mixed Model Results for Correct Responses per Second

	Correct Responses Per Second		
Predictors	*Estimates*	*SE*	*p-value*
Age	-0.020	0.030	0.507
Female - Male	0.085	0.032	**0.008****
High Penalty - Low Penalty	-0.029	0.005	**<0.001*****
High Reward - Low Reward	0.037	0.007	**<0.001*****
Average Congruency	-0.013	0.005	**0.004****
Reward X Penalty	0.010	0.004	**0.029**

Table 2. Mixed Model Results for Log-Transformed Reaction Time and Accuracy

	Log-transformed RT			**Accuracy**		
Predictors	*Estimates*	*SE*	*p-value*	*Odds Ratios*	*SE*	*p-value*
Age	0.009	0.008	0.243	0.937	0.111	0.559
Female - Male	-0.023	0.008	**0.003****	1.244	0.117	0.063
High Penalty - Low Penalty	0.015	0.002	**<0.001*****	1.379	0.051	**<0.001*****
High Reward - Low Reward	-0.012	0.002	**<0.001*****	1.037	0.037	0.331
Trial Congruency (Cong-Incong)	-0.020	0.002	**<0.001*****	1.107	0.042	**0.017***
Reward X Penalty	-0.003	0.001	**0.016***	1.015	0.038	0.702
Penalty X Congruency	0.001	0.001	0.404	1.040	0.035	0.258
Reward X Congruency	-0.001	0.001	0.471	1.043	0.035	0.224
Reward X Penalty X Congruency	0.001	0.001	0.512	1.082	0.034	**0.022***

Reward Rate-Optimal Control Allocation: Normative Predictions

To generate predictions about performance on the Stroop task, we parameterized the tasks as a process of noisy evidence accumulating towards one of two boundaries (correct vs. error), using the *drift diffusion model* (DDM; Ratcliff & McKoon, 2008; Musslick et al., 2015). Two of the parameters of the DDM that determine performance on a given trial are the rate of evidence accumulation (*drift rate*, v) and the decision threshold (a). As the drift rate increases, the likelihood of a correct response increases (error rate decreases), and responses are *faster*. As the threshold increases, responses are also more likely to be correct but are *slower* (Figure 3A; Bogacz et al., 2006).

Previous theoretical and empirical work has shown that participants can adjust parameters of this underlying decision process to maximize the rate at which they are rewarded over the course of an experiment (Bogacz et al., 2006; Simen et al., 2009). This reward rate is

determined by a combination of performance metrics (response time and error rate [ER]) and the incentives for performance (i.e., outcomes for correct vs. incorrect responses):

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

Here, the numerator (expected reward) is determined by the likelihood of a correct response (1-ER), scaled by the reward for a correct response (R), relative to the likelihood of an error (ER), scaled by the associated punishment (P) (Krueger et al., 2017). The denominator (response time) is determined by the time it takes to accumulate evidence for a decision (decision time [DT]) as well as additional time to process stimuli and execute a motor response (non-decision time [NDT]).

To correctly respond to a Stroop trial (i.e., name stimulus color), participants need to recruit cognitive control to overcome the automatic tendency to read the word (Cohen et al., 1990). Building on past work (Bogacz et al., 2006; Krueger et al., 2017; Simen et al., 2009), we can use the reward rate formulation above to identify how participants should normatively allocate control to maximize the reward rate (Figure 3B-C). To do so, we make three key assumptions. First, we assume that participants performing our task choose between adjusting two strategies for increasing their reward rate: (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 maximize reward rate. Third, we assume that increasing the drift rate incurs a nonlinear cost[1], 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, as shown in Figure 3B).

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

In this formula, Erepresent the weight of effort cost. Since the optimal drift rate and threshold are determined by the ratios R/E and P/E, the magnitude of effort costs is held constant ($E = 1$) for the reward rate optimization process, putting reward and punishment into units of effort cost. With this modified form of reward rate, the optimal drift rate is well-constrained (Figure 3C).

[1] While a quadratic cost term was chosen a priori based on previous work (Manohar et al., 2015; Musslick et al., 2015), follow-up analyses (not shown here) indicated that the predictions made by this quadratic function are also more consistent with our data than those for a linear function.

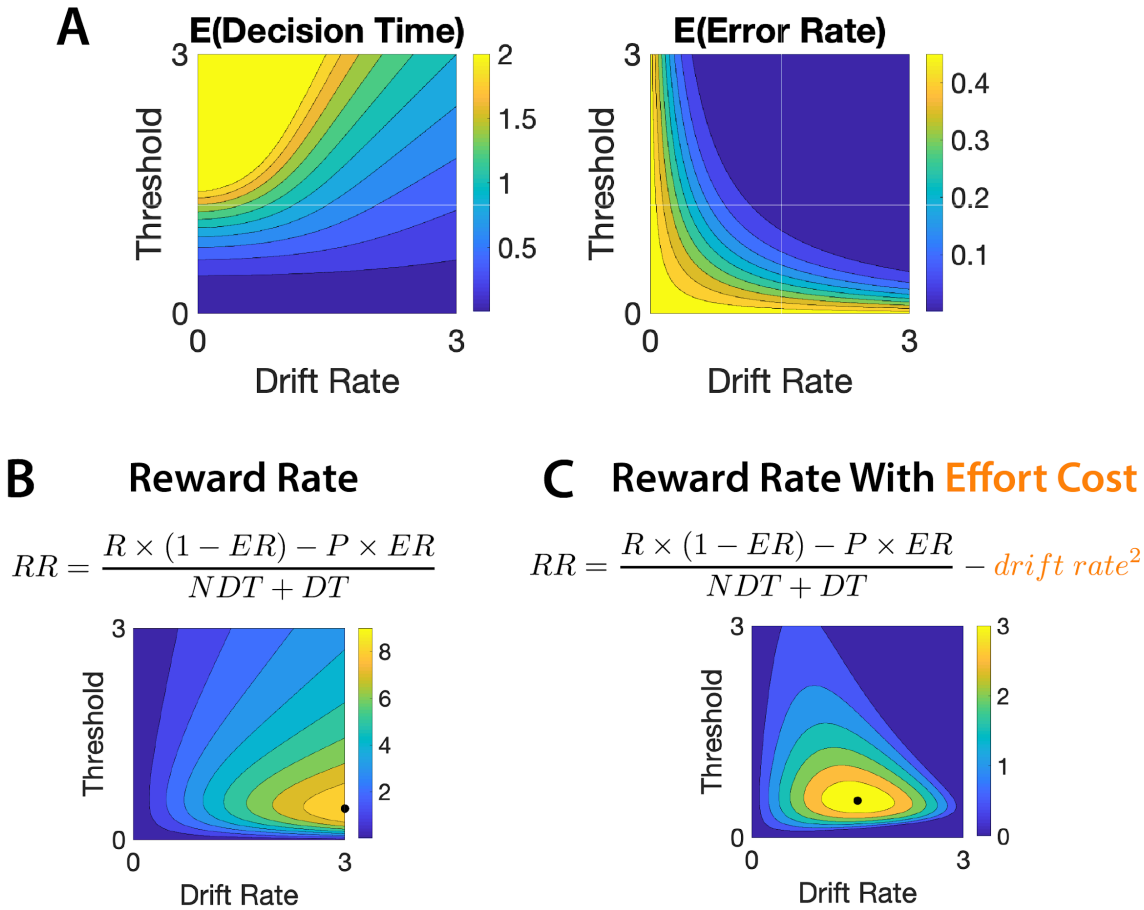

Figure 3. The influence of DDM parameter settings on estimates of reward rate. A) The expected error rate (ER) and decision time (DT) can be estimated as a function of drift rate and threshold. **B-C)** Reward rate is traditionally defined as a function of expected error rate, scaled by the value of correct vs. incorrect responses, and the overall response time (the combination of decision time and decision-unrelated processes; Bogacz et al., 2006). The combination of drift rate and threshold settings that maximizes reward rate (black dots) differs depending on whether drift rate is assumed to incur an effort cost or not. Without a cost (**B**), it is always optimal to maximize drift rate. With a cost (**C**), drift rate and threshold must both fall within a more constrained set of parameter values.

Using this formulation of reward rate (RR), we can generate predictions about the allocation of cognitive control (the combination of drift rate and threshold) that would be optimal under different reward and punishment conditions. To do so, we varied reward and punishment values and, for each pair, identified the pair of drift rate and threshold that would maximize reward rate. As reward increases, the model suggests that the optimal strategy is to increase the drift rate and reduce the threshold. As punishment increases, the optimal strategy is primarily to increase the threshold, while slightly increasing the drift rate (Figure 4A). 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 people will tend to increase drift rate the more they value receiving

a reward for a correct response. In contrast, people will adjust their threshold depending on how much they value receiving a reward for a correct response (decrease threshold) and receiving a punishment for an incorrect response (increase threshold).

Reward Rate-Optimal Control Allocation: Empirical Evidence

To test whether task performance was consistent with the predictions from our normative model, we fit behavioral performance on our task (reaction time and accuracy) with the HDDM. A systematic model comparison showed that the best-fitting parameterization of this model for our task allowed both drift rate and threshold to vary with trial-to-trial differences in congruency, reward level, and/or penalty level (Figure 4B). Critically, the parameter estimates from this model were consistent with predictions of our reward rate-optimal DDM (Figure 4C-E). Consistent with normative predictions, we found that drift rate increased with larger rewards as well as larger punishment. 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.

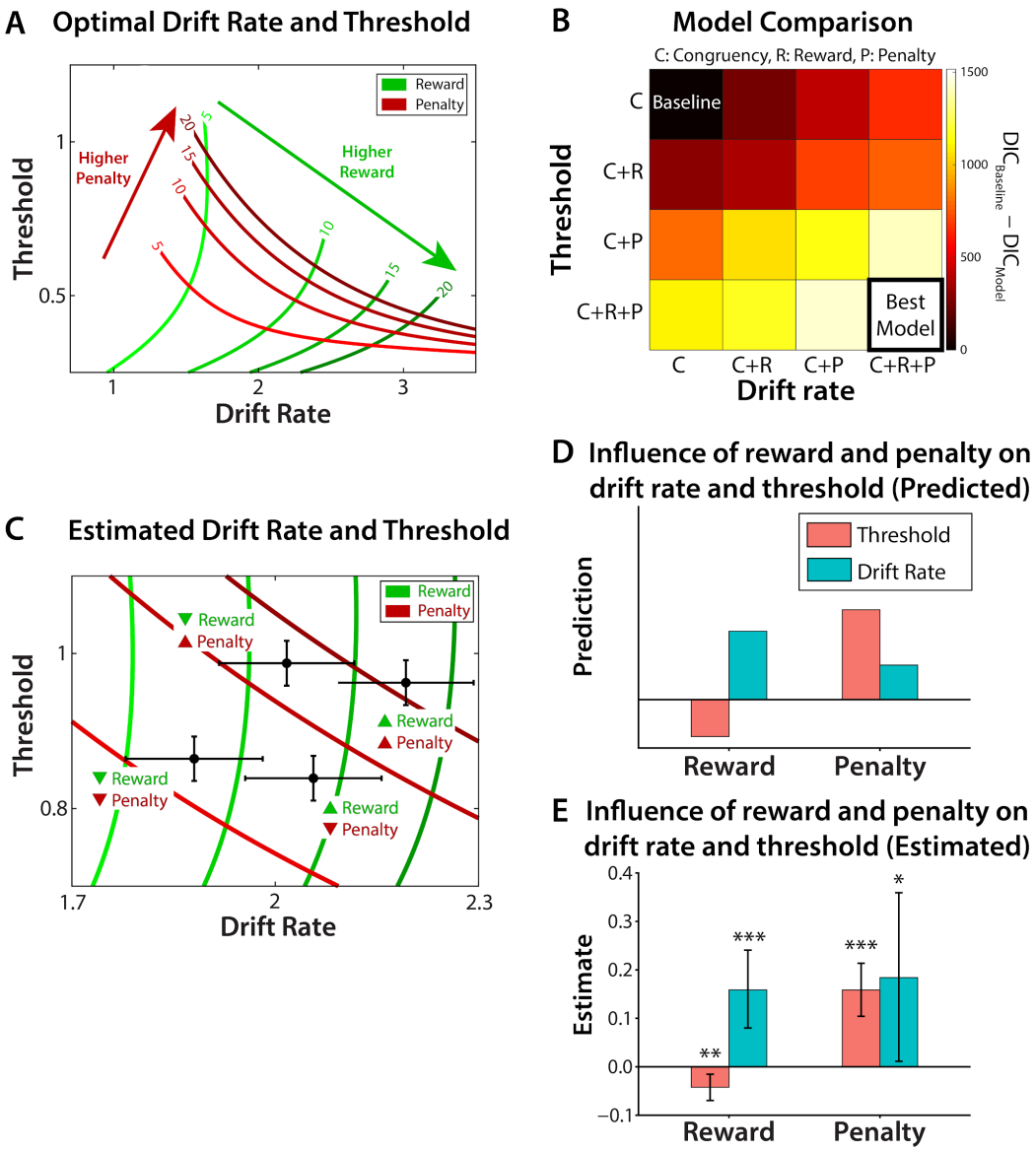

Figure 4. Normative and empirically observed estimates of incentive effects on DDM parameters. A) Combinations of drift rate and threshold that optimize (cost-discounted) reward rate, under different values of reward and penalty. **B)** We fit our behavioral data to different parameterizations of the DDM, with drift rate and/or threshold varying with reward, penalty, and/or congruency levels. The best-fitting model varied both DDM parameters with all three task variables. **C)** Estimated combination of drift rate and threshold for four conditions in the experiment. Error bars reflect s.d. **D-E)** Consistent with predictions based on reward-rate optimization (**D**, cf. panel **A**), we found that larger expected rewards led to increased drift rate and decreased threshold, whereas larger expected penalties led to increased threshold and (to a lesser extent) increased drift rate (**E**, cf. panel **C**). Error bars reflect 95% CI. *: p<0.05; ***: p<0.001.

Inferring Individual Differences in Reward and Punishment Valuation

Participants' subjective valuation of incentives is a latent variable that must be inferred from task performance. Since the reward-rate optimization model provides the mapping from incentives onto control configuration (i.e., reward-rate optimization) (Figure 5A), we can 'reverse-engineer' participants' subjective valuation of reward and punishment based on their DDM parameters (e.g., drift rate and threshold; 'inverse reward-rate optimization'). Simulation-based parameter recovery validated this approach for inferring individualized latent subjective valuation of reward and punishment incentive effects on adaptive cognitive control (See Supplementary Material).

Here, we used inverse reward-rate optimization to infer the individualized subjective weights of reward and punishment across the four task conditions based on participants' estimated DDM parameters. 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 and punishment that make the estimated DDM parameters the optimal strategy (Figure 5C). This workflow can be summarized as follows:

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

A repeated-measures ANOVA on our estimates of R and P (log-transformed) revealed a main effect of incentive magnitude (F(1,251)=16.60, p=6.2e-5), with larger R on high-reward intervals (t(31)=4.9, p<2.8e-5) and larger P on high-punishment intervals (t(31)=4.25, p=1.8e-4). We also observed a main effect of valence, such that estimates of P were higher than estimates of R (F(1,251)=210.04, p<2e-16). The ANOVA also revealed a significant interaction between valence and magnitude (F(1,251)=7.68, p=0.006) (Figure 5D), such that P estimates differed more across punishment levels than R estimates differed across reward levels. 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).

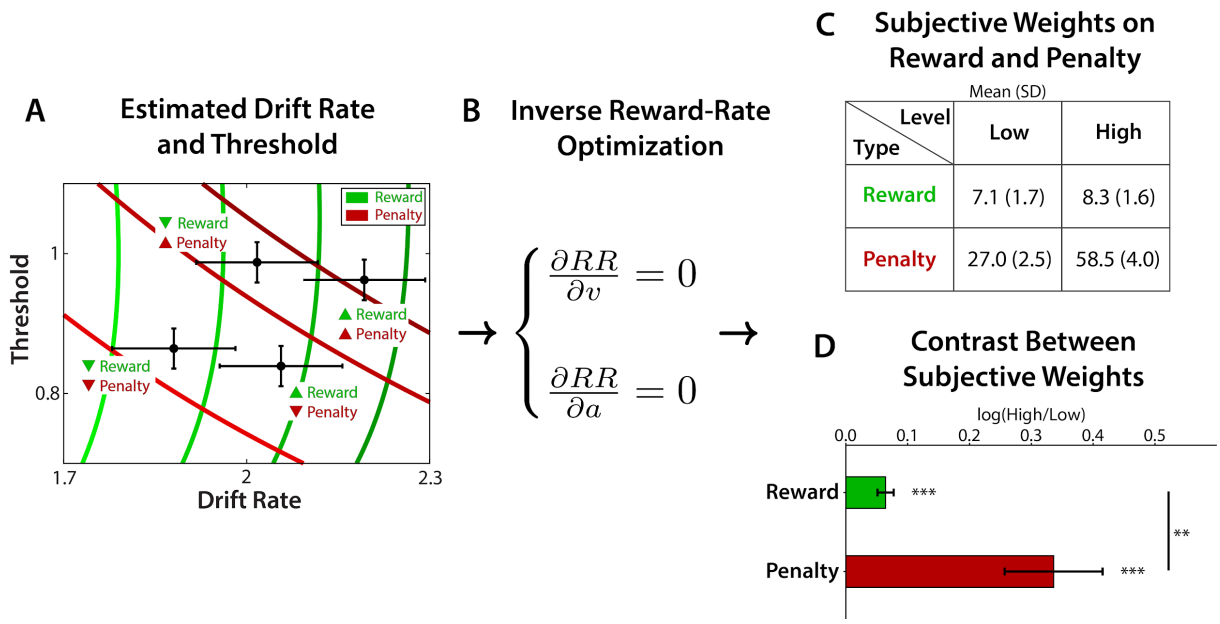

Figure 5. Inference of subjective weights on reward and penalty based on DDM estimates and reward rate optimization model. A) Estimated reward-rate optimal combinations of drift rate and threshold for the four conditions in the experiment. Error bars reflect s.d. **B)** To infer the subjective salience of reward and penalty for a given individual, we invert this reward-rate optimization procedure, estimating the set of reward and penalty weights (R and P) that best account for that person's pattern of behavior in a given condition. **C-D)** The resulting estimates of reward and penalty weights recapitulate our experimental manipulation, with higher subjective reward weights in the high vs. low reward condition, and higher subjective penalty weights for the high vs. low penalty condition. Panel **D** shows contrasts between subjectives weights of high vs. low reward and penalty. Error bars reflect s.e.m. **: p<0.01; ***: p<0.001.

Discussion

We investigated inverse 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 punishment increased, they responded slower and therefore completed fewer trials but were overall more accurate. We showed that these dissociable 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 that has shown that people flexibly recruit cognitive control 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, we incorporate the well-known cost of cognitive effort (Shenhav et al., 2017) into the reward rate optimization model (see below). 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.

In the current study, we demonstrate that adjustments of threshold and drift rate can vary as a function of 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, which is grounded in past research on mental effort (Manohar et al., 2015; Shenhav et al., 2017). Without this key assumption, the existing reward rate model would predict that individuals would maintain a maximal drift rate across incentive conditions, which would additionally guarantee consistently fast and accurate responding without constraint. Thus, 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 (Danielmeier et al., 2011; King et al., 2010; Musslick et al., 2015; Shenhav et al., 2013, 2016; van Maanen et al., 2011), threshold adjustments already carry an inherent cost in the form of a speed-accuracy tradeoff. Therefore, it 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.

While our modified reward rate optimization model was able to accurately characterize how reward and punishment incentives influenced cognitive control allocation our task, it is worth noting that a pure reward rate optimization model may not be ideally suited on its own due to its strong accuracy bias (Balci et al., 2011). To address these deviations from pure reward rate maximization, alternate variants of our model could down-weight response time and/or integrate expected reward and response time in an additive rather than multiplicative fashion.

In sum, 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. We found that people weighed punishments more heavily than rewards, despite the equivalent currency (i.e., amounts of monetary gain vs. loss). This finding is consistent with past work on loss aversion (Kahneman & Tversky, 1979) and motivation to avoid failure (Atkinson, 1957; Hajcak & Foti, 2008), and more generally, with the findings that distinct neural circuits are specialized for processing appetitive versus aversive outcomes (Bissonette et al., 2014; Pessiglione & Delgado, 2015). Critically, the current approach and findings hold promise for research into individual differences in sensitivity to rewards versus punishments (Boksem et al., 2006; Braem et al., 2013; Froböse et al., 2018; Kim et al., 2015; Locke & Braver, 2008;

McNaughton & Gray, 2000). 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; Cavanagh & Shackman, 2015; Grahek et al., 2019; Paulus, 2015).

References

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64, Part 1*(6), 359–372.

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

Balci, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J. A., Holmes, P., & Cohen, J. D. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Attention, Perception & Psychophysics*, *73*(2), 640–657.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *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.

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. *Frontiers in Systems Neuroscience*, *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. *Psychological Review*, *113*(4), 700–765.

Boksem, M. A. S., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Research*, *1101*(1), 92–101.

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113.

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

Braem, S., Duthoo, W., & Notebaert, W. (2013). Punishment sensitivity predicts the impact of punishment on cognitive control. *PloS One*, *8*(9), e74106.

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., Adcock, R. A., Barch, D. M., Botvinick, M. M., Carver, C. S., & Others. (2014). Mechanisms of motivation--cognition interaction: challenges and opportunities. *Cognitive, Affective & Behavioral Neuroscience*, *14*(2), 443–472.

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *Journal of Physiology, Paris*, *109*(1-3), 3–15.

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.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–361.

Cools, R. (2015). The cost of dopamine for dynamic cognitive control. *Current Opinion in Behavioral Sciences*, *4*, 152–159.

Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M., & Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(5), 1780–1789.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, 233.

Dixon, M. L., & Christoff, K. (2012). The decision to engage cognitive control is driven by expected reward-value: neural and behavioral evidence. *PloS One*, *7*(12), e51637.

Froböse, M. I., Swart, J. C., Cook, J. L., Geurts, D. E. M., den Ouden, H. E. M., & Cools, R. (2018). Catecholaminergic modulation of the avoidance of cognitive control. In *Journal of Experimental Psychology: General* (Vol. 147, Issue 12, pp. 1763–1781).

Frömer, R., Lin, H., Dean Wolf, C., Inzlicht, M., & Shenhav, A. (2020). *When effort matters: Expectations of reward and efficacy guide cognitive control allocation* (p. 2020.05.14.095935). https://doi.org/10.1101/2020.05.14.095935

Grahek, I., Shenhav, A., Musslick, S., Krebs, R. M., & Koster, E. H. W. (2019). Motivation and cognitive control in depression. *Neuroscience and Biobehavioral Reviews*, *102*, 371–381.

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., Halpern, D., Hamrick, J. B., & Chan, P. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, *48*(3), 829–842.

Hajcak, G., & Foti, D. (2008). Errors Are Aversive. In *Psychological Science* (Vol. 19, Issue 2, pp. 103–108).

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291.

Kim, S. H., Yoon, H., Kim, H., & Hamann, S. (2015). Individual differences in sensitivity to reward and punishment and neural activity during reward and avoidance learning. *Social Cognitive and Affective Neuroscience*, *10*(9), 1219–1227.

King, J. A., Korb, F. M., von Cramon, D. Y., & Ullsperger, M. (2010). Post-error behavioral

adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(38), 12759–12769.

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.

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 Zener, Trans.). *New York, NY, US: McGraw-Hill*.

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

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: behavior, brain activation, and individual differences. *Cognitive, Affective & Behavioral Neuroscience*, *8*(1), 99–112.

Manohar, S. G., Chong, T. T.-J., Apps, M. A. J., Batla, A., Stamelou, M., Jarman, P. R., Bhatia, K. P., & Husain, M. (2015). Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Current Biology: CB*, *25*(13), 1707–1716.

McNaughton, N., & Gray, J. A. (2000). Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety. *Journal of Affective Disorders*, *61*(3), 161–176.

Musslick, S., Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2015). A computational model of control allocation based on the expected value of control. *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.

Paulus, M. P. (2015). Cognitive control in depression and anxiety: out of control? *Current Opinion in Behavioral Sciences*, *1*, 113–120.

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). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Ritz, H., DeGutis, J., Frank M.J., Esterman, M., & Shenhav, A. (2020). An evidence accumulation model of motivational and developmental influences over sustained attention. In Proceedings of the 42nd Annual Meeting of the Cognitive Science Society. Toronto, CA.

Ritz, H., Shenhav, A. (2019) Parametric control of distractor-oriented attention. Proceedings of the 41st Annual Conference of the Cognitive Science Society. Montreal, CA.

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., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286–1291.

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*, 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. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(6), 1865–1897.

van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E.-J., Ho, T., Serences, J., & Forstmann, B. U. (2011). Neural Correlates of Trial-to-Trial Fluctuations in Response Caution. In *Journal of Neuroscience* (Vol. 31, Issue 48, pp. 17488–17495).

van Veen, V., Krug, M. K., & Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *Journal of Cognitive Neuroscience*, *20*(11), 1952–1965.

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*(3), e91008.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*, 14.

Wrosch, C., Scheier, M. F., Carver, C. S., & Schulz, R. (2003). The Importance of Goal Disengagement in Adaptive Self-Regulation: When Giving Up is Beneficial. In *Self and Identity* (Vol. 2, Issue 1, pp. 1–20).

Yee, D. M., & Braver, T. S. (2018). Interactions of motivation and cognitive control. In *Current Opinion in Behavioral Sciences* (Vol. 19, pp. 83–90).