# On the misery of cognitive effort

**Jiazhou Chen[1-2], Quentin J.M. Huys[2,4], Argyris Stringaris[3,5], Dylan M. Nielson[6]**

**1** Section on Development and Affective Neuroscience, National Institute of Mental Health, Bethesda, MD, USA

**2** Applied Computational Psychiatry Lab, Mental Health Neuroscience Department, Division of Psychiatry and Max Planck Centre for Computational Psychiatry and Ageing Research, Queen Square Institute of Neurology, UCL

**3** Division of Psychiatry, University College London

**4** Camden and Islington NHS Foundation Trust, London, UK

**5** First Department of Psychiatry, Aiginiteion Hospital, National and Kapodistrian University of Athens **6** Machine Learning Team, National Institute of Mental Health, Bethesda, MD, USA

**Corresponding Author**
Mr. Jiazhou Chen
Division of Psychiatry
University College London
Maple House
149 Tottenham Ct Rd
London, W1T 7NF, U.K.
jiazhou.chen.20@ucl.ac.uk

**Abstract**

The effect of cognitive effort on mood is unclear. Expending cognitive effort is generally avoided but expenditure of cognitive effort is the primary feature of some popular recreational activities, such as sudoku or video games. However, one common confound is that previous studies mostly looked into mood changes before and after cognitive effort manipulation. Therefore, it is unknown what immediate impact cognitive effort has on momentary mood. To investigate this, in two studies, we used a letter sorting paradigm to test whether momentary mood ratings change with different levels of experimentally manipulated task difficulty. In study 1 (N = 105), we found that increased difficulty, and consequently cognitive effort, leads to more errors and lower mood ratings. In study 2 (N = 210), we replicated our results from study 1 and further demonstrated that cognitive effort influences mood independently from immediate reward. In conclusion, controlling for reward conditions, we found that cognitive effort reduces people's moment to moment mood.

# 1 Introduction

Whether we plan our wedding, think through a problem or engage in fun activities, our brain has to perform complex computational operations. Such neural computations incur a variety of costs. While devoting our neural machinery to one problem, we forego opportunities that may have been reaped by focusing on alternatives. Furthermore, neural computations incur hard energy costs. Whether to engage in neural computation is hence a complex problem, one that requires potential gains to be traded off against potential costs (Russell & Wefald, 1991). Indeed, humans appear to be sensitive to cognitive effort costs, and often avoid cognitive effort both implicitly (e.g. Callaway et al. 2022; Huys et al. 2012) and explicitly (e.g. Botvinick et al. 2009; Inzlicht et al. 2018; Vogel et al. 2020; Westbrook et al. 2013). In fact, individuals can show preference for pain over a task demanding a high level of cognitive effort (Vogel et al., 2020). This strong level of avoidance suggests that expending cognitive effort, or the prospect of doing so, is both implicitly and explicitly unpleasant and punishing in some form (Aw et al., 2011; Ellis et al., 2022; Grèzes et al., 2021; Het & Wolf, 2007; Mintz, 2010).

However, this simplistic view of cognitive effort as computationally costly and aversive belies a more complex picture. First, the experienced feeling of expending cognitive effort does not necessarily reflect true energetic costs. The amount of calories consumed during a cognitive task is negligible (Kurzban, 2010), and measures of brain blood glucose do not change with extensive cognitive effort (Madsen et al., 1995). The experienced feeling of cognitive effort rather appears to reflect a subjective and biased assessment of the value of resource investment (Bijleveld, 2018; Shenhav et al., 2017).

Secondly, studies on work productivity and sports training have also shown that people seek out and appear to enjoy effortful activities, ranging from moutain climbing to a variety of intellectual pursuits (Inzlicht et al., 2018; Kurniawan et al., 2013; Shenhav et al., 2017; Wang et al., 2017). Here, the expenditure of cognitive effort, and the prospect of doing so, are associated with a positive affective experience. Although, in part, this may arise when the immediate costs of cognitively effortful behaviors are offset by the promise of longer-term or alternative rewards. For instance, during the effortful process of planning, current expenditure of cognitive effort is associated with expected gain in future reward (Daw et al., 2011; Ferstl et al., 2022; Ho et al., 2022). In social situations, individuals might opt to engage in cognitively effortful pro-social behavior to pursue better social outlook or others' reciprocation (Batson, 1987; Godman et al., 2014; Vaish et al., 2016). Indeed, planning tendencies reduce when without longer-term benefits (Callaway et al., 2022; Kool et al., 2016), and individuals reduce their pro-social behavior with increasing cognitive effort demands (Lockwood et al., 2017).

Lastly, there are strong suggestions of a bidirectional relationship between mood and effort. On one hand, mood clearly affects willingness to exert cognitive and physical effort. In the extreme, psychopathologcial states such as mania feature increased willingness to expend cognitive and physical effort.(American Psychiatric Association, 2013). Similarly, low mood and pathological states such as depression reduce the willingness to exert cognitive and physical effort (Anderson et al., 2019; Ganesan, 2020). Indeed, fatigue and difficulties in concentration are diagnostic symptoms of depression, and both speak to a reduced propensity to expend effort (American Psychiatric Asso-

ciation, 2013). Outside of psychopathology, studies have shown that experimentally induced mood impacts cognitive effort expenditure as well (Brinkmann & Gendolla, 2008; Gendolla & Krüsken, 2002; Joana, 2009; Treadway et al., 2009). On the other hand, while mood states are responsive to appetitive and aversive events (Diener et al., 2009; Grosscup & Lewinsohn, 1980; Philip, 1971; Stone & Neale, 1984), it is unknown how cognitive effort influences mood at the moment of expenditure. There is strong evidence for the influence of appetitive events on mood, and current evidence suggests that momentary mood reflects the statistics of recent rewards (Keren et al., 2021; Liuzzi et al., 2021). Considering that cognitive effort is both aversive and closely linked to reward, one could speculate that cognitive effort itself should have an impact on mood.

Furthermore, there is a question as to what the direction of this effect might be. As mentioned above, both a positive (mood-enhancing) direction, and a negative (mood-lowering) direction are conceivable. To our knowledge, this has not been examined. While previous studies examined how cognitive effort changes affective experience (Erber & Tesser, 1992; Ferstl et al., 2022; R. J. Larsen et al., 1986; S. E. Larsen & Berenbaum, 2014; Robinson & Morsella, 2014; Sonnentag & Grant, 2012), there are no parametric manipulations of cognitive effort while assessing momentary mood in an experimental setting.

Here, we therefore examined the impact of parametrically varied cognitive effort demand on the momentary affective experience at a fine temporal scale. Specifically, we were interested in how changes in cognitive effort exertion over time relate to temporal fluctuations in affective states.

To examine the impact of cognitive effort, we combined this momentary mood assessment with a letter sorting paradigm commonly used for the study of the neurobiology of cognitive effort (Jansma et al., 2006). This task allowed a parametric manipulation of cognitive effort by varying short-term memory load. The combination of these two tasks allowed us to study how variations in cognitive effort impact on momentary affective states.

# 2 Methods

## 2.1 Tasks

Two cognitively demanding tasks were used to investigate the effects of cognitive effort, as indexed by difficulty, on momentary mood. In both tasks, we asked participants to perform letter sorting tasks with varying difficulty and report their trial-by-trial mood ratings. A large body of literature also demonstrated the influence of reward processes on affective experience (Keren et al., 2021; O'Callaghan & Stringaris, 2019; Rutledge et al., 2017; Rutledge et al., 2014). Reward process is is also closely linked to the motivation to expend cognitive effort (Frömer et al., 2021; Massar et al., 2018; Yang et al., 2014). Therefore, we also included an explicit manipulation on reward magnitude as a control condition to account for possible reward effects.

### 2.1.1 Multi-Attempt Letter Task

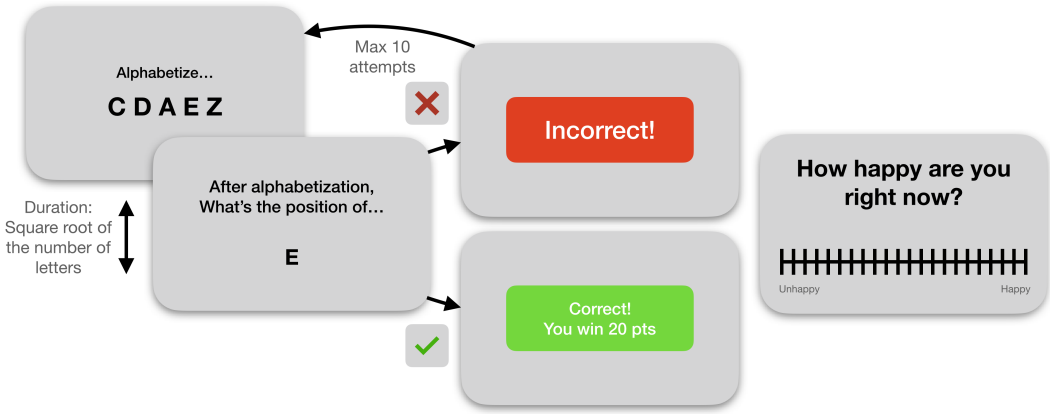


Figure 1: The Multi-attempt Letter Task. In each trial, participants viewed a string of letters of length $n$ for $\sqrt{n}$ seconds. Participants then had to mentally order these letters alphabetically and were asked to respond about the position of one specific randomly chosen letter. The difficulty of each trial was manipulated by the length of the string presented on that trial, number of letters ($n$; ranging from 3 to 9) . Feedback was provided after each response. If the response was incorrect, participants were returned to the same problem for up to a maximum of 10 attempts. After 10 failed attempts, the trial would end without bonus points being awarded. If the response was correct, participants saw the bonus points earned. The magnitude of the bonus points ranged from 30 to 90 in 10 increments. If no responses were provided, the trial is counted as missed and participants would proceed as if it was incorrect. At the end of each trial, participants were asked to self-report their momentary mood level using a visual-analog scale ranging from 0 to 10 (0 labeled unhappy and 10 labeled happy), as done in previous work (Keren et al., 2021; Liuzzi et al., 2021; Rutledge et al., 2014, 2015).

An overview of the Multi-attempt Letter Task is shown in Fig. 1. There were 168 trials, each with a unique strings. The difficulty and reward magnitude were fully randomized and temporally orthogonal to each other, such that the participants experienced the full range of reward magnitude for each level of difficulty.

### 2.1.2 Single-Attempt Letter Task

The Single-Attempt Letter Task was identical to the Multi-Attempt Letter Task except that, as the name suggests, participants were allowed only one attempt for each trial. In addition, the task also only provided sparse performance feedback. Feedbacks were only provided randomly on a total of 7 trials, with a between feedback interval ranging from 15 and 45 trials in 5 increments. Feedback consisted of both percentage and counts of correct trials (Fig. 2). To ensure identical overall duration between the Single-Attempt and the Multi-Attempt Letter Tasks, the number of trials was increased to 210. The Single-Attempt Letter Task was preregistered prior to data collection and analysis (osf link).

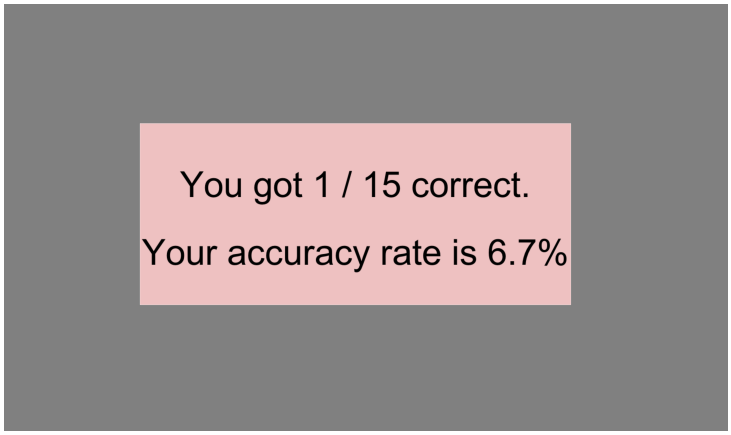


Figure 2: An example of the feedback on participant performance in the Single-Attempt Letter Task.

## 2.2 Participants and Procedure

Participants were recruited through online recruitment platforms Amazon Mechanical Turk (MTruk) ("Amazon Mechanical Turk", 2023) and Prolific ("Prolific", 2023) following standard online recruitment protocols (Aguinis et al., 2021; Mortensen & Hughes, 2018).

Before their participation, recruited individuals were presented with informed consent on a web page, where they must fully read and agree before continuing. Because both MTruk and Prolific mask personally identifiable information, including email, and that we only collected behavioral responses and no clinical assessments, our studies were determined to be exempt from institutional review board review by the National Institutes of Health (NIH) Office of Human Subjects Research Protections. The protocol on the consent processes and collection of behavioral responses was approved by the NIH Office of Human Subjects Research Protections.

At the beginning of the study, participants were first instructed on the task design and that their bonus earning would be correlated with their task performance. Participants were reimbursed for their time via respective recruitment platforms, along with a performance-based bonus payment (later assigned). For the Multi-Attempt Letter Task, the sample consisted of 27 pilot participants, along with 118 unique participants for the main sample. For the Single-Attempt Letter Task, we collected data from 320 unique participants that had not participated in the Multi-Attempt Letter Task previously.

## 2.3 Data Analysis and Modelling

We excluded participants based on three criteria:

1. **percentage of missing data:** A trial was considered missing if no response was provided within the time limit. In the case of the Multi-Attempt Letter Task, where multiple attempts were allowed, a trial was considered missing if all 10 attempts had no responses. A percentage was calculated based on the number of trials missed and participants who were missed more than 20% of the total data were

excluded.

2. **number of consecutive identical key-presses**: The number of consecutive key-presses was defined as the number of repetition of the same key. A high number of repeated key-presses was indicative of a lack of engagement. Thus, we excluded participants who had consecutive identical key presses on more than 10% of the total trials.

3. **probability of the average number of errors per trial / accuracy count** : For participants who were actively engaged, they should perform better than chance in this task. To determine if a participant was performing better than chance, we implemented a simulation-based method to generate synthetic behaviors as if responses were made randomly. Based on the probability density function of generated data, we calculated metrics that a participant needed to exceed to be considered statistically unlikely that they were providing answers randomly (significance level = 0.05). Because each task has different design elements, we employed different metrics for each tasks. For the Multi-Attempt Letter Task, the metric was the average number of errors per trial, and its threshold was 5.23. For the Single-Attempt Letter Task, the metric was total accuracy count, and the threshold was 49.

Using these criteria, we included 105 participants in the Multi-Attempt Letter Task sample and 210 participants in the Single-Attempt Letter Task sample.

We calculated Pearson correlation to investigate the association between difficulty ($n$) and number of errors ($e$, number of repeats before a correct response). We calculated the correlation estimate, test statistics and probability using the cor.test function in the R (version 4.1.3) stat package (R Core Team, 2020). To examine to what extent trial-by-trial mood ratings $m$ could be predicted by preceding trial events we built linear mixed-effects regression models using the lme4 (Bates et al., 2015) package from R (R Core Team, 2020). We built four linear mixed effects models as follows. The full model explained the mood $m_{t,i}$ reported by subject $i$ on trial $t$ as:

$$m_{t,i} = (\alpha_0 + \beta_{0,i}) + (\alpha_n + \beta_{n,i})n_{t,i} + (\alpha_e + \beta_{e,i})e_{t,i} + (\alpha_r + \beta_{r,i})r_{t,i} + (\alpha_\tau + \beta_{\tau,i})t + \epsilon_{t,i} \quad (1)$$

where $n_{t,i}$ is the $i$'th subject's difficulty level on trial $t$, $e_{t,i}$ indicates whether the trial was an error (1 = error, 0 = no error), $r_{t,i}$ indicates reward bonus obtained on that trial. Fixed effects are denoted by $\alpha$ and random effects varying for each subject by $\beta$. Three component models were examined, containing either difficulty, error or reward terms in addition to the intercept ($\alpha_0, \beta_0$) and time passage effects ($\alpha_\tau, \beta_\tau$), such as the mood drift over time effect (Jangraw et al., 2023). In each model, mood ratings were the dependent variable, and task conditions were independent variables.

To examine the effects of task events in the Single-Attempt Letter Task, we constructed another similar model:

$$m_{t,i} = (\alpha_0 + \beta_{0,i}) + (\alpha_n + \beta_{n,i})n_{t,i} + (\alpha_\lambda + \beta_{\lambda,i})\lambda_{t,i} + (\alpha_f + \beta_{f,i})f_{t,i} + (\alpha_\tau + \beta_{\tau,i})t + \epsilon_{t,i} \quad (2)$$

where the term $f_{t,i}$ indicates whether a feedback was displayed ($f = 1$) or not ($f = 0$), $\lambda_{t,i}$ indicates the number of trials passed since last feedback was displayed. We included the $f^i_{t,i}$ and $\lambda_{t,i}$ terms to capture the effect of feedback as well as its temporal influence on momentary mood.

All models included trial number as a predictor to capture time-related mood changes. Tests for linearity are shown in the supplemental results.

The fixed effect size $d$ was estimated as:

$$d = \frac{\Delta\mu}{\sigma^2_{\text{rand}} + \sigma^2_{\text{res}}} \tag{3}$$

where $\Delta\mu$ is the difference between the means, $\sigma^s_{\text{rand}}$ is the random effects and $\sigma^2_{\text{res}}$ the residual variance.

# 3 Results

## 3.1 Multi-Attempt Letter task

We first performed a manipulation check and examined whether longer strings were indeed more difficult and led to more errors. Figure 3B shows that a higher number of letters $n$ indeed led to a higher number of errors $e$ (correlation estimate = 0.225, 95% confidence interval (C.I.) = [0.211, 0.239], $t_{17617} = 30.69$, p $<$ 0.001), suggesting that longer strings were more difficult for participants to order alphabetically.

Next, we examined which aspects of the trials influenced momentary mood ratings. The component models showed that number of letters ($n$), errors ($e$), and reward ($r$) were individually associated with the mood ratings (Table 1, figure 3A). Importantly, while reward had a positive effect, both difficulty and errors had a negative effect. The effect of time was always negative, such that with increased in trial number, self-reported trial-by-trial mood rating decreased, which is consistent with previous studies showing a mood drift effect in task (Jangraw et al., 2023).

However, number of letters ($n$), error ($e$) and reward term ($r$) were partially correlated, in part by design. To control for this, we examined the full model containing all effects (Eqn. 1).

This model showed that while the error and the reward terms remained significant positive and negative predictors of momentary mood, respectively, the number of letters term was no longer significant (Table 1).

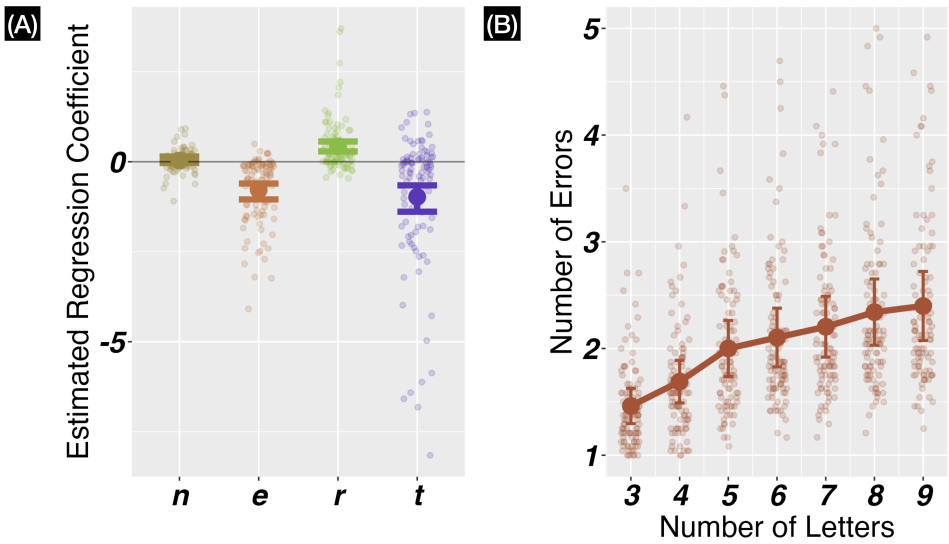


Figure 3: Multi-Attempt Letter Task Results. A): The estimated fixed effects of number of letters ($n$), number of errors ($e$), reward magnitude ($r$), and trial ($t$) on trial-by-trial mood rating. Each data point represents the estimated random effects varying for each subject. Error bar illustrates the standard error of the mean. B): Number of error ($e$) for difference levels of number of letters ($n$). Longer strings were more difficult. The number of errors increased with the length of the letter strings. Dots show the mean number of errors for each participant for a particular number of letters.

| Model | $R^2_{marginal}$ | $R^2_{conditional}$ | Effect | Estimate | $t$-value | $p$-value | Effect size |
|---|---|---|---|---|---|---|---|
| Difficulty | 0.023 | 0.827 | $n$ | -0.300 | -5.823 | <0.001 | 0.082 |
| | | | $t$ | -0.894 | -4.815 | <0.001 | 0.244 |
| Error | 0.040 | 0.859 | $e$ | -0.784 | -8.705 | <0.001 | 0.207 |
| | | | $t$ | -0.970 | -5.313 | <0.001 | 0.256 |
| Reward | 0.027 | 0.840 | $r$ | 0.491 | 6.193 | <0.001 | 0.135 |
| | | | $t$ | -0.917 | -4.965 | <0.001 | 0.253 |
| Full | 0.043 | 0.873 | $n$ | 0.028 | 0.662 | 0.509 | 0.007 |
| | | | $e$ | -0.758 | -8.538 | <0.001 | 0.203 |
| | | | $r$ | 0.417 | 5.679 | <0.001 | 0.111 |
| | | | $t$ | -0.974 | -5.337 | <0.001 | 0.260 |

Table 1: Linear mixed effect models predicting trial-by-trial mood ratings. $n$ denotes the difficulty or number of letters; $e$ indicates the number of errors; $r$ represents the reward magnitude; $t$ is the trial number. The degree of freedom for all the independent variables in the models is 104.

## 3.2 Single-Attempt Letter Task

In the first version of the task, the effect of cognitive effort appeared to be accounted for by the more frequent experience of accuracy negative feedback. However, this may have been a statistical artifact due to issues of collinearity, such that participants both made increased number of errors and experience higher degrees of cognitive effort in more difficult trials . To test this, we designed a new version of the letter task with sparser feedback that allowed a better distinction between the negative impact of cognitive effort and negative accuracy feedback on mood.

Participants' accuracy rate in the sparse feedback version was again lower in trials with more letters $n$ (Figure 4B). This was confirmed by a mixed effects logistic regression model, showing that the effects of number of letters on trial-by-trial accuracy is significant (estimate = -0.484, std. error = 0.028, z = -17.145, p $<$ 0.001, effect size = 0.484), controlling for trial number.

To investigate if participants' accuracy is impacted by the joint effect of time-in-task and difficulty, we included an interaction between trial number and number of letters. The mixed effects logistic regression model also revealed a significant negative interaction (estimate = -0.061, std. error = 0.013, z = -4.831, p$<$0.001, effect size = 0.061), such that trial-by-trial accuracy is even lower in later and more difficult trials. These effects persist even after adjusting for chance probability of each difficulty level (Figure 4C).

With the removal of trial-by-trial feedback in this task, we found that difficulty ($n$) was a significant negative predictor of momentary mood, (Table 2, Figure 4A). We also examined the effect of the sparse feedbacks ($f$) on momentary mood. Mood ratings were lower after feedback regardless of its valence. It is important to note that 64.63% of the feedback showed below average accuracy rate. However, participants' mood slowly recovered from the negative impact of feedback afterwards, as indicated by the significant positive effect of the term number of trial since last feedback ($\lambda$).

| Model | $R^2_{marginal}$ | $R^2_{conditional}$ | Effect | Estimate | $t$-value | $p$-value | Effect size |
|---|---|---|---|---|---|---|---|
| Full | 0.005 | 0.875 | $n$ | -0.168 | -7.826 | $<$0.001 | 0.166 |
| | | | $f$ | -0.174 | -2.449 | 0.015 | 0.172 |
| | | | $\lambda$ | 0.028 | 2.510 | 0.013 | 0.027 |
| | | | $t$ | -0.146 | -3.645 | $<$0.001 | 0.144 |

Table 2: Linear mixed effect models predicting trial-by-trial mood ratings. $n$ denotes the difficulty or number of letters; $f$ indicates whether performance feedback was displayed; $\lambda$ represents the number of trials elapsed since last feedback was displayed; $t$ is the trial number, which we use to infer the mood drift over time effect and time in task effect. The degree of freedom for all the independent variables in the models is 209.

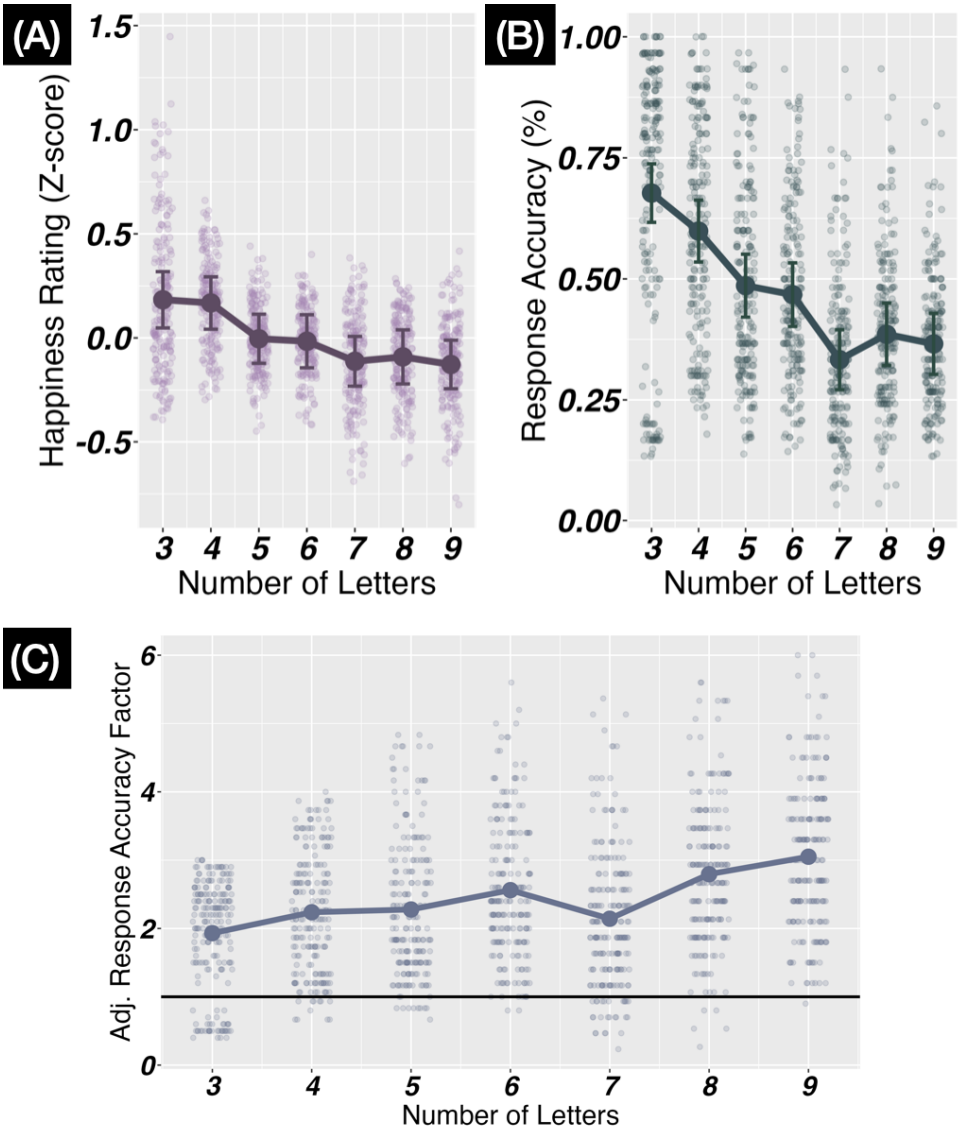


Figure 4: A): The estimated fixed effects of number of letters ($n$) on trial-by-trial mood rating (z-scored). Each data point represents the estimated random effects varying for each subject. Error bar illustrates the standard error of the mean. B): Responses to longer strings were less accurate. Accuracy decreases with the length of the letter strings. Dots indicate the mean accuracy rate given participant and number of letters. C): Participants accuracy are all above chance across all difficulty conditions. On the y axis is the adjusted response accuracy factor which was calculated as the accuracy rate divided by chance. Chance is different for each number of letter. As the number of letter increases, the chance probability will decrease, it can be calculated as $1/n$.

# 4 Discussion

We examined the relationship between cognitive effort and momentary mood. In two tasks, a parametric increase in the exertion of cognitive effort reduced momentary mood. In the first study, increased cognitive effort led to an increase in errors, and subsequent repeat of the same trial. Therefore, in this task, effort was partially confounded by other factors such as duration, negative feedback and possibly fatigue. These confounds were removed in the second study allowing a direct estimate of the impact of cognitive effort on momentary mood. This confirmed that effort exertion has an overall negative effect on momentary affective states on short time scales.

A negative effect of cognitive effort on mood is consistent with both the view of cognitive effort as computational cost, and of momentary mood ratings as running estimates of reward rates (Bennett et al., 2021; Rutledge et al., 2014). In terms of the former, we have noted previous failures to show an obvious energetic correlate of cognitive effort expenditure (Kurzban, 2010; Madsen et al., 1995). As such, the negative affective consequence of cognitive effort could be due to opportunity costs. Opportunity costs arise when the execution of one activity means that the rewards for other activities cannot be earned. Indeed, when comparing the two studies, the impact of cognitive effort on mood in the first study was particularly marked when subjects were forced to repeat the same task after errors. While this result is obviously confounded by the negative feedback itself, the fact that errors in this task cost participants additional time for each attempt support the idea that opportunity costs contribute to the negative affective response to cognitive effort. In the absence of a clear energetic cost for cognitive effort, the opportunity cost itself is then likely to derive from the limitations on cognitive capacity (Inzlicht et al., 2018; Sandra & Otto, 2018). The existence of a opportunity cost on cognitively effortful actions is necessary for the effective allocation of cognitive resources. Otherwise, prioritization of cognitive resources for one task would not mean that alternative tasks cannot be processed.

As for the latter, the question is why should momentary mood or affective judgements reflect the opportunity costs of cognitive effort, i.e., why should exerting cognitive effort change how we feel? Numerous studies over the past decade have shown that momentary mood ratings quantitatively reflect recent and history of rewards and losses (Eldar et al., 2018; Emanuel & Eldar, 2023; Keren et al., 2021; Liuzzi et al., 2021; Rutledge et al., 2017; Rutledge et al., 2014, 2015). Specifically, momentary mood is predicted by not only reward history but also an average of recent positive and negative prediction errors, i.e. whether current experience exceeded expectations or failed to meet them (Keren et al., 2021; Liuzzi et al., 2021; Rutledge et al., 2014). Monitoring this rate of reward has been argued to be useful for inferring underlying environmental changes (de Boer et al., 2017; Kumar et al., 2018; Packheiser et al., 2021) and to facilitate future behavior (Farrell et al., 2022; Gläscher et al., 2010; Rouhani & Niv, 2021; Schultz et al., 1997). This is because the average reward rate is a measure of the ongoing opportunity cost: an environment with high opportunity cost is one with high reward rate, i.e., one in which acting slowly causes many rewards to be missed (Niv et al., 2007). Thus, this motivates high vigor and fast action. A similar argument can be made for cognitive effort (Brinkmann & Gendolla, 2008; Treadway et al., 2012; Treadway et al., 2009). In this case, low effort—or slow computations—would lead to a major reduction in reward, hence, motivating high cognitive effort expenditure.

In our study, we assumed a linear relationship between cognitive effort and task difficulty, and we did not directly measure the sensation of cognitive effort. There is a possibility that mapping between cognitive effort and difficulty is not linear. First, while difficulty creates demand for cognitive effort, it is possible that various factors, such as individual ability, disengagement, and fatigue, could all lead to low cognitive effort exertion in the face of high difficulty. However, we demonstrated that reaction time, which is often used as a approximation of cognitive effort exertion (Ganesan, 2020; Robinson & Morsella, 2014; Robles & Johnson, 2017), is indeed modulated by task difficulty (see supplemental material: Reaction Time Analysis).

The study has some limitations. It relies on a single subjective mood rating to measure momentary mood, and we cannot exclude that other aspects of momentary mood could be affected differently. However, mood ratings to appear to index a general state of well-being (Eldar et al., 2016; Keren et al., 2021; Liuzzi et al., 2021; Rutledge et al., 2014), suggesting that the findings may generalize to a certain extent. Additionally, initial mood ratings in both studies showed an above average score (see supplemental material: Baseline Mood Rating Analysis). Although, we showed that initial mood has no impact on the effects of cognitive effort on mood.

Our study also has strengths. One of the main strength is the large sample size based on detailed power analyses; c.f. supplemental material). Such large samples have become more common with online cognitive testing (Gillan & Daw, 2016). Importantly, the demographics of online recruitment platforms are similar to the general population (Huff & Tingley, 2015; McCredie & Morey, 2019; Redmiles et al., 2019), with the main discrepancies in terms of lower affect and social engagement (McCredie & Morey, 2019; Shapiro et al., 2013). The design involving two studies allowed us to replicate our main findings, and we note that the second study (the Single-Attempt Letter Task) was pre-registered. Importantly, the second study also showed that the effects remain when controlling for (or removing) explicit rewards and losses.

In conclusion, across two studies, cognitive effort induced by task difficulty resulted in reduced self-report moment-to-moment mood ratings. Viewed from a formal setting of reinforcement learning, the impact of cognitive effort on mood is in keeping with the putatively normative roles of both effort and mood, providing insight into potential mechanism for cognitive resource allocation.

# 5 Acknowledgement

The authors would like to thank Hanna Keren, Lucrezia Liuzzi, Song Qi, and Charles Zheng for their early feedback on the study design; Isabel Berwian, Tore Erdmann, Jenny Fielder, Anna Hall, Mina Ilieva, Jolanda M Malamud, Anahit Mkrtchian, Agnes Norbury, Jakub Onysk, Jade Serfaty, Lana Tymchyk for their feedback on the manuscripts.

# 6 Declarations

## 6.1 Funding support and Conflict of Interests

J.C. was supported by the Intramural Training Program of the National Institute of Mental Health (NIMH, part of the National Institutes of Health, Bethesda, MD, U.S.A.), as part of the UCL-NIMH Joint Doctoral Program in Neuroscience. D.M.N was supported by the Intramural Research Program of the NIMH (grant no. ZICMH002968, to Francisco Pereira). Q.J.M.H has received consultancy fees and options from Aya Health and Alto Neuroscience and a research grant from Koa Health. J.C., D.M.N. and A.S. have no conflict of interests to report. The work was supported by the UCL NIH BRC (London, U.K.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The views expressed in this article do not necessarily represent the views of the NIMH, the Department of Health and Human Services or the United States Government.

## 6.2 Ethics approval

Experimental protocol is approved by the National Institutes of Health Office of Human Subjects Research Protection (Protocol Number: P194594).

## 6.3 Data and Analysis Scripts

Data from both studies are made available at this Open Science Framework (osf.io) depository. The experiment code and analysis scripts are available at this public Github repository.

## 6.4 Author Contribution

Conceptualization: J.C., Q.J.M.H, A.S., and D.M.N; Experiment Program written by J.C., and D.M.N; Data collection: J.C.; Data analysis: J.C. and D.M.N; Manuscript written and revised by J.C., Q.M.H, A.S., and D.M.N.

# 7 Supplemental Material

## 7.1 Baseline Mood

To ensure that our samples did not contain a disproportionately high number of depressed individuals, we first calculated the summary statistics of the baseline momentary mood rating collected prior to the task, and conducted two tailed t tests on whether the true mean of the ratings is different than average (ratings = 5). Additionally, we also conducted a not paired two sample t test on the rating difference between the two sample. We found that our participants had reported significantly higher than average momentary mood ratings and that there is no significant difference in starting mood between the two samples. The summary and t-test statistics are in the table below:

| Sample | N | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | t | D.F. | P |
|---|---|---|---|---|---|---|---|---|---|---|
| LT | 105 | 0.140 | 5.040 | 7.070 | 6.904 | 8.890 | 9.990 | 8.574 | 104 | $< 0.001$ |
| LTSF | 210 | 0.150 | 6.100 | 7.980 | 7.472 | 9.020 | 10.000 | 11.966 | 209 | $< 0.001$ |
| Difference | | | | | | | | 0.175 | 206 | 0.861 |

Table 3: Summary statistics and one tailed t test results for the Multi-Attempt Letter Task and the Single-Attempt Letter Task samples. The summary statistics include the minimum (min.), 1st quartile (1st Qu.), median, mean, 3rd quartile (3rd Qu.) and maxium (max.).

Therefore, we conclude that our sample do not contain disproportionately high number of depressed individuals and that our samples are equivalent in terms of baseline mood. Although, it does suggest that, in both of our sampels, participants on average have elevated mood to begin with.

## 7.2 Reaction Time Analysis

To affirm that difficulty manipulation does lead to higher cognitive effort expenditure, we analyzed the effect of difficulty on another metric that is linked to cognitive effort: reaction time. Because in the Multi-Attempt Letter Task, the participants were allowed to repeat a given trial, so their reaction times were not suitable for this analysis. We then only applied this analysis to the reaction time data from the Single-Attempt Letter Task.

We first log-transformed all the reaction time data and constructed a mixed effect linear regression model as following:

$$log(RT_{t,i}) = (\alpha_0 + \beta_{0,i}) + (\alpha_n + \beta_{n,i})n_{t,i} + (\alpha_\tau + \beta_{\tau,i})t + \epsilon_{t,i} \tag{4}$$

where $RT_{t,i}$ is the reaction time for a given participant ($i$) for a given trial ($t$). The rest of the notations are similar to our mixed linear effect model for momentary mood ratings, as illustrated in the Methods section.

We found that indeed increased number of letters ($n$) leads to higher reaction time (est. = 0.061, std. error = 0.006, df = 209, t = 10.157, p $<$ 0.001).

We conclude that indeed higher difficulty leads to higher reaction time, which indicates increased exertion of cognitive effort.

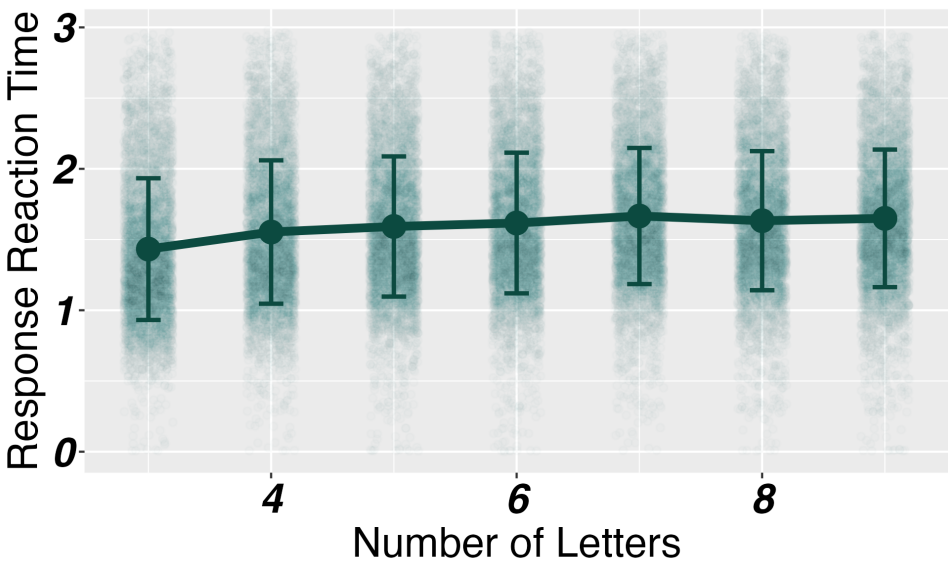


Figure 5: Reaction Time data from the Single-Attempt Letter Task.

## 7.3 Time-in-task related effects

To further investigate how participants' performance and mood changed over the course of task, we provided additional figures (Figure 6) to illustrate that the main effects we have reported is stable and do not change over the course of the task.

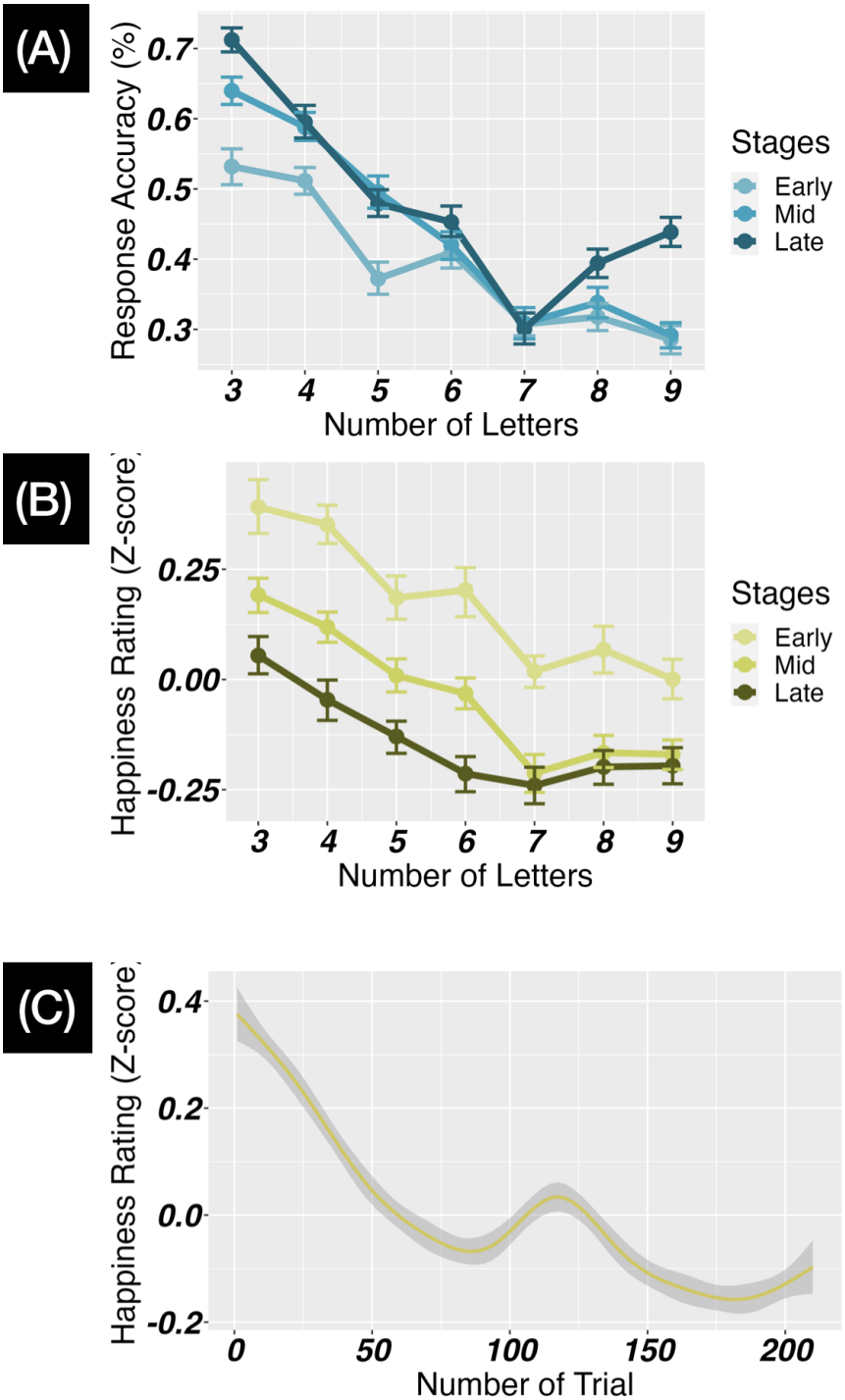


Figure 6: The effects of number of letters on accuracy and mood over the course of the task. Stage (early, mid late) indicates a factorized form of the time elapsed since the beginning of the task (evenly split into 3 sections: early/mid/late). A) Accuracy over the course of the task showed no significant change; B) Difficulty's impact on mood is stable over the course of the task; and C) Mood showed downward drifting over time.

# References


Aguinis, H., Villamor, I., & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, *47*(4), 823–837. https://doi.org/10.1177/0149206320969787

*Amazon mechanical turk*. (2023). https://www.mturk.com/

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-V)* (Fifth Edition). https://doi.org/10.1176/appi.books.9780890425596

Anderson, E. C., Carleton, R. N., Diefenbach, M., & Han, P. K. J. (2019). The relationship between uncertainty and affect. *Frontiers in Psychology*, *10*, 2504. https://doi.org/10.3389/fpsyg.2019.02504

Aw, J. M., Vasconcelos, M., & Kacelnik, A. (2011). How costs affect preferences: Experiments on state dependence, hedonic state and within-trial contrast in starlings. *Animal Behaviour*, *81*(6), 1117–1128. https://doi.org/10.1016/j.anbehav.2011.02.015

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

Batson, C. D. (1987). Prosocial Motivation: Is it ever Truly Altruistic? In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (pp. 65–122). Academic Press. https://doi.org/10.1016/S0065-2601(08)60412-8

Bennett, D., Radulescu, A., Zorowitz, S., Felso, V., & Niv, Y. (2021). *Affect-congruent attention drives changes in reward expectations* (tech. rep.). PsyArXiv. https://doi.org/10.31234/osf.io/vu2cw

Bijleveld, E. (2018). The feeling of effort during mental activity. *Consciousness and Cognition*, *63*, 218–227. https://doi.org/10.1016/j.concog.2018.05.013

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 16–27. https://doi.org/10.3758/CABN.9.1.16

Brinkmann, K., & Gendolla, G. H. E. (2008). Does depression interfere with effort mobilization? effects of dysphoria and task difficulty on cardiovascular response. *Journal of Personality and Social Psychology*, *94*(1), 146–157. https://doi.org/10.1037/0022-3514.94.1.146

Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., & Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, *6*(8), 1112–1125. https://doi.org/10.1038/s41562-022-01332-8

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

de Boer, L., Axelsson, J., Riklund, K., Nyberg, L., Dayan, P., Bäckman, L., & Guitart-Masip, M. (2017). Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age (W. Schultz, Ed.). *eLife*, *6*, e26424. https://doi.org/10.7554/eLife.26424

Diener, E., Lucas, R. E., & Scollon, C. N. (2009). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. In E. Diener (Ed.), *The science of well-being:*

*The collected works of ed diener* (pp. 103–118). Springer Netherlands. https://doi.org/10.1007/978-90-481-2350-6_5

Eldar, E., Roth, C., Dayan, P., & Dolan, R. J. (2018). Decodability of Reward Learning Signals Predicts Mood Fluctuations. *Current Biology*, *28*(9), 1433–1439.e7. https://doi.org/10.1016/j.cub.2018.03.038

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as representation of momentum. *Trends in Cognitive Sciences*, *20*(1), 15–24. https://doi.org/10.1016/j.tics.2015.07.010

Ellis, J. M., Midgette, E. P., Capiaghi, A. J., Schoemann, A. M., Sears, S. F., Kyle, B. N., Carels, R. A., & Whited, M. C. (2022). Mobile assessment of experiential avoidance, mood, stress, and attendance in cardiopulmonary rehabilitation. *Health Psychology*, *41*(12), 955–963. https://doi.org/10.1037/hea0001204

Emanuel, A., & Eldar, E. (2023). Emotions as computations. *Neuroscience & Biobehavioral Reviews*, *144*, 104977. https://doi.org/10.1016/j.neubiorev.2022.104977

Erber, R., & Tesser, A. (1992). Task effort and the regulation of mood: The absorption hypothesis. *Journal of Experimental Social Psychology*, *28*(4), 339–359. https://doi.org/10.1016/0022-1031(92)90050-T

Farrell, K., Lak, A., & Saleem, A. B. (2022). Midbrain dopamine neurons signal phasic and ramping reward prediction error during goal-directed navigation. *Cell Reports*, *41*(2), 111470. https://doi.org/10.1016/j.celrep.2022.111470

Ferstl, M., Teckentrup, V., Lin, W. M., Kräutlein, F., Kühnel, A., Klaus, J., Walter, M., & Kroemer, N. B. (2022). Non-invasive vagus nerve stimulation boosts mood recovery after effort exertion [Publisher: Cambridge University Press]. *Psychological Medicine*, *52*(14), 3029–3039. https://doi.org/10.1017/S0033291720005073

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

Ganesan, K. (2020, September 1). *Effort-related decision-making and its underlying processes during childhood* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/uxjq6

Gendolla, G. H. E., & Krüsken, J. (2002). The joint effect of informational mood impact and performance-contingent consequences on effort-related cardiovascular response. [Publisher: US: American Psychological Association]. *Journal of Personality and Social Psychology*, *83*(2), 271. https://doi.org/10.1037/0022-3514.83.2.271

Gillan, C. M., & Daw, N. D. (2016). Taking Psychiatry Research Online. *Neuron*, *91*(1), 19–23. https://doi.org/10.1016/j.neuron.2016.06.002

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Godman, M., Nagatsu, M., & Salmela, M. (2014). The Social Motivation Hypothesis for Prosocial Behavior. *Philosophy of the Social Sciences*, *44*(5), 563–587. https://doi.org/10.1177/0048393114530841

Grèzes, J., Erblang, M., Vilarem, E., Quiquempoix, M., Van Beers, P., Guillard, M., Sauvet, F., Mennella, R., & Rabat, A. (2021). Impact of total sleep deprivation

and related mood changes on approach-avoidance decisions to threat-related facial displays. *Sleep*, *44*(12), zsab186. https://doi.org/10.1093/sleep/zsab186

Grosscup, S. J., & Lewinsohn, P. M. (1980). Unpleasant and pleasant events, and mood. *Journal of Clinical Psychology*, *36*(1), 252–259. https://doi.org/10.1002/1097-4679(198001)36:1⟨252::AID-JCLP2270360131⟩3.0.CO;2-S

Het, S., & Wolf, O. T. (2007). Mood changes in response to psychosocial stress in healthy young women: Effects of pretreatment with cortisol [Place: US Publisher: American Psychological Association]. *Behavioral Neuroscience*, *121*, 11–20. https://doi.org/10.1037/0735-7044.121.1.11

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*(7912), 129–136. https://doi.org/10.1038/s41586-022-04743-9

Huff, C., & Tingley, D. (2015). "who are these people?" evaluating the demographic characteristics and political preferences of MTurk survey respondents [Publisher: SAGE Publications Ltd]. *Research & Politics*, *2*(3), 2053168015604648. https://doi.org/10.1177/2053168015604648

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLOS Computational Biology*, *8*(3), e1002410. https://doi.org/10.1371/journal.pcbi.1002410

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences*, *22*(4), 337–349. https://doi.org/10.1016/j.tics.2018.01.007

Jangraw, D. C., Keren, H., Sun, H., Bedder, R. L., Rutledge, R. B., Pereira, F., Thomas, A. G., Pine, D. S., Zheng, C., Nielson, D. M., & Stringaris, A. (2023). A highly replicable decline in mood during rest and simple tasks. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-023-01519-7

Jansma, J., Ramsey, N., de Zwart, J., van Gelderen, P., & Duyn, J. (2006). fMRI study of effort and information processing in a working memory task. *Human Brain Mapping*, *28*(5), 431–440. https://doi.org/10.1002/hbm.20297

Joana, D. B. D. L. R. (2009). Mood and mental effort : Informational mood impact on cardiovascular reactivity and the context-dependency of moods [Publisher: Université de Genève]. https://doi.org/10.13097/ARCHIVE-OUVERTE/UNIGE:16908

Keren, H., Zheng, C., Jangraw, D. C., Chang, K., Vitale, A., Rutledge, R. B., Pereira, F., Nielson, D. M., & Stringaris, A. (2021). The temporal representation of experience in subjective mood (J. Roiser & T. E. Behrens, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, *10*, e62051. https://doi.org/10.7554/eLife.62051

Kool, W., Cushman, F. A., & Gershman, S. J. (2016). When Does Model-Based Control Pay Off? *PLOS Computational Biology*, *12*(8), e1005090. https://doi.org/10.1371/journal.pcbi.1005090

Kumar, P., Goer, F., Murray, L., Dillon, D. G., Beltzer, M. L., Cohen, A. L., Brooks, N. H., & Pizzagalli, D. A. (2018). Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology*, *43*(7), 1581–1588. https://doi.org/10.1038/s41386-018-0032-x

Kurniawan, I. T., Guitart-Masip, M., Dayan, P., & Dolan, R. J. (2013). Effort and valuation in the brain: The effects of anticipation and execution. *Journal of Neu-*

*roscience*, *33*(14), 6160–6169. https://doi.org/10.1523/JNEUROSCI.4777-12.2013

Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? [Publisher: SAGE Publications Inc]. *Evolutionary Psychology*, *8*(2), 14747049100080020 https://doi.org/10.1177/147470491000800208

Larsen, R. J., Diener, E., & Emmons, R. A. (1986). Affect intensity and reactions to daily life events. *Journal of Personality and Social Psychology*, *51*(4), 803–814. https://doi.org/10.1037/0022-3514.51.4.803

Larsen, S. E., & Berenbaum, H. (2014). Substantial symptom changes in naturalistic recovery from aversive events. *Journal of Clinical Psychology*, *70*(10), 967–978. https://doi.org/10.1002/jclp.22092

Liuzzi, L., Chang, K. K., Zheng, C., Keren, H., Saha, D., Nielson, D. M., & Stringaris, A. (2021). Magnetoencephalographic correlates of mood and reward dynamics in human adolescents. *Cerebral Cortex*, bhab417. https://doi.org/10.1093/cercor/bhab417

Lockwood, P. L., Hamonet, M., Zhang, S. H., Ratnavel, A., Salmony, F. U., Husain, M., & Apps, M. A. J. (2017). Prosocial apathy for helping others when effort is required — Nature Human Behaviour. *Nature Human Behaviour*, *1*(7), 0131. https://doi.org/10.1038/s41562-017-0131

Madsen, P. L., Hasselbalch, S. G., Hagemann, L. P., Olsen, K. S., Bülow, J., Holm, S., Wildschiødtz, G., Paulson, O. B., & Lassen, N. A. (1995). Persistent resetting of the cerebral oxygen/glucose uptake ratio by brain activation: Evidence obtained with the kety—schmidt technique [Publisher: SAGE Publications Ltd STM]. *Journal of Cerebral Blood Flow & Metabolism*, *15*(3), 485–491. https://doi.org/10.1038/jcbfm.1995.60

Massar, S. A. A., Csathó, Á., & Van der Linden, D. (2018). Quantifying the motivational effects of cognitive fatigue through effort-based decision making [Publisher: Frontiers]. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00843

McCredie, M. N., & Morey, L. C. (2019). Who are the turkers? a characterization of MTurk workers using the personality assessment inventory [Publisher: SAGE Publications Inc]. *Assessment*, *26*(5), 759–766. https://doi.org/10.1177/1073191118760709

Mintz, N. L. (2010). Effects of esthetic surroundings: II. prolonged and repeated experience in a “beautiful” and an “ugly” room [Publisher: Taylor & Francis Group]. *The Journal of Psychology*. https://www.tandfonline.com/doi/abs/10.1080/00223980.1956.9713018

Mortensen, K., & Hughes, T. L. (2018). Comparing amazon’s mechanical turk platform to conventional data collection methods in the health and medical research literature. *Journal of General Internal Medicine*, *33*(4), 533–538. https://doi.org/10.1007/s11606-017-4246-0

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. https://doi.org/10.1007/s00213-006-0502-4

O’Callaghan, G., & Stringaris, A. (2019). Reward processing in adolescent depression across neuroimaging modalities. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *47*(6), 535–541. https://doi.org/10.1024/1422-4917/a000663

Packheiser, J., Donoso, J. R., Cheng, S., Güntürkün, O., & Pusch, R. (2021). Trial-by-trial dynamics of reward prediction error-associated signals during extinction learning and renewal. *Progress in Neurobiology*, *197*, 101901. https://doi.org/10.1016/j.pneurobio.2020.101901

Philip, B. (1971). Hedonic relativism and planning the good society. *Adaptation level theory*, 287–301.

*Prolific*. (2023). https://www.prolific.co/

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Redmiles, E. M., Kross, S., & Mazurek, M. L. (2019). How well do my results generalize? comparing security and privacy survey results from MTurk, web, and telephone samples [ISSN: 2375-1207]. *2019 IEEE Symposium on Security and Privacy (SP)*, 1326–1343. https://doi.org/10.1109/SP.2019.00014

Robinson, M. M., & Morsella, E. (2014). The subjective effort of everyday mental tasks: Attending, assessing, and choosing. *Motivation and Emotion*, *38*(6), 832–843. https://doi.org/10.1007/s11031-014-9441-2

Robles, C. F., & Johnson, A. W. (2017). Disruptions in effort-based decision-making and consummatory behavior following antagonism of the dopamine d2 receptor. *Behavioural Brain Research*, *320*, 431–439. https://doi.org/10.1016/j.bbr.2016.10.043

Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory (T. Kahnt & C. Büchel, Eds.). *eLife*, *10*, e61077. https://doi.org/10.7554/eLife.61077

Russell, S., & Wefald, E. (1991). Principles of metareasoning. *Artificial Intelligence*, *49*(1), 361–395. https://doi.org/10.1016/0004-3702(91)90015-C

Rutledge, R. B., Moutoussis, M., Smittenaar, P., Zeidman, P., Taylor, T., Hrynkiewicz, L., Lam, J., Skandali, N., Siegel, J. Z., Ousdal, O. T., Prabhu, G., Dayan, P., Fonagy, P., & Dolan, R. J. (2017). Association of neural and emotional impacts of reward prediction errors with major depression. *JAMA Psychiatry*, *74*(8), 790. https://doi.org/10.1001/jamapsychiatry.2017.1713

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being [Publisher: National Academy of Sciences Section: Biological Sciences]. *Proceedings of the National Academy of Sciences*, *111*(33), 12252–12257. https://doi.org/10.1073/pnas.1407535111

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2015). Dopaminergic modulation of decision making and subjective well-being [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, *35*(27), 9811–9822. https://doi.org/10.1523/JNEUROSCI.0702-15.2015

Sandra, D. A., & Otto, A. R. (2018). Cognitive capacity limitations and need for cognition differentially predict reward-induced cognitive effort expenditure. *Cognition*, *172*, 101–106. https://doi.org/10.1016/j.cognition.2017.12.004

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using mechanical turk to study clinical populations [Publisher: SAGE Publications Inc]. *Clinical Psychological Science*, *1*(2), 213–220. https://doi.org/10.1177/2167702612469015

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Sonnentag, S., & Grant, A. M. (2012). Doing good at work feels good at home, but not right away: When and why perceived prosocial impact predicts positive affect. *Personnel Psychology*, *65*(3), 495–530. https://doi.org/10.1111/j.1744-6570.2012.01251.x

Stone, A. A., & Neale, J. M. (1984). Effects of severe daily events on mood. *Journal of Personality and Social Psychology*, *46*(1), 137–144. https://doi.org/10.1037/0022-3514.46.1.137

Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: A translational model of motivational anhedonia. *Journal of Abnormal Psychology*, *121*(3), 553–558. https://doi.org/10.1037/a0028813

Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the 'EEfRT'? the effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS ONE*, *4*(8). https://doi.org/10.1371/journal.pone.0006598

Vaish, A., Carpenter, M., & Tomasello, M. (2016). The Early Emergence of Guilt-Motivated Prosocial Behavior. *Child Development*, *87*(6), 1772–1782. https://doi.org/10.1111/cdev.12628

Vogel, T. A., Savelson, Z. M., Otto, A. R., & Roy, M. (2020). Forced choices reveal a trade-off between cognitive effort and physical pain (C. Büchel, J. Roiser, M. Inzlicht, & A. Westbrook, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, *9*, e59410. https://doi.org/10.7554/eLife.59410

Wang, L., Zheng, J., & Meng, L. (2017). Effort provides its own reward: Endeavors reinforce subjective expectation and evaluation of task performance. *Experimental Brain Research*, *235*(4), 1107–1118. https://doi.org/10.1007/s00221-017-4873-z

Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? load, trait, and aging effects revealed by economic preference. *PloS One*, *8*(7), e68210. https://doi.org/10.1371/journal.pone.0068210

Yang, X.-h., Huang, J., Zhu, C.-y., Wang, Y.-f., Cheung, E. F., Chan, R. C., & Xie, G.-r. (2014). Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry Research*, *220*(3), 874–882. https://doi.org/10.1016/j.psychres.2014.08.056