

Short-Term Memory Capacity Predicts Willingness to Expend Cognitive Effort for Reward

 Brandon J. Forys,¹  Catharine A. Winstanley,^{1,2}  Alan Kingstone,¹ and  Rebecca M. Todd^{1,2}

¹Department of Psychology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada and ²Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada

Abstract

We must often decide whether the effort required for a task is worth the reward. Past rodent work suggests that willingness to deploy cognitive effort can be driven by individual differences in perceived reward value, depression, or chronic stress. However, many factors driving cognitive effort deployment—such as short-term memory ability—cannot easily be captured in rodents. Furthermore, we do not fully understand how individual differences in short-term memory ability, depression, chronic stress, and reward anticipation impact cognitive effort deployment for reward. Here, we examined whether these factors predict cognitive effort deployment for higher reward in an online visual short-term memory task. Undergraduate participants were grouped into high and low effort groups ($n_{\text{HighEffort}} = 348$, $n_{\text{LowEffort}} = 81$; $n_{\text{Female}} = 332$, $n_{\text{Male}} = 92$, $M_{\text{Age}} = 20.37$, $\text{Range}_{\text{Age}} = 16\text{--}42$) based on decisions in this task. After completing a monetary incentive task to measure reward anticipation, participants completed short-term memory task trials where they could choose to encode either fewer (low effort/reward) or more (high effort/reward) squares before reporting whether or not the color of a target square matched the square previously in that location. We found that only greater short-term memory ability predicted whether participants chose a much higher proportion of high versus low effort trials. Drift diffusion modeling showed that high effort group participants were more biased than low effort group participants toward selecting high effort trials. Our findings highlight the role of individual differences in cognitive effort ability in explaining cognitive effort deployment choices.

Received Feb. 13, 2024; revised May 26, 2024; accepted May 31, 2024.

The authors declare no competing financial interests.

Author Contributions: B.J.F. and R.M.T. designed research; B.J.F. performed research; B.J.F. analyzed data; B.J.F., C.A.W., A.K., and R.M.T. wrote the paper.

We acknowledge Alex Terpstra for creating the delayed monetary incentive task, Rita Jin for assisting in data analyses, and Claire Hales for her input on the drift diffusion model results. This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Grant (#F19-05182) to R.M.T. and the University of British Columbia Djavad Mowafaghian Centre for Brain Health Innovation Fund Kickstart Research Grant (#F19-05932), as well as an NSERC Canada Graduate Scholarship—Doctoral (CGS D) award to B.J.F. and a Michael Smith Foundation for Health

Continued on next page.

Significance Statement

We must often make decisions about when cognitive effort is worth the potential reward. Reward value, depression, and chronic stress in rodents can impact cognitive effort deployment decisions for reward, but factors like short-term memory ability can only be easily characterized in humans. We examined whether short-term memory ability, depression, chronic stress, and reward anticipation predict cognitive effort decisions for reward. In a short-term visual memory task with a choice of easier or harder trials for low versus high reward, we found that only short-term memory ability predicted more choices of high versus low effort trials. This research suggests that cognitive effort decisions could be driven by cognitive effort ability more than motivational factors like depression or chronic stress.

Introduction

We are often faced with difficult choices about work and life. For example, we may choose to spend time thoroughly studying for an exam to gain a few extra percent points on a grade; alternatively, we may trade off this reward to work less and be able to do other tasks. These choices require trading off more or less effort for a larger or smaller reward, and thus involve deciding how much cognitive effort to deploy. In general, the motivation

to deploy cognitive effort can be influenced by the potential reward to be gained (Shenhav et al., 2013, 2016; Yee et al., 2021). Beyond this, an individual's willingness to expend cognitive effort can also be linked to individual differences in factors that influence overall motivation, such as mood disorder levels (Grahek et al., 2019; Yee et al., 2019; Pruessner et al., 2020). Research in rodents has revealed patterns of individual differences in motivation to exert cognitive effort, where rats are classified as "workers" (high effort group) or "slackers" (low effort group) depending on whether they are willing to work more or less for reward, respectively (Hosking et al., 2016; Silveira et al., 2020, 2021). However, factors influencing individual differences in willingness to deploy cognitive effort for reward have yet to be fully characterized in humans.

In both rodents and humans, chronic stress (Birn et al., 2017; Watt et al., 2017) and depressive traits (Grahek et al., 2019; Silveira et al., 2020) have been shown to negatively impact cognitive effort deployment for reward. In rodents, chronic stress dampens reward anticipation even as acute stress heightens it (Ironsides et al., 2018; Kúkel'ová et al., 2018). Here, chronic stress is typically induced by prolonged social defeat (Kúkel'ová et al., 2018) or non-social chronic mild stress, such as modifications to housing (Slattery and Cryan, 2017). In humans, chronic stress—unlike acute stress—cannot ethically be induced. Research into the impact of chronic stress on reward motivation primarily focuses on reports early childhood stress (Birn et al., 2017; Watt et al., 2017), and not shorter-term chronic stress of the kind that university students may experience (Towbes and Cohen, 1996). This latter form of chronic stress is more widely experienced than early childhood stress, and yet its impact on our willingness and ability to complete everyday cognitive tasks is not well understood.

Alongside chronic stress, depression is another motivational factor that has been shown to dampen reward anticipation (Westbrook et al., 2018, 2020; Terpstra et al., 2023). In particular, anhedonia—a key symptom of depression impacting interpretations of reward (Slattery and Cryan, 2017)—can drive reduced reward anticipation and negatively impact willingness to deploy effort in rodents (Scheggi et al., 2018). Similarly, in humans, trait-level anhedonia (Treadway et al., 2012, 2009) and anticipation of anhedonia (Sherdell et al., 2012) generally predict reduced motivation to deploy physical effort for reward (Culbreth et al., 2018). However, many tasks in life are cognitive as opposed to physical in nature; cognitive effort tasks may offer more naturalistic value as a result.

Decisions on when to deploy more or less cognitive effort require attentional control in both humans and rodents. Such decisions can be driven by executive processes such as visuospatial working memory (Engstrom et al., 2013). Greater working memory ability can, in turn, drive better attentional control (Unsworth and Robison, 2020). Furthermore, during cognitively effortful tasks, reward anticipation modulates activation in neural circuitry that supports working memory (Fuentes-Claramonte et al., 2015). Reductions in executive function capability—for example, given high levels of chronic stress and depressive traits—could drive excessive reliance on reward incentives, leading to maladaptive effort deployment (Kool et al., 2010; Sandra and Otto, 2018). Although visuospatial working memory tasks can be conducted in rodents, including those exhibiting chronic stress and depression phenotypes, cognitive factors impacting cognitive effort deployment in humans are more complex than those that we can evaluate with rodents (Stephan et al., 2019). Acknowledging and accounting for these fundamental differences between human and rodent experiences of the world is an important yet underexamined element of translational research. Therefore, it is important to translate rodent cognitive effort tasks (rCETs) into human tasks, allowing us to test human-based theories of cognitive and attentional control while maintaining comparability with rodent work.

In humans, the expected value of control theory (Shenhav et al., 2013) predicts that, in general, people will deploy more cognitive effort when they know that they can effectively obtain greater rewards by doing so (Frömer et al., 2021). Additionally, the theory predicts that greater reward anticipation will lead to a higher expected value of deploying cognitive effort to obtain a reward (Grahek et al., 2020). Past work suggests that dopaminergic processes motivate physical and cognitive effort deployment decisions to maximize rewards (Michely et al., 2020) and to weigh rewards more heavily than effort costs. These processes also interplay with serotonergic processes that promote reward learning and biasing people toward more cognitively effortful options (Westbrook et al., 2020) even when it is suboptimal to deploy effort earlier—which could occur more often in people with low levels of short-term working memory ability (Raghubath et al., 2021). However, the degree to which individual

Research Scholar Award to R.M.T.

Correspondence should be addressed to Brandon J. Forys at brandon.forys@psych.ubc.ca.

Copyright © 2024 Forys et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

differences in reward anticipation and executive function capacity influence the likelihood of choosing to deploy more cognitive effort for higher reward is not known.

In rodents, different levels of motivation between “workers” and “slackers” have also been associated with differences in learning rate and bias toward choosing high effort trials, which may be linked to differences in cognitive capacity. Drift diffusion modeling (DDM) in rodents completing a cognitive effort task has revealed that “workers” with high accuracy may accumulate evidence more quickly toward selecting high effort, high reward (HR) trials than “slackers” or “workers” with low accuracy (Hales et al., 2024). In an equivalent task in humans, such evidence could include awareness of one’s effort deployment willingness and ability as well as fatigue. For example, participants may build experience with how the task demands align with their short-term memory capabilities, as well as their willingness to push these abilities further for higher reward. Overall, the above drift diffusion model findings suggest that effort cost computations could vary with a combination of effort demands and reward value as factors we rely on to make judgments about when to deploy more or less cognitive effort for reward. In turn, individual differences in reward anticipation may be linked to depressive traits or chronic stress, and rodent research indicates that executive capacity may also influence choice behavior (Eichenbaum, 2017). The goal of the present study was to examine the degree to which individual differences in reward anticipation, chronic stress levels, and depressive traits on the one hand, and visuospatial short-term memory as an index of executive function on the other, predict human choice behavior independently or in interaction.

The present study

One way of examining factors that motivate humans to deploy more or less effort for reward is to offer participants the choice of completing an easier or harder cognitive effort task for a low or HR, respectively (Treadway et al., 2009). In the current study, we built on an existing rodent choice task to investigate whether visual short-term memory capacity, chronic stress, depressive traits, and reward anticipation predict one’s choices in deploying cognitive effort for reward. A variety of human tasks exist to evaluate the costs of effort decisions. Some of these tasks examine the point at which high versus low effort and reward options are equally preferred, including the cognitive effort discounting task (Westbrook et al., 2013; Crawford et al., 2021). In contrast, physical effort deployment tasks, such as the effort expenditure for rewards task (Treadway et al., 2009), typically offer a binary effort choice structure with an easy trial for a lower reward or a harder trial for a higher reward. A key aim of our study was to examine whether human cognitive effort decisions can be modeled based on established animal models of cognitive effort deployment (Cocker et al., 2012; Hosking et al., 2016; Silveira et al., 2021). Furthermore, the use of two effort levels encourages participants to strategize according to the trial difficulties available rather than fully discounting excessively easy or difficult trials offered in a parametric design. Thus, to ensure continuity with rCETs—including a binary choice paradigm—while maintaining a similar incentive structure to existing human cognitive effort tasks, we adapted the rCET (Cocker et al., 2012; Hosking et al., 2016; Silveira et al., 2020) for use in humans. In the rodent paradigm, rodents first learn to associate illuminated lights with the opportunity to obtain a reward. In the main phase of the task, rodents are given the choice between a low effort, low reward (LR) trial or a high effort, HR trial. Then, in a basic memory task, they must poke their noses in a hole that lights up for either 1000 ms (LR) or 200 ms (HR) (Cocker et al., 2012). Rodents must rapidly encode the location of the illuminated light to succeed. Such a task can be scaled up to human working memory capabilities by including more stimuli and features that must be encoded to obtain a reward. For our study, we used harder and easier conditions (larger vs smaller set size) from a visual short-term memory task (Luck and Vogel, 1997) to serve as high and low cognitive effort choices within the same kind of choice paradigm offered in the rCET. In each trial, participants could choose to perform an easier visual short-term memory task for a lower reward or a harder task for a higher reward. We examined overall visuospatial short-term memory capacity as well as indices of depressive traits (Beck et al., 1996), chronic stress (Levenstein et al., 1993), anhedonia (Rizvi et al., 2015), and reward anticipation (Terpstra et al., 2023) as predictors of the tendency to choose lower or higher effort tasks. Furthermore, we used DDM (Ratcliff, 1978; Hales et al., 2024) to examine factors that may contribute to choice biases.

Materials and Methods

Participants

We powered our study using a power analysis through the `pwrss` package (Bulus, 2023) to achieve an expected power of 80% at an odds ratio of 0.70 for a logistic regression evaluating the probability of a participant selecting 70% or more high effort trials (being in the high effort group in the study). This analysis gave us a target sample size of $N = 487$. We recruited $N = 570$ participants from the Human Subject Pool of psychology undergraduate students at the University of British Columbia; they were each compensated with bonus points for their courses as well as a CAD (Canadian dollar) \$5.00 Starbucks gift card. Of these, $n = 42$ participants did not complete the initial survey and $n = 7$ participants did not complete the reward anticipation task. Furthermore, $n = 11$ participants completed the change detection task more than once, $n = 2$ participants spent more than 30 s choosing any one trial—as this could indicate disengagement with task demands, and $n = 79$ participants performed at or below chance (50%) during the choice phase of the task. In total, we analyzed data from $N = 429$ participants ($n = 92$ male, $n = 332$ female, $n = 5$ other; Table 1). The study was approved by the Behavioural Research Ethics Board at the University of British Columbia, approval code H20-01388.

Table 1. Demographic information for all participants, by sex and effort deployment group

Group	Sex	<i>n</i>	<i>M</i> _{age}	<i>SD</i> _{age}	<i>Min</i> _{age}	<i>Max</i> _{age}	<i>M</i> _{BDIprop}	<i>SD</i> _{BDIprop}	<i>M</i> _{PSS}	<i>SD</i> _{PSS}	<i>M</i> _{DARS}	<i>SD</i> _{DARS}	<i>M</i> _{ant.}	<i>SD</i> _{ant.}
High effort	Female	263	20.34	2.37	16	42	0.27	0.18	20.83	6.21	35.61	11.71	3.94	1.81
	Male	80	20.70	2.34	17	32	0.19	0.14	17.69	5.30	37.69	9.62	4.29	1.75
	Other	5	20.40	0.89	19	21	0.17	0.10	19.60	3.65	43.00	10.07	4.80	1.64
Low effort	Female	69	20.01	1.94	16	26	0.28	0.16	20.68	6.54	38.04	8.94	4.00	1.79
	Male	12	20.92	2.97	18	28	0.19	0.14	21.08	4.38	39.17	8.61	3.96	1.51

BDI prop: beck depression inventory II proportion score (score divided by max score); PSS: perceived stress scale score; DARS: dimensional anhedonia rating scale; Ant.: mean reward anticipation in the behavioral monetary incentive delay task; high effort group: >70% HR trials selected; low effort group: ≤70% HR trials selected.

Stimulus presentation

All participants completed the study online on their own devices, via the Pavlovia online study platform using PsychoPy 2021.2.3 (RRID: SCR_006571; Peirce et al., 2019). Participants were not allowed to complete the study on mobile phones or tablets.

Stimuli

All stimuli used in the study were generated by and implemented in PsychoPy (Peirce et al., 2019). In the first phase of the study, participants performed an online behavioral monetary incentive delay task as a measure of reward anticipation (Fig. 1A) as outlined by Terpstra et al. (2023). In brief, during this initial task phase, participants saw a series of scales that they could click to indicate how excited they were to receive a reward. After a brief wait, participants would see a small happy face appear on one side of the screen. Afterward, they were shown whether they received a reward or not, and were asked to view another scale where they would indicate how excited they felt. They completed this task for a series of eight trials.

In the second phase of the study, as a measure of each individual's visuospatial short-term memory capacity, we presented a visuospatial short-term memory task (Fig. 1B,C) modified from Luck and Vogel (1997) and adapted from an open-source version of the task on Pavlovia (de Oliveira, 2023). In this task, we presented a series of trials with between two and eight colored squares that were presented for 500 ms. Each square subtended a visual angle of approximately 0.05° on the screen. A mask with multi-colored squares would then appear at each of the original squares' locations for another 500 ms, followed by a single colored square appearing in one of the positions of the original squares. This final square had a 50% chance of being the same or a different color from the square appearing in the same position in the initial part of the trial. After indicating with a keyboard press whether the square was the same or a different color from the initial square, they would see whether or not they gained points toward a monetary reward.

Procedure

Questionnaires and monetary incentive delay task. Participants began the study by completing an online questionnaire. After giving consent and demographic information, they were asked how about their history of depression and anxiety, COVID-related stress, the perceived stress scale (PSS; Levenstein et al., 1993); the beck depression inventory II (BDI; Beck et al., 1996), and the dimensional anhedonia rating scale (DARS; Rizvi et al., 2015). Afterward, they were redirected to the first phase of our study, the online behavioral monetary incentive delay task (Fig. 1A; Terpstra et al., 2023). This task measured participants' reward anticipation via the mean of the excitement ratings given before each trial.

Visuospatial short-term memory task. Versions of this task served both as a means to measure individual differences in visuospatial short-term memory as well as the tasks of varied cognitive effort to be subsequently chosen for high/low effort trials. After completing the monetary incentive delay task, participants were redirected to the second phase of our study, the visuospatial short-term memory task (Luck and Vogel, 1997). Here, this task served as a means to measure individual differences in visuospatial short-term memory as an index of working memory. After receiving instructions on which stimuli would appear and how to respond to them, participants first completed a series of 10 practice trials presented in a random order. In half of these trials, the probe square was the same color as that of the cue square (congruent/no change trial); on the other half, the probe square was a different color from that of the cued square (incongruent/change trial). Of these practice trials, four had a set size = 2, another four had a set size = 4, and another two had a set size = 6. In each trial, the probe square could be anywhere in the array, and participants were required to hold the whole array in visual short-term memory in order to successfully indicate whether the probe color had changed. Participants did not receive feedback on whether their responses were correct or not in the practice block, in keeping with Luck and Vogel (1997). Following these practice trials, participants completed (Fig. 1B) a series of 120 trials of the visuospatial short-term memory task with 60 change and 60 no change trials. In total, 30 trials of each set size (2, 4, 6, and 8 squares onscreen) were presented in a randomized order. Although the original task by Luck and Vogel (1997) contained trials with up to 10 squares onscreen, this largest set size was determined to be too difficult for participants to reliably complete correctly following piloting; as such, the maximum set size was eight. Furthermore, although the main task had a maximum set size of eight, the practice block

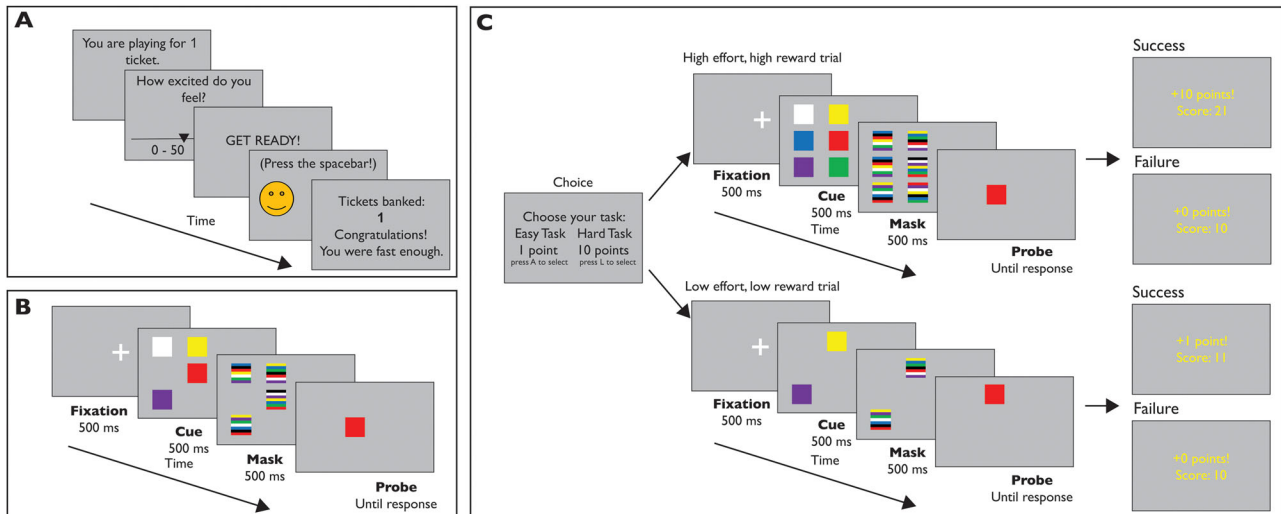


Figure 1. Layout of the experimental tasks. **A**, The monetary incentive delay task, in which participants indicate their excitement in playing for tickets toward a monetary reward. **B**, The calibration phase of the change detection task, where participants saw an array of either two, four, six, or eight squares and indicated whether the final square that appeared (the probe) was of the same or a different color from the square that appeared in the same location in the previously show array. **C**, The reward phase of the change detection task, where participants saw an array of either two or six, or four or eight, shapes onscreen depending on their performance in the calibration phase of the task. Once again, participants indicated with a keyboard press whether the final square that appeared (the probe) is the same or a different color from the square that appeared in the same location in the array. Here, if they made the correct decision, they would receive 1 point (low effort trial) or 10 points (high effort trial); they would receive 0 points for an incorrect decision.

only included a maximum set size of six so as to give a brief overview of the trials and required responses without giving excessive practice with a large set size, for which performance would be evaluated in the subsequent block.

Based on participants' performance in this phase of the task, a K estimate of their visuospatial working memory capability (Rouder et al., 2011) was calculated for each set size, as follows:

$$\hat{k}_p = N \left(\frac{\hat{h} - \hat{f}}{1 - \hat{f}} \right),$$

where N is the set size, \hat{h} is the hit rate, and \hat{f} is the false alarm rate. As participants had to evaluate the whole display to judge the color of the resultant single probe (Rouder et al., 2011), we used a whole-display K estimate measure (Pashler, 1988).

In addition to serving as indices of individual differences in visuospatial short-term memory, K estimate scores were used to calibrate the tasks used in the effort choice task (Fig. 1C), which was based on the rCET (Hosking et al., 2016; Silveira et al., 2021). Here, participants completed a series of 30 trials. They would receive rewards for making a correct response, and they could choose a high effort, HR trial or a low effort, LR trial at each trial. Importantly in this task, trials with small and large arrays from the visuospatial short-term memory task served as the LR and HR trials. HR trials used a larger set size, but gave a reward of 10 points toward a monetary reward. LR trials had a smaller set size than the high effort/HR trials, but yielded a reward of one point toward a monetary reward. We set the high effort trial reward to 10 points to motivate participants to continue deploying effort even for a much more difficult task. In order to ensure that the task difficulty in this phase was balanced by participants' working memory capability, we used participants' K estimate score at set size = 4 from the initially presented visuospatial short-term memory task. This set size was determined through piloting to be the maximum set size before performance dropped off. Performance at set size = 4 was used to set a criterion for the available set sizes in HR and LR trials. Specifically, if the K estimate at set size four was ≤ 3 , participants could choose an LR trial with a set size of two or an HR trial with a set size of six. If the K estimate at set size four was > 3 , participants could choose an LR trial with a set size of four or an HR trial with a set size of eight. Although participants were told that the number of points they gained in this phase was proportion to the monetary reward they would receive, all participants received the same reward (a \$5 CAD Starbucks gift card) at the end of the study. Afterward, participants were redirected to a debriefing survey and received their course credit and monetary reward.

Our primary outcome measure of interest was the proportion of time participants chose high versus low effort trials. In order to conduct secondary analyses evaluating whether participants' tendencies to choose high effort trials for HR explained performance (accuracy in the reward phase of the task) and choice latency (time until participants selected a difficulty level in the reward phase), we classified participants into two categories according to the criteria discussed by Silveira et al. (2021): participants who chose the HR option for more than 70% of all trials in the reward phase were

in the high effort preference group, while participants who chose the HR option for less than or equal to 70% of all trials in the reward phase were in the low effort preference group. This grouping was chosen to ensure continuity with comparable cognitive effort studies in rodents (Hosking et al., 2016; Silveira et al., 2021; Hales et al., 2024). Although the grouping binarizes the outcomes of our analyses, it offers a clear point of translation from the rodent literature on which our task is based, and allows us to predict whether participants exhibit a high effort motivation versus low effort motivation phenotype.

Analyses

All analyses were conducted using R 4.3.1 “Beagle Scouts” (R Core Team, 2023) through RStudio (Posit team, 2024).

The primary predictor variables in our study were (1) working memory ability, operationalized as a participant’s K estimate at a set size of four; (2) depressive traits, operationalized as a participant’s BDI proportion score; (3) chronic stress levels, operationalized as a participant’s PSS score; and (4) reward anticipation, operationalized as a participant’s mean excitement before playing for tickets in the monetary incentive delay task (Terpstra et al., 2023). As our overall depression score measures were of more translational interest than anhedonia, given the role of depression in downweighing effort deployment and reward anticipation in both rodents (Slattery and Cryan, 2017) and humans (Grahek et al., 2019), we focus on reporting this measure as opposed to the anhedonia measure we also collected. The primary dependent variable in our study was (1) proportion of HR trials chosen in the reward phase of the task, operationalized as whether participants chose the HR option for more or less than 70% of all trials in the reward phase. In additional analyses, we further examined the effects of willingness to expend high effort for HR on the following dependent variables in the rewarded visuospatial short-term memory task: (2) accuracy, operationalized as the proportion of correct responses (hits and correct rejections), and (3) choice latency, operationalized as the time in seconds until participants selected a difficulty level for each trial. To assess whether participants who consistently chose higher difficulty trials were overall better at the task compared to those who chose low difficulty trials, we compared accuracy between high and low effort groups. To examine whether, as predicted by Hales et al. (2024), choice latency was longer for all LR versus HR trials—as opposed to whether the tendency to choose high versus low effort trials explained how long participants took to make a decision when selecting a trial—we further compared choice latency between high and low effort groups. Finally, drift diffusion model parameters of drift rate, starting point, boundary separation, and non-decision time (Hales et al., 2024) during choices in the effort choice (reward) phase of the task were additional outcome variables.

For our analyses, we first conducted a series of *t* tests to evaluate sex differences in BDI and PSS scores. Past work suggests sex differences in depressive traits (Altemus et al., 2014; Forys et al., 2023) and chronic stress levels (Watt et al., 2017), with women presenting with higher depression and chronic stress scores than men. For sex difference comparisons, only male and female participants were included as we were under-powered to report sex differences from those reporting sex as “other”. As participants may express high levels of both depressive traits and chronic stress—potentially impacting reward anticipation (Alloy et al., 2016)—we also evaluated the extent to which BDI scores, PSS scores, and reward anticipation were correlated with each other. Next, we divided participants into high and low effort groups to ensure translational comparability with existing rodent work. As by Cocker et al. (2012), participants who chose the HR option for more than 70% of all trials in the reward phase of the task were in the high effort preference group, while those who chose the HR option for 70% or less of all trials in this phase were in the low effort preference group. We first conducted a binomial linear regression to determine whether sex, working memory ability, depressive traits, chronic stress levels, and reward anticipation significantly predicted whether a participant was in the high or low effort group in terms of their effort choices. As choices for high effort trials may also vary continuously with visual short-term memory ability, we then conducted a linear regression to evaluate whether the above regressors also predict the proportion of high effort trials selected. We then conducted two 2 × 2 within-between ANOVAs (trial type × group) using `anova_test` through the `rstatix` package (Kassambara, 2023) to evaluate whether accuracy or choice latency significantly differed by trial effort level (HR vs LR) and motivation group (high effort vs low effort groups). We examined accuracy on each trial and choice latency for selecting the trial difficulty to evaluate whether participants were matched for performance and time spent selecting a trial (Cocker et al., 2012), regardless of how many high versus low effort trials they selected. We then ran two multi-level models through the `lmerTest` package (Kuznetsova et al., 2017) to evaluate whether sex, working memory ability, depressive traits, chronic stress levels, reward anticipation, and effort level significantly predict choice latency or accuracy.

Lastly, in order to evaluate whether the high and low effort groups differed in response strategies and biases toward selecting HR versus LR trials, we fit a hierarchical Bayesian drift diffusion model, adapted from Ratcliff (1978) and run for 2,000 iterations with 1,000 warmup iterations and four Markov chains for Monte Carlo sampling, to data from all participants using the `hBayesDM` package in R (Ahn et al., 2017). We used this model to compare drift rate, starting point, boundary separation, and non-decision time between participants who chose high levels of effort versus those who chose low levels of effort. The drift rate captures the rate at which participants drift toward a decision boundary (selecting an HR or LR trial). The starting point is expressed as a percentage of the distance between the lower and upper boundaries; it captures the initial bias, at the beginning of each trial, that participants have toward selecting an HR over an LR trial. The boundary separation, expressed relative to zero, captures a tradeoff between speed and selecting the HR trial as opposed to the LR trial. Lastly, non-decision time captures the part of choice latency that is not related to cognitive effort decisions, such as the time required to execute a motor response upon making a decision.

Results

Depressive traits and chronic stress levels

Summary statistics for depressive traits (BDI) and chronic stress levels (PSS) can be found in Table 1. After correcting for multiple comparisons using the Bonferroni method, women had significantly higher depressive trait scores ($t(261.23) = 5.35, p \leq 0.001, d = 0.45$; Fig. 2A) and chronic stress scores ($t(270.03) = 4.93, p < 0.001, d = 0.40$) than men (Fig. 2B). However, men and women did not significantly differ on anhedonia scores ($t(253.79) = -2.05, p = 0.414, d = -0.17$; Fig. 2C) or levels of mean reward anticipation ($t(233.53) = -2.50, p = 0.131, d = -0.23$; Fig. 2D). Levels of depressive traits (BDI), chronic stress (PSS), and reward anticipation were significantly correlated with each other; neither measure was significantly correlated with working memory ability (K estimate; Fig. 3).

Predictors of overall high versus low effort choices

Our primary research question focused on factors that predict the likelihood of choosing high versus low effort/reward options (Table 2 and Fig. 4). Here, results of the binomial regression showed that only visual short-term memory

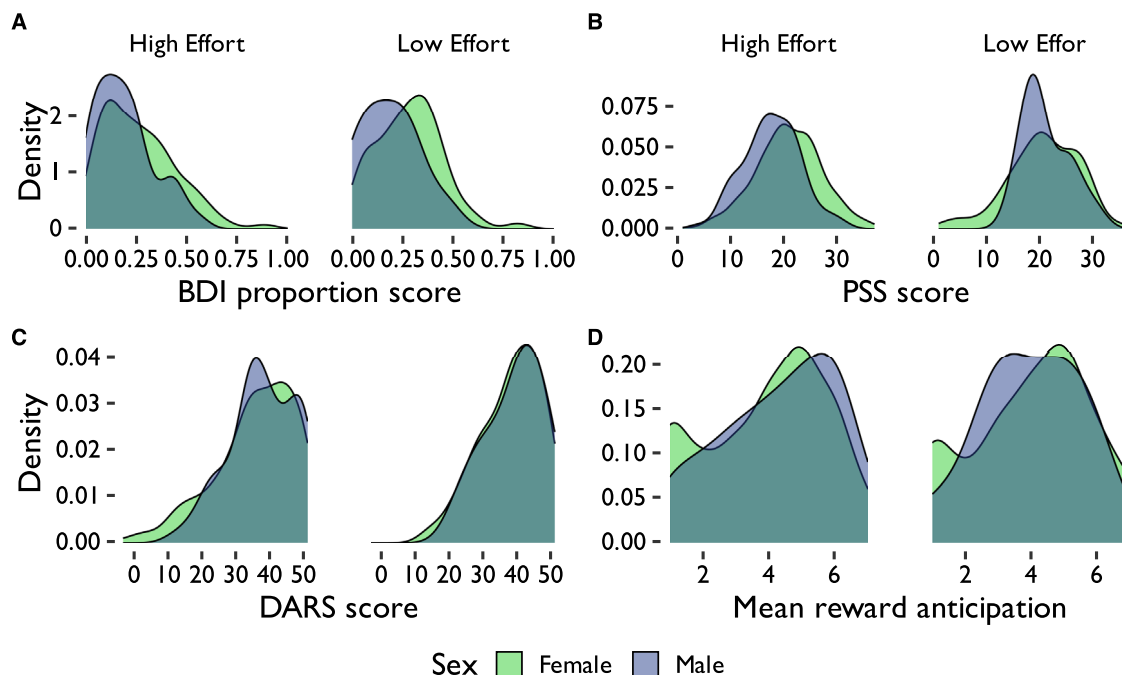


Figure 2. Distribution of **A**, depressive trait (BDI) proportion scores (score divided by total possible score), **B**, perceived stress (PSS) scores, **C**, anhedonia (DARS) scores, and **D**, mean reward anticipation scores by sex and effort deployment group. BDI, beck depression inventory; PSS, perceived stress scale; DARS, dimensional anhedonia rating scale.

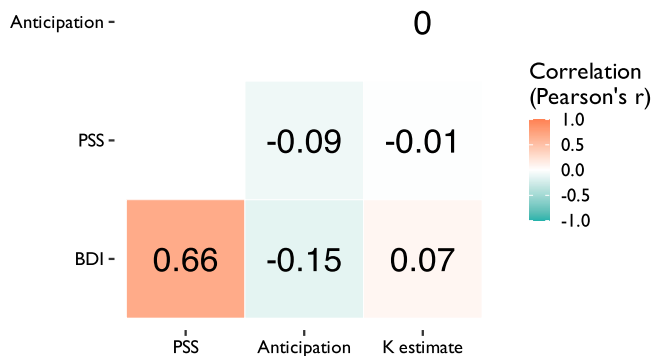


Figure 3. Correlations among depressive traits (BDI), perceived stress (PSS), reward anticipation, and short-term memory ability (K estimate). BDI, beck depression inventory; PSS, perceived stress sale; anticipation, mean reward anticipation; K estimate, visual short-term memory ability (K score).

Table 2. Binomial logistic regression predicting whether participants were in the high effort versus low effort group

Binomial logistic regression: high versus low effort group	
Working memory ability (K estimate)	0.21* (0.09)
BDI (prop. score)	-0.78 (0.96)
PSS	0.00 (0.03)
Reward anticipation	0.00 (0.07)
Num. Obs.	429
AIC	419.7
BIC	440.0
Log. Lik.	-204.833
RMSE	0.39

High effort group, >70% HR trials selected; low effort group, ≤70% HR trials selected; BDI (prop. score), depression score on the beck depression inventory; PSS, perceived stress scale. Proportion scores are scores divided by total possible score. AIC, Akaike information criterion; BIC, Bayesian information criterion; ICC, intraclass correlation; RMSE, root mean squared error.
[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

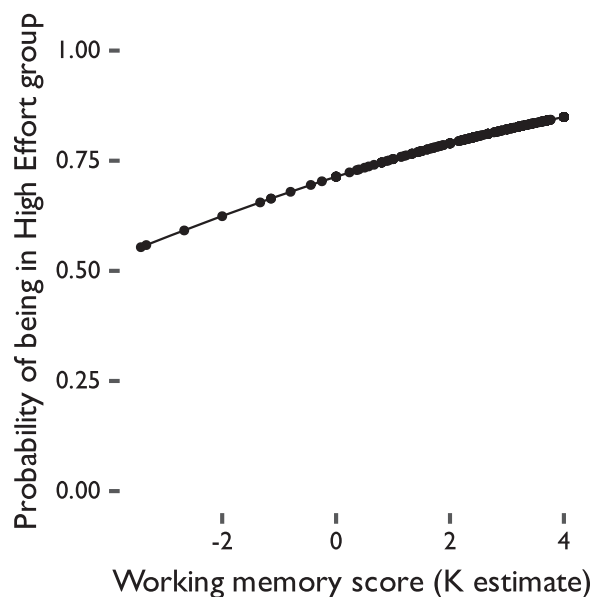


Figure 4. Binomial logistic regression predicting whether participants were in the high effort versus low effort group. High effort group, >70% high reward (HR) trials selected; low effort group, ≤70% HR trials selected.

(K estimate) score significantly predicted whether a participant was in the high effort or the low effort group ($z = 2.26$, $p = 0.024$, $e^{\beta} = 1.24$). Participants with higher working memory ability had a higher probability of being in the high effort group. Depressive traits, chronic stress levels, and reward anticipation did not significantly predict whether participants were in the low versus high effort groups. Furthermore, a linear model analysis with the same predictors revealed that, once again, only visual short-term memory (K estimate) scores significantly predicted the proportion of high effort trials that a participant selected (Table 3).

Accuracy

In all ANOVAs reported below, all p -values were corrected for multiple comparisons using Greenhouse–Geiser correction. A 2×2 within-between ANOVA (group \times trial type; Table 4 and Fig. 5A) revealed a main effect of trial type on accuracy such that participants showed poorer performance on the more difficult HR trials compared to the easier LR trials ($F(1, 273) = 724.10$, $p < 0.001$, $\gamma = 0.515$). This was qualified by a trial type \times group interaction such that participants in the high effort group who selected high effort trials more than 70% of the time performed better than those in the low effort group on LR but not HR trials ($F(1, 273) = 11.81$, $p < 0.001$, $\gamma = 0.017$).

Table 3. Linear regression of predictors for proportion of high effort trials selected

	Linear model: proportion of high effort trials selected
(Intercept)	0.84*** (0.04)
Working memory ability (K estimate)	0.02** (0.01)
BDI (prop. score)	-0.03 (0.07)
PSS	0.00 (0.00)
Reward anticipation	0.00 (0.00)
Num. Obs.	429
R^2	0.018
R^2 Adj.	0.009
AIC	-274.2
BIC	-249.8
Log. Lik.	143.105
RMSE	0.17

High effort group, >70% HR trials selected; low effort group, ≤70% HR trials selected; BDI (prop. score), depression score on the beck depression inventory; PSS, perceived stress scale. Proportion scores are scores divided by total possible score. AIC, Akaike information criterion; BIC, Bayesian information criterion; ICC, intraclass correlation; RMSE, root mean squared error.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Table 4. Summary table of accuracy and choice latency by group and effort level chosen

Group	Effort level	$M_{Accuracy}$	$SD_{Accuracy}$	$M_{Latency}$	$SD_{Latency}$
High effort	High	0.70	0.11	1.27	1.04
High effort	Low	0.95	0.11	2.54	1.99
Low effort	High	0.70	0.11	1.23	1.10
Low effort	Low	0.90	0.11	1.60	1.92

High effort group, >70% HR trials selected; low effort group, ≤70% HR trials selected.

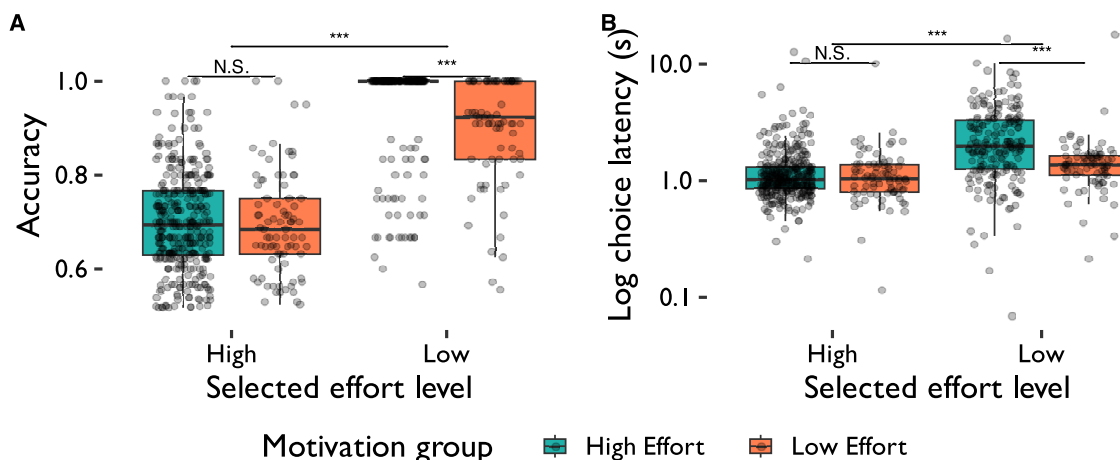


Figure 5. Accuracy and choice latency in the choice phase of the change detection task by motivation group (high effort vs low effort group). The choice latency plot has been log-transformed on the y-axis to more clearly show small choice latency values. High effort group, >70% HR trials selected; low effort group, ≤70% HR trials selected.

Choice latency

A 2 × 2 within-between ANOVA (group × trial type; Table 4 and Fig. 5B) revealed a main effect of trial type such that participants were slower to choose HR trials compared to LR trials ($F(1, 273) = 37.50, p < 0.001, \gamma = 0.063$). Furthermore, high effort participants were slower than low effort participants to make choices across all trial types ($F(1, 273) = 10.10, p = 0.002, \gamma = 0.019$). There was also a trial type × group interaction, such that participants in the high effort group took longer to select LR trials than HR trials, while choice latency between high effort and low effort groups did not differ on HR trials ($F(1, 273) = 37.50, p < 0.001, \gamma = 0.021$).

Additional predictors of accuracy and choice latency

We additionally evaluated predictors of accuracy on each trial and choice latency for selecting the trial difficulty from individual differences in participant visuospatial short-term memory capacity, mood disorder score, and reward anticipation by sex. A multi-level model analysis revealed that the effort level required in the selected trial significantly predicted accuracy across all trials such that accuracy was higher on LR trials compared to HR trials. Importantly, of the predictors of interest, only visuospatial short-term memory score predicted accuracy, such that higher short-term memory ability was associated with higher accuracy (Table 5 and Fig. 6A). Trial effort level also predicted choice latency, with slower choices on low relative to high effort trials. Depression scores also predicted choice latency such that higher depression scores predicted faster choices across all trials (Table 6 and Fig. 6B). Sex, chronic stress scores, and reward anticipation did not significantly predict accuracy or choice latency.

Drift diffusion model

To evaluate group differences in evidence accumulation and bias toward deploying higher amounts of cognitive effort in decision-making when choosing HR or LR trials, we fit a hierarchical Bayesian drift diffusion model (Ahn et al., 2017) to choice (HR vs LR trial selected) and corresponding choice latency data on each trial from all participants who had at least one correct response on an HR and an LR trial and who chose at least four trials of each type ($N = 136$, $n_{\text{High effort}} = 94$, $n_{\text{Low effort}} = 42$). A drift diffusion model can only be fit to participants who selected both trial choices. For this reason, we excluded data from $n = 2$ participants who selected only low effort trials and $n = 86$ participants who selected only high effort trials. We then evaluated this overall model fitting data from these participants to explore group-level differences in whether participants selected HR or LR trials more often. For each participant in the high effort and low effort groups, we evaluated four parameters: drift rate (speed of evidence accumulation in deciding on an HR vs LR trial), starting point (bias toward HR vs LR trial), boundary separation (extent to which speed and effort choice are traded off between HR and LR trials), and non-decision time (choice latency time unrelated to trial choice, such as time for motor response; Fig. 7). These parameters capture individual differences in the speed and motivation to make a decision about deploying more or less cognitive effort for reward. After Bonferroni correction for multiple comparisons, we found that high effort group participants had a starting point closer to the HR trial decision boundary ($t(83.21) = 5.80$, $p < 0.001$, $d = 1.05$); had wider decision boundaries ($t(68.49) = 4.38$, $p < 0.001$, $d = 0.87$) than low effort group participants; and had a significantly higher drift rate ($t(82.94) = 13.54$, $p < 0.001$, $d = 2.46$) than low effort group participants. However, participants did not significantly differ in non-decision time ($t(105.65) = 2.34$, $p = 0.213$, $d = 0.39$). These findings suggest that participants who were in the high effort group required more evidence to select low but not high effort trials, were more strongly biased toward selecting high effort/HR trials, and were quicker to accumulate evidence for these choices, than those in the low effort group.

Table 5. Multi-level model analysis coefficients and standard errors for accuracy in the reward phase of the task

	Multi-level model: accuracy
(Intercept)	0.64*** (0.02)
Sex	0.01 (0.01)
Working memory ability (K estimate)	0.01** (0.00)
BDI (prop. score)	-0.01 (0.04)
PSS	0.00 (0.00)
Reward anticipation	0.00 (0.00)
Effort level	0.25*** (0.01)
SD (intercept participant)	0.04
SD (observations)	0.10
Num. Obs.	542
R^2 Marg.	0.592
R^2 Cond.	0.656
AIC	-840.0
BIC	-801.3
ICC	0.2
RMSE	0.09

BDI (prop. score), depression score on the beck depression inventory; PSS, perceived stress scale. Proportion scores are scores divided by total possible score. AIC, Akaike information criterion; BIC, Bayesian information criterion; ICC, intraclass correlation; RMSE, root mean squared error.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

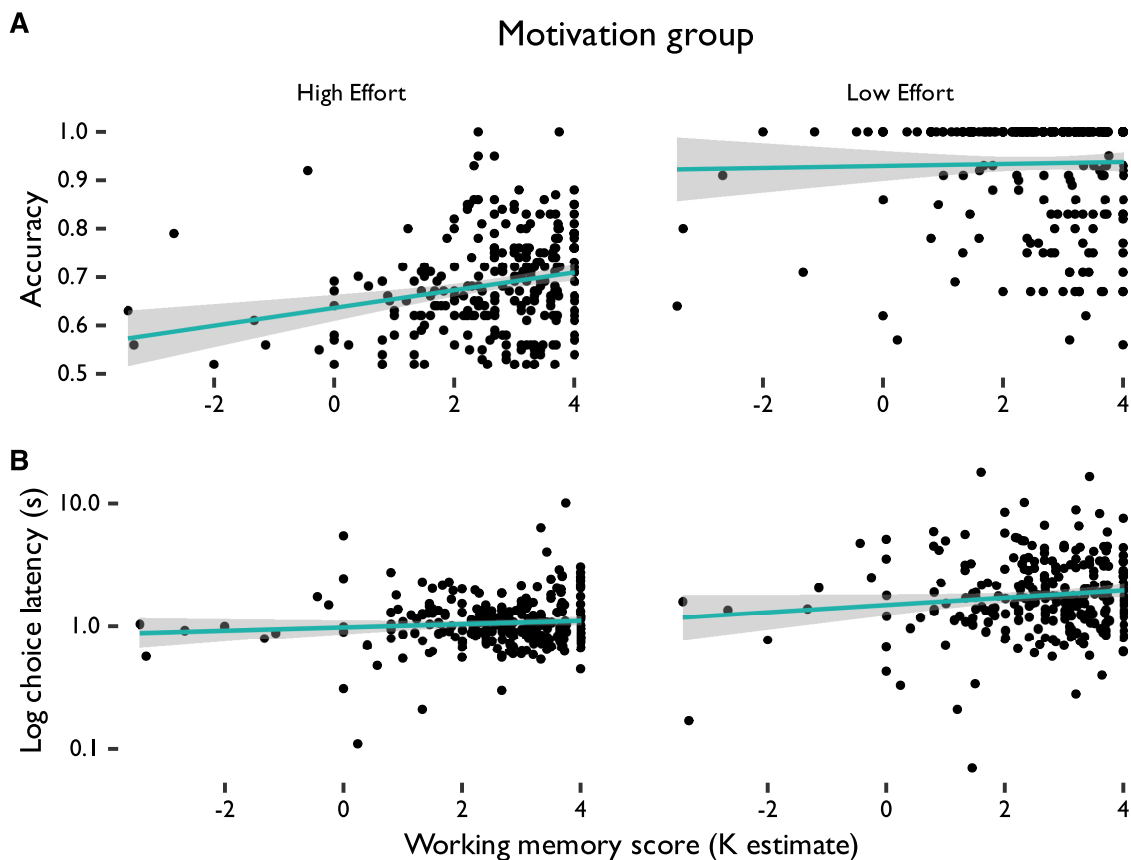


Figure 6. A, Accuracy and **B**, choice latency in the reward phase of the change detection task by working memory ability (K estimate). The choice latency plot has been log-transformed on the y-axis to more clearly show small choice latency values.

Table 6. Multi-level model analysis coefficients and standard errors for choice latency in the reward phase of the task

	Multi-level model: choice latency
(Intercept)	1.25*** (0.34)
Sex	-0.07 (0.17)
Working memory ability (K estimate)	0.05 (0.05)
BDI (prop. score)	-1.11* (0.54)
PSS	0.01 (0.01)
Reward anticipation	-0.04 (0.04)
Effort level	1.06*** (0.13)
SD (intercept participant)	0.18
SD (observations)	1.53
Num. Obs.	542
R ² Marg.	0.114
R ² Cond.	0.127
AIC	2042.9
BIC	2081.6
ICC	0.0
RMSE	1.51

BDI (prop. score), depression score on the beck depression inventory; PSS, perceived stress scale. Proportion scores are scores divided by total possible score. AIC, Akaike information criterion; BIC, Bayesian information criterion; ICC, intraclass correlation; RMSE, root mean squared error.

* $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

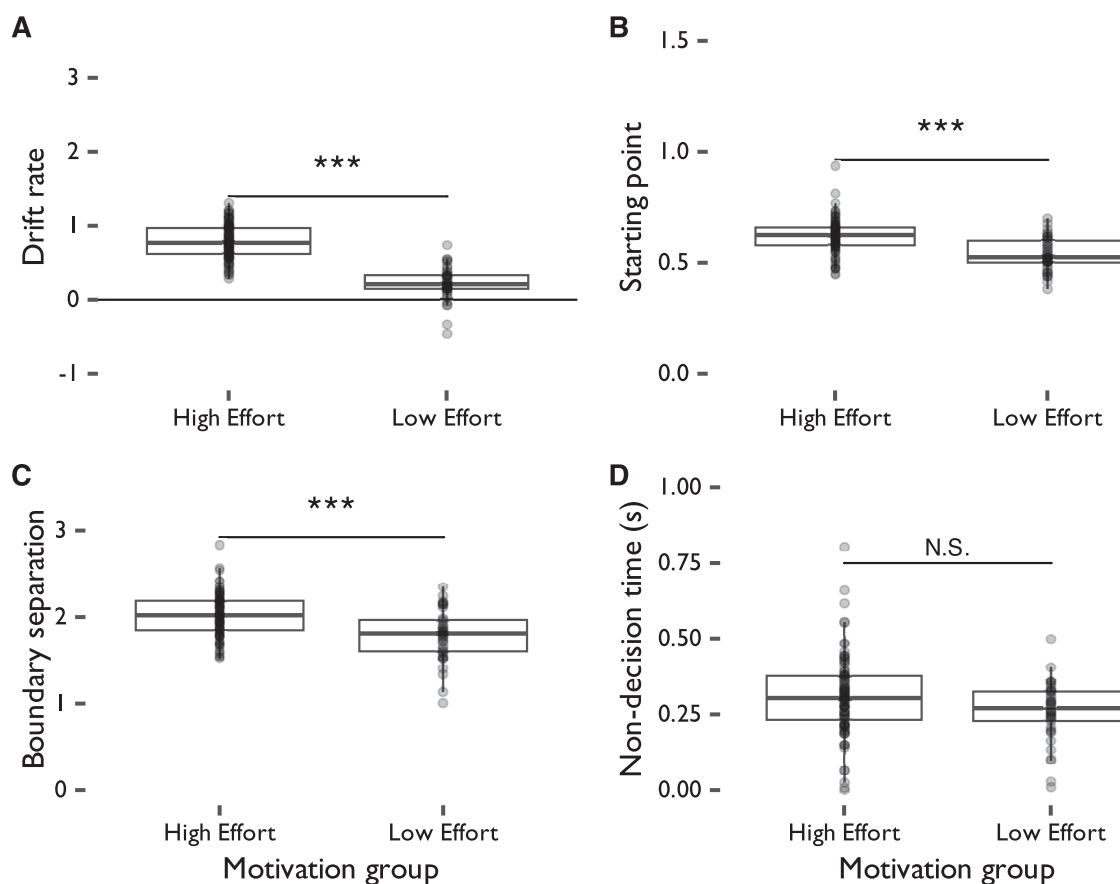


Figure 7. Outputs of a hierarchical Bayesian drift diffusion model fit to all participants in the effort choice task, divided by effort group. High effort participants had a significantly higher **A**, drift rate, **B**, starting point and **C**, boundary separation compared to low effort participants, but did not significantly differ on **D**, non-decision time compared to low effort participants.

To evaluate the quality of our DDM analysis, we conducted a parameter recoverability test by fitting this DDM to simulated binary choice and response time data that we generated using a Wiener process in the *rwiener* R package (Wabersich and Vandekerckhove, 2014) with the same average drift rate, starting point, boundary separation, non-decision time, and number of samples as our actual data. We evaluated the goodness of fit of the model to the simulated and the actual data using the leave-one-out cross validation information criterion (LOOIC) measure. Our model showed better parameter recoverability to the simulated data (LOOIC = $-8,997.31$) compared to the actual data (LOOIC = $8,080.26$). However, the parameter fits were significantly correlated between simulated and actual data ($t(134) = 2.94$, $p = 0.039$, $r = 0.25$), suggesting good parameter recoverability in our model.

Discussion

In this study, we investigated whether working memory performance, depressive traits, and chronic stress levels influence people's motivation and ability to consistently deploy cognitive effort for reward. Results showed that greater working memory ability significantly predicted the likelihood of systematically choosing high effort/HR trials (whether the participant was classified in the high vs low effort group), whereas chronic stress, depression trait levels, and reward anticipation were not predictive of effort deployment choices. Additionally, participants required more evidence to select LR trials compared to HR trials. DDM further indicated that those who were categorized as high effort participants required less evidence to select high effort/HR trials, and were more strongly biased toward selecting these trials, than low effort participants. Furthermore, working memory ability significantly predicted accuracy, while depressive traits predicted choice latency; both factors significantly differed depending on the effort required in the trial.

These results partially recapitulate findings from rodent research using the rCET regarding differences in performance and decision-making biases between those who select high effort trials more often and those who select them less often. As by Cocker et al. (2012), participants who chose high effort trials over 70% of the time performed better than those who chose high effort trials 70% or less of the time, on LR—although not HR—trials. Continually deploying high amounts of cognitive effort could be fatiguing; the high effort group's lower performance on HR trials could reflect the ineffectiveness

of this high effort choice strategy, or the fact that those with higher visuospatial short-term memory ability ($K > 3$) had a more difficult task with higher effort demands on high effort trials.

High effort group participants, who generally selected HRs more often and who were more biased toward selecting HR trials than low effort group participants, appear to have spent more time deliberating before selecting an LR trial than low effort group participants. Choice latency indicates how long participants spend when deciding which effort level to choose. Thus, they may have been trading off the benefits and costs of selecting a LR trial more consciously than low effort group participants were (Sandra and Otto, 2018; Pruessner et al., 2020), even as they performed significantly better than low effort group participants on LR but not HR trials. These tradeoffs could shift over time in the task as participants become fatigued or bored with the task, or found a new effort deployment strategy that better optimized rewards or task enjoyment.

Furthermore, our findings align with past work in rodents (Hales et al., 2024) in that our drift diffusion model results showed wider decision boundaries between HR and LR trials for high effort group compared to low effort group participants, as well as a starting point closer to HR trials for high effort group participants. Thus, we are the first to show in humans that evidence accumulation for low effort short-term memory tasks can be greater if participants consistently deploy high levels of effort. Given that the decision available on each trial is the same, individual differences in this initial level of bias toward selecting high effort trials—together with random noise—drive differences in trial difficulty choice between participants as modeled by the drift diffusion model. Therefore, the drift diffusion model may not be capturing meaningful, systematic patterns of effort deployment choices. However, it still provides an exploratory evaluation of the underlying cognitive processes that could be driving effort deployment choices. Additionally, as with rodents, participants in the low effort group made faster choices than those in the high effort group on LR trials. Furthermore, high effort group participants showed steeper drift rates than low effort group participants—while rodents in the equivalent high effort group had steeper drift rates when they subsequently made correct responses for their given choice (Hales et al., 2024). Taken together, these findings could suggest similarities between humans and rodents in the influence of cognitive capabilities on performance when deploying more cognitive effort, as well as similarities in biases toward deploying more effort or wanting higher rewards. Furthermore, given the conditionality of the drift rate findings in rodents, strategies driving impulsiveness or choice motivations to select an HR or LR trial could differ between humans and rodents. This could be explained by differences in the strength of reinforcement provided by the reward, which was a secondary reinforcement (points toward a monetary reward) in our task but a primary reinforcement (sugar pellets) in the rCET (Cocker et al., 2012; Hales et al., 2024), or by other differences between our task and the rCET as well as between humans and rodents in terms of how reward motivation is processed. These differences were not driven by reward anticipation, which did not differ between high and low effort groups ($t(122.95) = 0.17$, $p = 1$, $d = 0.02$).

Additionally, previous studies with the rCET have shown no overall accuracy differences between high and low effort group participants across either LR or HR trials (Hosking et al., 2016; Silveira et al., 2020). However, in a recent, larger analysis, Hales et al. (2024) showed that rodents choosing more HR trials were slightly more accurate on HR than LR trials when completing the trials themselves, while rodents choosing more LR trials were slightly more accurate on LR than HR trials when completing these trials. Although participants practiced task contingencies ahead of the main choice phase and the task difficulty was calibrated to working memory ability, individual differences in working memory—which predicted whether a participant selected high or low effort trials more often—still predicted these performance differences in humans. Furthermore, chronic stress—which can negatively impact availability of executive resources—did not explain accuracy or choice latency in this task. Thus, within the constraints of a binary effort choice task, participants' willingness to deploy cognitive effort was explained more by individual differences in executive function (working memory capability; Sandra and Otto, 2018) rather than affective factors or the effects of affective factors on executive function (Slattery and Cryan, 2017). Note that although we did not measure executive control as a global construct in this study, visuospatial working memory ability is typically described as a process that falls within the suite of executive functions (Miyake and Friedman, 2012).

Our drift diffusion model results also suggest that high effort group participants had to overcome a bias toward HR trials in order to select LR trials. This process could be driven by adjustments to effort deployment strategies based on performance, as suggested by recent interpretations of the expected value of control theory (Shenhav et al., 2021). Through this framework, choosing a high proportion of low effort trials could be seen as optimizing effort deployment and making slower and more considered decisions based on performance—as reflected in increased choice latency among high effort group participants for LR trials—whereas selecting a high proportion of high effort trials could entail more rapid and impulsive decisions. In order to evaluate whether participant performance was driven by trial-to-trial, model-free strategies—as opposed to a more considered strategy that incorporates performance over multiple trials—we conducted a follow-up linear regression analysis to evaluate whether performance on the previous trial predicts performance on the current trial as a function of effort motivation group and trial type. Performance on the prior trial, but not effort motivation group, significantly predicted performance on the current trial ($t(68955) = 8.45$, $p < 0.001$). This suggests that, overall, participants maintained a constant performance-based strategy that was not explained by effort choice decisions. Furthermore, participants with higher levels of depressive traits responded significantly faster across all trials, a result that runs counter to past work suggesting increased choice latency in reward tasks in rodents (Hales et al., 2023) or humans (Di Schiena et al., 2013) with depressive-like traits. However, after Bonferroni correction, depressive traits and choice latency were not

significantly correlated across all participants and trials ($t(702) = -1.21, p = 1, r = -0.05$). As depressive traits did not impact accuracy, these findings could instead suggest a lack of engagement with decision-making in the task characterized by higher levels of automatic responding given high levels of depressive traits (Teachman et al., 2012).

Our findings suggest that choices to deploy cognitive effort for reward are primarily driven by participants' cognitive capacity (as measured by visuospatial working memory ability) and not by depressive traits, chronic stress, or reward anticipation, in participants within a non-clinical range of depression traits and chronic stress. Potentially, the aversiveness of expending cognitive effort for participants with lower working memory capacity may have overridden overall sensitivity for reward. Alternatively, participants with lower working memory abilities may have determined that they were unlikely to do well on HR trials, and so may have been less incentivized to deploy effort on HR trials. Additionally, although our population exhibited a wide range of depression scores—including those above clinical thresholds (Wang and Gorenstein, 2013)—this population was non-clinical. In contrast, much work on the impact of reward anticipation on cognitive flexibility and related constructs has focused on participants who were clinically diagnosed with depression (Alloy et al., 2016; Terpstra et al., 2023) or on rodents in which chronic stress was directly induced (Watt et al., 2017). In order to obtain the highest accuracy and reward possible, participants may have titrated the task's difficulty—through choosing a higher proportion of LR trials—to a level at which they can consistently complete the task and at which the value of their effort was greatest (Shenhav et al., 2013).

We found no significant sex differences in performance or choice latency in our study. Women generally present more with depression than men (Parker and Brotchie, 2010) and are also more impacted by chronic stress across the lifespan (Hodes and Epperson, 2019). However, the results of our study were primarily driven by visuospatial short-term memory ability, where sex differences are overall smaller and women may be better than men at memory for location (Voyer et al., 2017)—a relevant measure in our task. Our undergraduate psychology student sample strongly skewed female; in future studies, it would be important to obtain a more balanced sample in order to further explore potential sex differences.

There are a number of caveats we must consider when interpreting these results. First, the study was conducted fully online on participants' own devices. While this allowed us to obtain a larger and better powered sample, differences in screen size and background distractions could have impacted participants' ability to encode the positions and colors shapes on each trial. These differences may have made the task more difficult than originally intended for some participants. However, participant performance was still near ceiling on LR trials and very high on HR trials, so this was unlikely to have a large effect.

Second, high effort trials never gave LRs, or vice versa. This may impact our ability to evaluate whether it was sensitivity to effort or to reward that more strongly drove participant choices in the task. However, reward anticipation did not predict choices for high effort deployment. This suggests that effort costs could be weighed more strongly than rewards by participants when making effort choices. In future studies, varying levels of effort and reward independently would allow us to explore whether differences in reward anticipation for small versus large rewards differentially influence decisions to deploy effort for reward. For example, if a participant only chooses to deploy effort for large rewards as opposed to for all rewards, they may pursue a strategy of only selecting trials that give large rewards regardless of the required effort level. This would allow us to better understand the role of reward magnitude in motivating effort choices.

Third, the set sizes used on the change detection task were modified from those originally used by Luck and Vogel (1997). The set size of 10 was removed from the pool of trials as pilot participants had very low performance at this set size. Although this change reduced the possible variety of trials used, participants still reported a maximum set size of eight to be highly challenging. Furthermore, the choice structure of the task incorporated only two of the set sizes. This could have made the task excessively easy for some participants—as reflected in the ceiling effect on LR trials—and excessively difficult to others. However, we did titrate task difficulty according to participants' visual short-term memory ability, reducing the likelihood of the task difficulty not matching participant ability. Future studies could use a finer-grained, parametric design that offers choices at more difficulty levels to examine whether motivation continuously varies with effort levels (Sayali et al., 2023).

Fourth, the method used to calculate the K estimate slightly differed between the task—used to set the criterion level—and our analysis—used to establish the K estimate for each participant. In the task, the K estimate was calculated based on the current hit and false alarm rate after every trial, with the final K estimate being used to set the criterion for the choice phase of the task. In our analysis, we calculated the K estimate based on the full hit and false alarm rate at each set size. Because of rounding differences in these calculations, the K estimate values used in our analysis differ from those in our task by $M = 0.01, SD = 0.63$. We consider this difference to be small enough as to not have an effect on the difficulty of the task.

Fifth, participants were told that the points they received would be proportional to the monetary reward received, which could motivate them toward higher effort choices regardless of their effort capabilities—shifting their effort/reward trade-offs. However, such a strong effort incentive was necessary to motivate participant performance given the difficulty of the high effort trials for many participants. Participants may have also individually differed in compliance with reading instructions or in their motivation for reward. Such factors could be investigated in a follow-up study incorporating qualitative evaluations of participant experiences of the task.

Sixth, differences in performance that are explained by the K estimate could also be explained by the differences in set sizes available associated with being above or below the K criterion of 3 on set size = 4. However, the set sizes at each level

of criterion were established with piloting such that the difficulty of each criterion level was matched as closely as possible. Last, although our task only had a single probe whose color had to be compared to the original probe, participants had to attend to the whole display to compare the single probe to the existing square at any given location onscreen. As such, we used the whole-display K estimate measure to calibrate participant performance in our task (Rouder et al., 2011). A single-probe K estimate measure reflects lower effort demands than the whole-display measure, as the whole-display measure requires the status of all shapes onscreen to be held in memory for the probe—whereas the single-probe measure only involves evaluating one shape in one position. However, as the position of the probe shape in our task is randomized, participants must still attend to the whole display to correctly identify any changes in the probe's color. As such, the whole-display K estimate measure is more relevant for our task's structure and cognitive effort demands, and the results could differ if we use the single-probe measure.

Future studies could make further use of information about effort deployment choices that can only be captured in human studies. As many human cognitive effort tasks—unlike rodent tasks—do not require substantial training, a future study could examine effort choices given a wider range of effort and reward levels, as well as test how preference for effort changes given high effort trials with LR and vice versa. Such a task could more richly capture individual differences in human cognitive effort deployment decisions. Furthermore, an expanding field of research uses qualitative approaches to investigate why participants make the judgments they do about effort deployment, as well as ask about participant experiences of task difficulty and the value of rewards versus the effort required to obtain them (Vásquez-Rosati et al., 2019). Understanding the phenomenology of cognitive effort deployment choices would help us extend our understanding of how and why we make decisions to deploy more or less cognitive effort for reward, beyond traditional self-report measures. Participant descriptions of their experiences in the task could indicate whether boredom and fatigue, in addition to reward motivation, impacted effort deployment choices. This approach would also allow us to further explore how rodent-based constructs of cognitive effort deployment are reflected in human decision-making, while extending our understanding of these constructs with human-specific experiential information. Additionally, exploring how the brain represents information about effort task difficulty—especially in regions relevant to cognitive control like the dorsal anterior cingulate cortex (Shenhav et al., 2016; Yee et al., 2021)—through functional magnetic resonance imaging would help us better understand how motivational states in cognitive effort decision-making are reflected in neural circuitry.

Our study suggests that, as in rodents, significant individual differences exist in the human tendency to choose harder but more rewarding options in cognitive effort tasks. Using a visuospatial working memory task, we show that individual differences in accuracy on the trial and choice latency for the trial type chosen are primarily driven not by depressive traits or chronic stress—as has previously been shown in the rodent literature—but instead by working memory capability. These findings may help to inform clinical interventions aimed at increasing motivation to seek rewards and engage in work in everyday life, and illustrate the importance of translating and extending rodent work with human-specific measures.

Data and Code Availability

All task code, stimuli, data, and code used to generate this manuscript and the figures are available at <https://osf.io/c4h7s/>.

References

- Ahn W-Y, Haines N, Zhang L (2017) Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Comput Psychiatry (Cambridge, MA)* 1:24–57.
- Alloy LB, Olino T, Freed RD, Nusslock R (2016) Role of reward sensitivity and processing in major depressive and bipolar spectrum disorders. *Behav Ther* 47:600–621.
- Altemus M, Sarvaiya N, Epperson CN (2014) Sex differences in anxiety and depression clinical perspectives. *Front Neuroendocrinol* 35:320–330.
- Beck AT, Steer RA, Brown GK (1996) *Manual for the beck depression inventory-II*. San Antonio (TX): Psychological Corporation.
- Birn RM, Roeber BJ, Pollak SD (2017) Early childhood stress exposure, reward pathways, and adult decision making. *Proc Natl Acad Sci USA* 114:13549–13554.
- Bulus M (2023) Pwrss: statistical power and sample size calculation tools. R package version 0.3.1, <https://CRAN.Rproject.org/package=pwrss>.
- Cocker PJ, Hosking JG, Benoit J, Winstanley CA (2012) Sensitivity to cognitive effort mediates psychostimulant effects on a novel rodent cost/benefit decision-making task. *Neuropsychopharmacology* 37:1825–1837.
- Crawford JL, Eisenstein SA, Peelle JE, Braver TS (2021) Domain-general cognitive motivation: evidence from economic decision-making. *Cognit Res Princ Implic* 6:4.
- Culbreth AJ, Moran EK, Barch DM (2018) Effort-cost decision-making in psychosis and depression: could a similar behavioral deficit arise from disparate psychological and neural mechanisms? *Psychol Med* 48:889–904.
- de Oliveira JRV (2023) Luck and Vogel change detection task. <https://gitlab.pavlovlab.org/joaorobertoventura/luck-and-vogel-change-detection-task>.
- Di Schiena R, Luminet O, Chang B, Philippot P (2013) Why are depressive individuals indecisive? Different modes of rumination account for indecision in non-clinical depression. *Cognit Ther Res* 37:713–724.
- Eichenbaum H (2017) Prefrontal–hippocampal interactions in episodic memory. *Nat Rev Neurosci* 18:547–558.
- Engstrom M, Landtblom A-M, Karlsson T (2013) Brain and effort: brain activation and effort-related working memory in healthy participants and patients with working memory deficits. *Front Hum Neurosci* 7:1–17.
- Forys BJ, Tomm RJ, Stamboliyska D, Terpstra AR, Clark L, Chakrabarty T, Floresco SB, Todd RM (2023) Gender impacts

- the relationship between mood disorder symptoms and effortful avoidance performance. *Eneuro* 10:ENEURO.0239–22.2023.
- Frömer R, Lin H, Wolf CKD, Inzlicht M, Shenhav A (2021) Expectations of reward and efficacy guide cognitive control allocation. *Nat Commun* 12:1030.
- Fuentes-Claramonte P, Ávila C, Rodríguez-Pujadas A, Ventura-Campos N, Bustamante JC, Costumero V, Rosell-Negre P, Barrós-Loscertales A (2015) Reward sensitivity modulates brain activity in the prefrontal cortex, ACC and striatum during task switching. *PLoS One* 10:e0123073.
- Grahek I, Musslick S, Shenhav A (2020) A computational perspective on the roles of affect in cognitive control. *Int J Psychophysiol* 151:25–34.
- Grahek I, Shenhav A, Musslick S, Krebs RM, Koster EHW (2019) Motivation and cognitive control in depression. *Neurosci Biobehav Rev* 102:371–381.
- Hales CA, Silveira MM, Calderhead L, Mortazavi L, Hathaway BA, Winstanley CA (2024) Insight into differing decision-making strategies that underlie cognitively effort-based decision making using computational modeling in rats. *Psychopharmacology* 241:947–962.
- Hales CA, Stuart SA, Griffiths J, Bartlett J, Arban R, Hengerer B, Robinson ES (2023) Investigating neuropsychological and reward-related deficits in a chronic corticosterone-induced model of depression. *Psychoneuroendocrinology* 147:105953.
- Hodes GE, Epperson CN (2019) Sex differences in vulnerability and resilience to stress across the life span. *Biol Psychiatry Neurobiol Resilience* 86:421–432.
- Hosking JG, Cocker PJ, Winstanley CA (2016) Prefrontal cortical inactivations decrease willingness to expend cognitive effort on a rodent cost/benefit decision-making task. *Cereb Cortex* 26:1529–1538.
- Ironside M, Kumar P, Kang M-S, Pizzagalli DA (2018) Brain mechanisms mediating effects of stress on reward sensitivity. *Curr Opin Behav Sci* 22:106–113.
- Kassambara A (2023) Rstatix: pipe-friendly framework for basic statistical tests. <https://CRAN.R-project.org/package=rstatix>.
- Kúkel'ová D, Bergamini G, Sigrist H, Seifritz E, Hengerer B, Pryce CR (2018) Chronic social stress leads to reduced gustatory reward salience and effort valuation in mice. *Front Behav Neurosci* 12:1–14.
- Kool W, McGuire JT, Rosen ZB, Botvinick MM (2010) Decision making and the avoidance of cognitive demand. *J Exp Psychol Gen* 139:665–682.
- Kuznetsova A, Brockhoff PB, Christensen RHB (2017) {lmerTest} Package: Tests in linear mixed effects models. *J Stat Softw* 82:1–26.
- Levenstein S, Prantera C, Varvo V, Scribano ML, Berto E, Luzi C, Andreoli A (1993) Development of the perceived stress questionnaire: a new tool for psychosomatic research. *J Psychosom Res* 37:19–32.
- Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279–281.
- Michely J, Viswanathan S, Hauser TU, Delker L, Dolan RJ, Grefkes C (2020) The role of dopamine in dynamic effort-reward integration. *Neuropsychopharmacology* 45:1448–1453.
- Miyake A, Friedman NP (2012) The nature and organization of individual differences in executive functions: four general conclusions. *Curr Dir Psychol Sci* 21:8–14.
- Parker G, Brotchie H (2010) Gender differences in depression. *Int Rev Psychiatry* 22:429–436.
- Pashler H (1988) Familiarity and visual change detection. *Percept Psychophys* 44:369–378.
- Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Kastman E, Lindeløv JK (2019) PsychoPy2: experiments in behavior made easy. *Behav Res Methods* 51:195–203.
- Posit team (2024) RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. <http://www.posit.co/>.
- Pruessner L, Barnow S, Holt DV, Joormann J, Schulze K (2020) A cognitive control framework for understanding emotion regulation flexibility. *Emotion* 20:21–29.
- Raghnath N, Fournier LR, Kogan C (2021) Precrastination and individual differences in working memory capacity. *Psychol Res* 85:1970–1985.
- Ratcliff R (1978) A theory of memory retrieval. *Psychol Rev* 85:59–108.
- R Core Team (2023) R: a language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria.
- Rizvi SJ, Quilty LC, Sproule BA, Cyriac A, Bagby RM, Kennedy SH (2015) Development and validation of the dimensional anhedonia rating scale (DARS) in a community sample and individuals with major depression. *Psychiatry Res* 229:109–119.
- Rouder J, Morey R, Morey C, Cowan N (2011) How to measure working memory capacity in the change detection paradigm. *Psychon Bull Rev* 18:324–30.
- Sandra DA, Otto AR (2018) Cognitive capacity limitations and need for cognition differentially predict reward-induced cognitive effort expenditure. *Cognition* 172:101–106.
- Sayali C, Heling E, Cools R (2023) Learning progress mediates the link between cognitive effort and task engagement. *Cognition* 236:105418.
- Scheggi S, De Montis MG, Gambarana C (2018) Making sense of rodent models of anhedonia. *Int J Neuropsychopharmacol* 21:1049–1065.
- Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79:217–240.
- Shenhav A, Cohen JD, Botvinick MM (2016) Dorsal anterior cingulate cortex and the value of control. *Nat Neurosci* 19:1286–1291.
- Shenhav A, Fahey MP, Grahek I (2021) Decomposing the motivation to exert mental effort. *Curr Dir Psychol Sci* 30:307–314.
- Sherdell L, Waugh CE, Gotlib IH (2012) Anticipatory pleasure predicts motivation for reward in major depression. *J Abnorm Psychol* 121:51–60.
- Silveira MM, Wittekindt SN, Ebsary S, Winstanley CA (2021) Evaluation of cognitive effort in rats is not critically dependent on ventrolateral orbitofrontal cortex. *Eur J Neurosci* 53:852–860.
- Silveira MM, Wittekindt SN, Mortazavi L, Hathaway BA, Winstanley CA (2020) Investigating serotonergic contributions to cognitive effort allocation, attention, and impulsive action in female rats. *J Psychopharmacol* 34:452–466.
- Slattery DA, Cryan JF (2017) Modelling depression in animals: at the interface of reward and stress pathways. *Psychopharmacology* 234:1451–1465.
- Stephan M, Volkman P, Rossner MJ (2019) Assessing behavior and cognition in rodents, nonhuman primates, and humans: where are the limits of translation? *Dialogues Clin Neurosci* 21:249–259.
- Teachman BA, Joormann J, Steinman SA, Gotlib IH (2012) Automaticity in anxiety disorders and major depressive disorder. *Clin Psychol Rev* 32:575–603.
- Terpstra AR, Vila-Rodriguez F, LeMoult J, Chakrabarty T, Nair M, Humaira A, Gregory EC, Todd RM (2023) Cognitive-affective processes and suicidality in response to repetitive transcranial magnetic stimulation for treatment resistant depression. *J Affective Disord* 321:182–190.
- Towbes LC, Cohen LH (1996) Chronic stress in the lives of college students: scale development and prospective prediction of distress. *J Youth Adolescence* 25:199–217.
- Treadway MT, Bossaller NA, Shelton RC, Zald DH (2012) Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *J Abnorm Psychol* 121:553–558.
- Treadway MT, Buckholz JW, Schwartzman AN, Lambert WE, Zald DH (2009) Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS One* 4:1–9.
- Unsworth N, Robison MK (2020) Working memory capacity and sustained attention: a cognitive-energetic perspective. *J Exp Psychol Learn Mem Cognit* 46:77–103.

- Vásquez-Rosati A, Montefusco-Siegmund R, López V, Cosmelli D (2019) Emotional influences on cognitive flexibility depend on individual differences: a combined micro-phenomenological and psychophysiological study. *Front Psychol* 10:1138.
- Voyer D, Voyer SD, Saint-Aubin J (2017) Sex differences in visual-spatial working memory: a meta-analysis. *Psychonomic Bull Rev* 24:307–334.
- Wabersich D, Vandekerckhove J (2014) The {RW}iener package: An {R} package providing distribution functions for the Wiener diffusion model. *The R Journal* 6:49–56.
- Wang YP, Gorenstein C (2013) Psychometric properties of the beck depression inventory-II: a comprehensive review. *Revista Brasileira de Psiquiatria* 35:416–431.
- Watt MJ, Weber MA, Davies SR, Forster GL (2017) Impact of juvenile chronic stress on adult cortico-accumbal function: implications for cognition and addiction. *Prog Neuro-Psychopharmacol Biol Psychiatry* 79:136–154.
- Westbrook SR, Hankosky ER, Dwyer MR, Gulley JM (2018) Age and sex differences in behavioral flexibility, sensitivity to reward value, and risky decision-making. *Behav Neurosci* 132:75–87.
- Westbrook A, Kester D, Braver TS (2013) What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS One* 8:e68210.
- Westbrook A, van den Bosch R, Määttä JI, Hofmans L, Papadopetraki D, Cools R, Frank MJ (2020) Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science* 367:1362–1366.
- Yee DM, Adams S, Beck A, Braver TS (2019) Age-related differences in motivational integration and cognitive control. *Cognit Affective Behav Neurosci* 19:692–714.
- Yee DM, Crawford JL, Lamichhane B, Braver TS (2021) Dorsal anterior cingulate cortex encodes the integrated incentive motivational value of cognitive task performance. *J Neurosci* 41:3707–3720.