Effort Foraging Task reveals positive correlation between individual differences in the cost of cognitive and physical effort in humans and relationship to self-reported motivation and affect

Laura A. Bustamante[1] Temitope Oshinowo[1] Jeremy Lee[1] Elizabeth Tong[1]
Allison R. Burton[1] Amitai S. Shenhav [2] Jonathan D. Cohen [1]
Nathaniel D. Daw [1]

[1]Princeton Neuroscience Institute
[2]Brown University
laura.ana.bustamante@gmail.com

Abstract

Effort-based decisions, in which people weigh potential future rewards against effort costs required to achieve those rewards, have largely been studied separately for cognitive or physical effort, yet most real world actions incur both cognitive and physical effort costs. What is the relationship between cognitive and physical effort costs? Here we attempt to formalize the mechanisms underlying effort-based decisions and address methodological challenges to isolate and measure the factors contributing to such decisions (including sensitivity to reward and effort costs).

Patch foraging is an ecologically valid reward rate maximization problem with well developed theoretical tools to understand choices. We developed the Effort Foraging Task to isolate and quantify the cost of both cognitive and physical effort using a computational model. We embedded cognitive or physical effort costs into a patch foraging sequential decision task. Participants chose between harvesting a depleting patch, or traveling to a new patch that was costly in time and effort. Participants' exit thresholds (reflecting the reward they expected to receive by harvesting when they chose to travel to a new patch) were sensitive to cognitive and physical effort demands, allowing us to quantify the perceived effort cost in monetary terms. Individual differences in cognitive and physical effort costs were positively correlated, suggesting that these are perceived and processed in common terms across different domains. We found patterns of correlation of both cognitive and physical effort costs with self-reported cognitive function, anhedonia, depression, anxiety, and fatigue. This suggests that our task captures decision mechanisms closely associated with real-world motivation, and can be used to study individual variation in effort-based decisions across different domains of cost.

1 Introduction

People make *effort-based decisions* every day, weighing the potential rewards associated with an action against the effort it requires. Economic utility theories guide researchers' understanding of these choices: people seek to maximize reward while minimizing effort, which can be accomplished by computing an 'expected value' of effort (Shenhav et al., 2013; Shenhav et al., 2017). In these theories effort is described as costly, since it reduces the value of rewards. Evidence from cognitive psychology and neuroscience consistently shows that people factor effort into their decisions and that individuals approach tradeoffs between rewards and effort differently (Chong et al., 2016; Salamone et al., 2018; Treadway et al., 2012; Walton et al., 2007). Individuals choose to avoid environments with a higher demand for task switches (Kool et al., 2010) and are willing to forgo rewards to avoid tasks that impose a higher working memory load (Westbrook et al., 2013). Other studies have shown that participants can increase the amount of cognitive effort they exert with increased incentives, but generally hold back from doing so (Botvinick & Braver, 2015; Padmala & Pessoa, 2011).

Achieving many goals depends on the exertion of both cognitive and physical effort. For example you might need to plan a route to a reward (cognitive effort) and then walk there (physical effort). Research tends to treat cognitive and physical effort costs separately. However, real-world actions involve multiple types of costs including the cognitive and physical effort, time costs, and factors such as risk and uncertainty. Furthermore, these costs interact – for example, exerting greater physical effort (e.g., vigor) can reduce time costs, and exerting greater cognitive effort (i.e., planning an efficient action) can reduce both physical effort and time costs. The reinforcement learning literature has been concerned with the extent to which there is a common representation of 'value' that integrates different kinds of rewards (e.g., food, money Chib et al., 2009; Levy & Glimcher, 2011, 2012). To what extent is there also a common representation of 'cost' that integrates different forms of effort and costs associated with an activity (i.e. cognitive effort costs, physical effort costs, time costs Schmidt et al., 2012)? Human and animal research suggests that cognitive and physical effort decisions are controlled by shared neural populations (Borderies et al., 2020; Bornert & Bouret, 2021; Chong et al., 2017; Lopez-Gamundi et al., 2021; Schmidt et al., 2012). Cognitive and physical fatigue seem to interact (Giboin & Wolff, 2019; Marcora et al., 2009) and intermixing choices about cognitive and physical effort affects choices for both effort types (Toro-Serey et al., 2021).

One window into the relationship of different domains of cost is to investigate individual differences. Do individuals who avoid cognitive effort more also avoid physical effort more? Research is limited about the relationship between individual differences in cognitive and physical effort costs. Lopez-Gamundi & Wardle (2018) found a positive relationship (correlation = 0.43, df = 60, $p<0.001$) between the percent of hard task choices in the cognitive (task-switching task) and physical (rapid key-pressing task) versions of an effort based decision making task (this finding was replicated in Tran et al., 2020, correlation = 0.35, df = 79, $p<0.05$). In the present study we quantified individual differences in cognitive and physical effort costs and examined the relationship between them. To do so, we developed a novel Effort Foraging Task with two variants - cognitive and physical - and then fit a computational model to individual participants' behavior in order to estimate the costs associated with each form of effort. Using this model we evaluated the correlation between the estimated cognitive and physical effort costs. This task was designed to leverage the strengths of the patch foraging paradigm, an ecologically valid sequential decision and reward rate maximization problem with well developed theoretical tools for estimating costs. By examining the relationship between individual differences in these costs, we can learn more about the overlap between the computation of cognitive and physical effort. If individual differences in effort costs for these domains are unrelated, this would suggest that costs are separately computed. Conversely, if there is a strong relationship between them, this would suggest that cognitive and physical costs share some common sources. Furthermore, to assess the extent to which these cost measures are related to real-world behavior, we also used a data driven approach — canonical correlation analysis — to look for relationships between task behavior and self-report surveys of motivation and affect (including measures of apathy, anhedonia, depression, and others).

1.1 The Effort Foraging Task.

Valid and reliable measures of effort costs are needed to test theories concerning how individual differences in effort costs affect behavior and everyday effort exertion, and that can be used to help identify and characterize their neural underpinnings in future research. Here we introduce the Effort Foraging Task, an indirect, experiential, foraging-style task, designed to improve effort cost estimation and predictive validity compared to previous methods. Many previously reported tasks involved explicit effort-based decision making; for example, participants are directly if they would rather complete a high or low effort task (Kool et al., 2010; Lopez-Gamundi & Wardle, 2018; Westbrook et al., 2013). Such explicit decision making may be subject to secondary demand characteristics, in which participants behave according to what they think the experimenter wants, or with their self-image; decisions explicitly trading off numeric quantities are also susceptible to idiosyncratic arithmetic heuristics (Marzilli Ericson et al., 2015). In contrast, the influence of effort in the the Effort Foraging Task is measured indirectly; participants learn about the environment reward rate and effort costs through experience, and their choices reflect their ongoing evaluation of these quantities. In addition, some existing tasks concern hypothetical choices, or choices separated in time from a later the realization of cognitive effort (Lopez-Gamundi & Wardle, 2018; Westbrook et al., 2013). In the Effort Foraging Task participants directly experience the effort immediately after each choice to travel. Furthermore, foraging tasks are sequential decision tasks that are ecologically valid, and have proven to be valuable in understanding decision-making in formally rigorous terms, and relating it to underlying neural mechanisms, across a variety of species, including rodents(Carter & Redish, 2016; Kane et al., 2021), non-human primates (Hayden, 2018; Hayden et al., 2011) and humans (Constantino et al., 2017; Constantino & Daw, 2015; Kolling et al., 2012; Lenow et al., 2017). The Effort Foraging Task adapts a version of the paradigm developed by Constantino & Daw 2015, by embedding cognitive and physical effort costs into a patch foraging environment so as to quantify how these are traded off against monetary rewards.

In a standard patch foraging task the forager visits a 'patch' which can be harvested to yield rewards (here, a simulated orchard with apple trees). Rewards in a patch (apples) deplete over time spent harvesting, such that reward rate within the patch decreases over time. At any point the forager can travel to a new patch, which has replenished rewards, but it takes time to travel there. Deciding when to leave a patch in a foraging environment involves tradeoffs between harvesting rewards available from the current patch, and the time (and/or effort) spent traveling to a different (but richer) one. For this reason, the level at which the forager decides to exit the current patch (i.e., their 'exit threshold' or the number of apples they have last received before quitting) reflects the reward they are willing to forgo by leaving that patch and spending the time (and/or effort) to travel to another. In these respects, the exit threshold reveals the point of equivalence in the tradeoff between the cost of harvesting with diminishing rewards and the time (and/or effort) cost of traveling to a new patch.

These considerations are formalized in the Marginal Value Theorem (Charnov, 1976), which asserts that a simple threshold policy maximizes reward rate. The forager simply has to maintain an estimate of the average reward rate in the environment, and exit a patch when the instantaneous reward rate falls below the long run average. Constantino & Daw 2015 found that human participants playing a virtual foraging game used a threshold exit strategy consistent with the Marginal Value Theorem, which explained behavior better than other reinforcement learning models (e.g., Temporal Difference learning). Furthermore, they found that exit thresholds shifted reliably and in predicted ways when the environment changed (e.g., when travel time and/or reward depletion was experimentally manipulated). For example, when the travel time between patches was increased, participants' exit thresholds decreased, reflecting the increased opportunity cost of travel and an overall decrease in average reward rate.

For the Effort Foraging Task, we followed the design of 2015, but rather than manipulating travel time, instead we varied the *effort* – cognitive or physical – required to travel between patches, and compared exit thresholds in high versus low effort conditions. We predicted that contexts with higher effort costs would decrease participants' estimates of average reward rate, leading the exit threshold to be lower; that is, a greater willingness to accept diminishing rewards to avoid effortful travel. Accordingly, we used the difference in exit threshold between high and low effort conditions to infer the perceived costs of travel. More specifically, we used participants' decision thresholds to create a computational model based on the Marginal Value Theorem to quantify the added cost of high compared to low effort conditions in this task. Using this model, we found that most participants avoided the high effort tasks, treating them as costly. We used these cost estimates to directly fit the correlation between individual differences in cognitive and physical effort costs in the same currency (money) and found a moderate positive correlation. In developing the task, we conducted validation experiments and results confirmed that participants' thresholds were sensitive to standard foraging manipulations (i.e., patch resource richness) and also to two domains of cognitive effort (interference and working memory tasks).

We developed two effort variants of the Effort Foraging Task (Fig. 1, Experiment 1). Effort level was manipulated block-wise (4 minute long blocks). The cognitive effort variant required performing trials of the Multi-Source Interference Task (MSIT, Fig. 2 left panel, Bush & Shin, 2006). The high effort condition required completing interference trials (demanding more cognitive effort), and the low effort condition required completing congruent trials (demanding relatively less cognitive effort). The physical variant of the task required participants to rapidly press a key to reach a new patch (Fig. 2 right panel, based on previous research demonstrating that rapid key-pressing is considered to be physically effortful and costly Treadway et al., 2009). The high effort condition required participants to press the key the maximum number of times they could in the time allotted (individually-determined in a preceding calibration phase), and the low effort condition required half that number of key-presses. Travel time (i.e., time to complete the MSIT or key-pressing tasks) was fixed and the same across both variants of the task and the high and low effort conditions of each. We predicted participants would have a lower patch leaving threshold (in units of apples) in the high effort conditions compared to the low effort conditions, since travel (effort) costs were greater in the former. The measure of effort cost for an individual was the differential travel cost of the more effortful condition (incongruent MSIT, or Larger Number of Presses) compared to the less effortful condition (congruent MSIT, or Smaller Number of Presses).

2 Results

The primary dependent variables in our analyses were exit thresholds, which reflect the point when the cost of leaving just offsets the benefits of reaching a replenished patch (which is increasing as the current

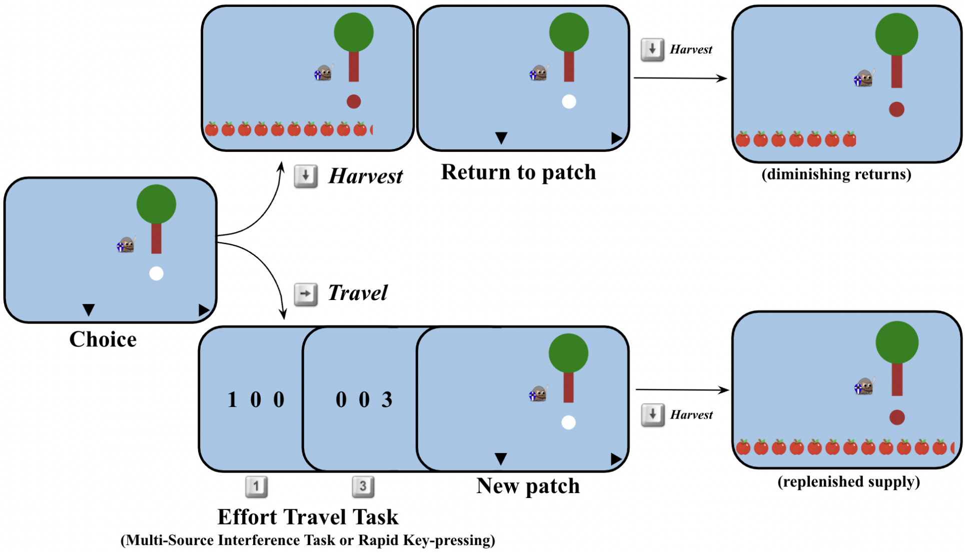

Figure 1: Foraging trial diagram. On each trial participants chose to harvest the tree they were at (down arrow key) or travel to a new tree (right arrow key), during the travel they completed an effortful task, after which they arrived at a new patch with a replenished supply of apples.

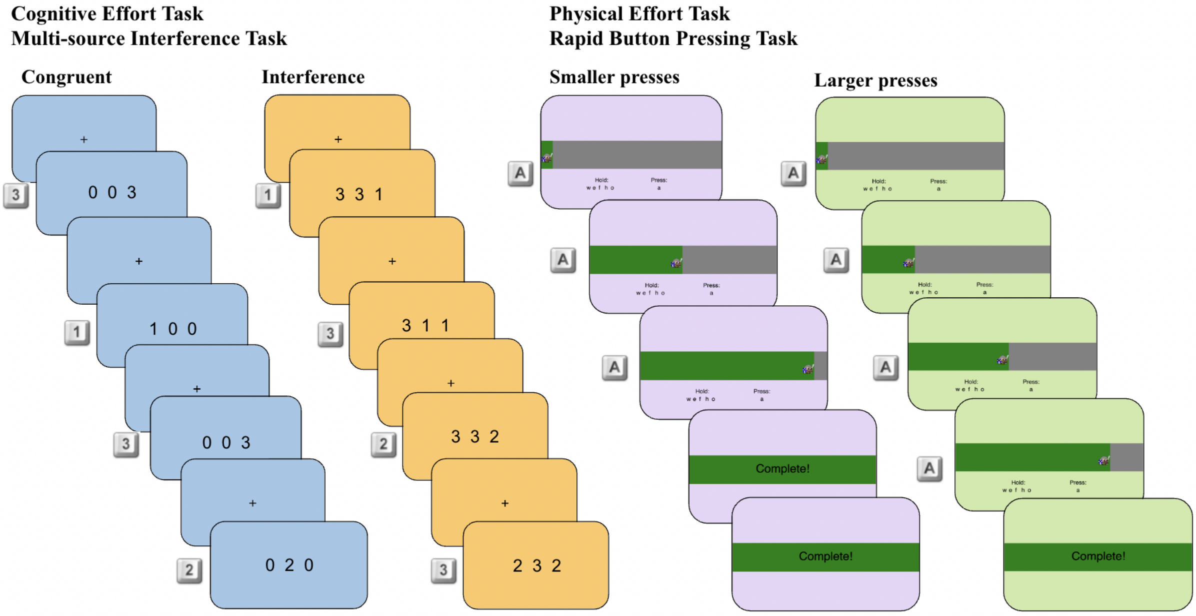

Figure 2: Effort travel tasks. Left panel: Cognitive Effort, Multi-source Interference Task. Participants identified which number was the oddball in a list of three numbers. The background color differed for the high effort (interference trials, orange) and low effort (congruent trials, blue) conditions. They responded with the '1', '2', '3' keys. The correct response for each example screen is displayed on the left of that example screen. Right panel: Physical Effort, Rapid Key-pressing Task. Participants rapidly pressed the 'a' key while holding down the 'w', 'e', 'f', 'h' and 'o' keys. Pressing the 'a' key moved the avatar rightwards and filled up the grey horizontal bar with green. When participants reached the goal number of presses 'Complete!' appeared in the horizontal bar and participants waited for the remainder of the travel time. The background color differed for the high effort (smaller presses, purple)and low effort (larger presses, green) conditions.

patch yields progressively less). We assumed that participants set their exit thresholds to maximize the rate of rewards minus costs per time step. This is given by the Marginal Value Theorem:

$$\rho = \frac{\sum r - \sum c}{T} \tag{1}$$

According to Eq. (1), the exit threshold (ρ, the reward the forager expects if they continue harvesting a patch at the time they exit), is equal to the long run average reward rate, which includes the sum of all rewards ($\sum r$) minus sum of (non-time) costs incurred in the environment ($\sum c$, i.e., energy spent extracting rewards or traveling to the next patch), divided by the total number of harvest periods (the time cost normalized by the harvest time, T = total time / harvest time). We implemented a computational model based on the Marginal Value Theorem and fit it to participants' exit thresholds to quantify the relative increase in travel cost between the high and low effort conditions for each effort type (see model methods in Section 4.4.1).

Summary of results. We found that differences in foraging decisions (viz., exit thresholds) is a useful indirect measure of motivation to exert both cognitive and physical effort. Foraging decisions were sensitive to task difficulty in both of the variants tested, in which travel costs were implemented as cognitive or physical effort. Consistent with our prediction, average exit threshold was lower in the higher travel effort than the lower travel effort conditions (Fig. 3). Participants (Experiment 1, N = 537) opted to stay longer in a patch, accepting diminishing rewards, in the high travel effort conditions to avoid the increased cost of travel. Validation experiments demonstrated that participants' exit thresholds were responsive to all of our manipulations; two types of cognitive effort (interference control in Experiment 1 (MSIT), working memory in Experiment 2 (N-Back)), environment richness (Experiment 3). Fits of the Marginal Value Theorem model to trial by trial behavior (Experiment 1) further confirmed the presence of high effort costs, and revealed an interesting mixture of effort-avoiding and effort-seeking participants. We found that cognitive and physical effort costs were moderately positively correlated, and that both effort costs had patterns of correlation to self-report measures related to motivation and affect.

2.1 Change in exit threshold by effort condition.

As a model-agnostic metric of high effort cost, we used the change in exit threshold from low to high effort conditions. For each participant we computed the average exit threshold per condition, and the difference between them (high effort - low effort mean threshold). We expected this value to be negative, reflecting effort-avoidance. If threshold increased for a participant, this suggested effort-seeking. Across participants, we found a mix of effort avoidance, effort seeking, and indifference to effort (values close to zero) (Fig. 3 right panel). We computed the group average change in threshold (Fig. 3 left panel), and used linear mixed-effects regression to test whether change in exit thresholds significantly differed from zero. As predicted, on average participants exited trees later in the high relative to low effort conditions (mixed-effects regression: interference - congruent MSIT, $\beta_{cognitive} = -0.236$ apples, df = 460.071, F = 50.062, $p < 0.001$, Larger - Smaller Number of Presses $\beta_{physical} = -0.379$ apples, df = 474.041, F = 87.326, $p < 0.001$). Next, we used the foraging behavior to formally quantify the additional cost of the high effort tasks using a model based on the Marginal Value Theorem.

2.2 Hierarchical Bayesian model to estimate effort costs for an individual.

We fit a hierarchical Bayesian logistic model based on the Marginal Value Theorem (Charnov, 1976), which predicted harvest versus exit decisions by comparing expected reward on the next harvest against the average reward rate (see model methods in Section 4.4.1). We defined the reward rate in terms of known reward rate values of the foraging environment per effort condition per participant (apples earned, time cost incurred, number of patches visited), as well as the unknown reward rate value (the cost of travel). The cost of travel in high effort blocks was expressed as the marginal increase in cost of travel from low to high effort. Defining this cost as a difference measure controls for any additional biases individual participants may have (such as differences in the subjective value of the reward) which are common to both conditions. The dependent individual differences measures in this task were the inferred cognitive and physical effort cost parameters. The other model parameters were the travel costs in the cognitive and physical low effort conditions, and the inverse temperature applied to the softmax function.

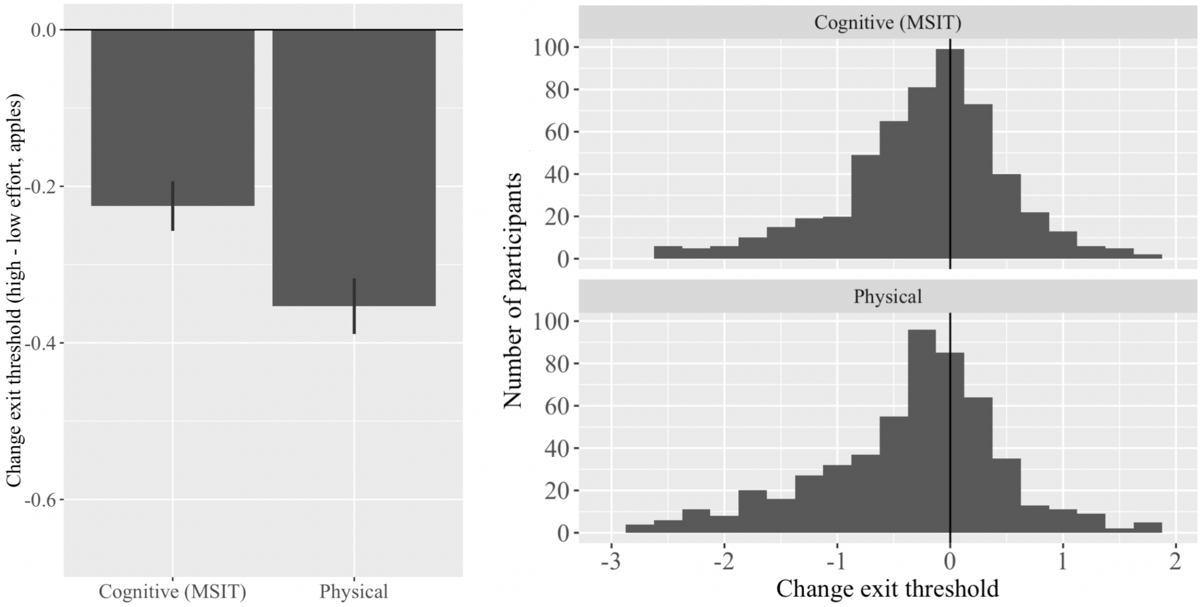

Figure 3: Change in exit thresholds by effort condition. Left panel: y-axis: Group-level mean change in exit threshold for cognitive and physical effort. x-axis: effort type. As predicted, on average participants exhibited lower exit thresholds in the high relative to low effort conditions. Error bars indicate standard error of the mean. Right panel: Individual variation in change in exit threshold. Top row: Histogram of participants mean change in exit threshold for cognitive high effort relative to cognitive low effort. Bottom row: mean change in exit threshold for physical high effort relative to physical low effort. Most participants were effort-avoiding (negative change in threshold), whereas some participants showed indifference to effort condition (near zero) or were effort-seeking (positive change in threshold).

Parameter	Mean	Lower CI bound	Upper CI bound
Cognitive Effort Cost	7.547	5.508	9.682
Cognitive Low Effort Travel Cost	94.646	90.706	98.761
Physical Effort Cost	13.527	11.045	16.061
Physical Low Effort Travel Cost	99.781	95.615	104.182
Inverse Temperature	0.260	0.238	0.282

Table 1: Group level posterior distributions. The group level average high cognitive effort cost was 7.5 apples. The group level average high physical effort cost was 13.5 apples. Table includes the mean of the group-level posterior distribution and the credible interval (95% HDI).

Consistent with the model-agnostic change in threshold metric, the group-level posterior parameter fit indicated that the high effort task was costly (Table 1). There was a range of individual differences (Fig. 4), cost was positive for most participants, some participants were indifferent to the effort manipulation (cost near zero), and some participants had a negative cost, suggesting that effort was valued.

Relationship between cognitive and physical effort costs. We computed the correlation between the estimated cognitive and physical high effort costs (again, each estimated as reflecting the additional cost of high effort relative to the low baseline), and found a moderate positive relationship (mean correlation = 0.566, 95% HDI = 0.355 - 0.766, Fig. S.6). This suggests a potential common representation for costs of different types used in effort-based decision-making.

2.3 Cognitive task performance predicts cognitive effort cost.

We tested whether cognitive task performance contributes to the cognitive effort cost measured by foraging choices (see relationship between costs and performance in Fig. 6). We regressed the difference in (log transformed) MSIT error rate and (log transformed) reaction time onto cognitive effort costs. We found that the difference in error rate significantly predicted cognitive effort cost (estimate = 15.313,

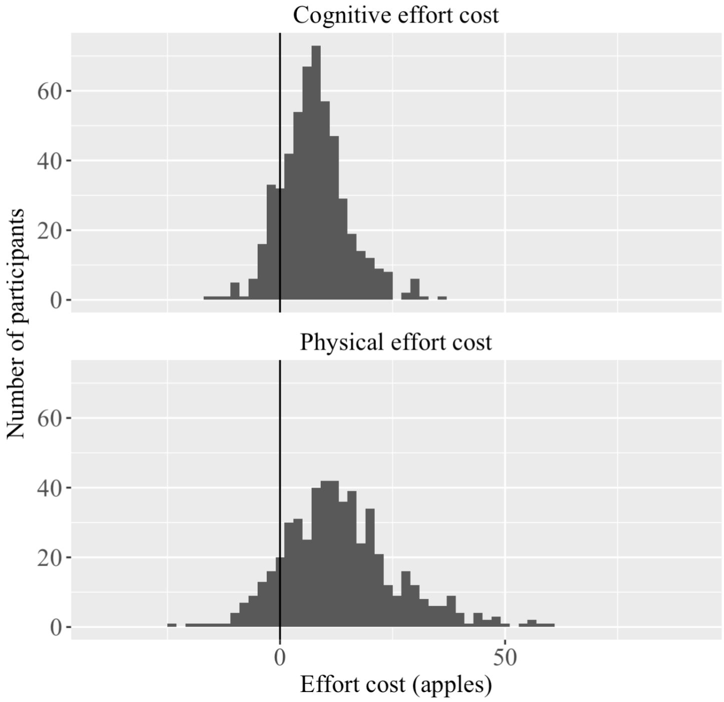

Figure 4: Individual differences in effort costs. There were individual differences in the high effort travel costs (expressed as the additional cost of the high relative to the low effort condition). Paralleling the pattern of exit thresholds most participants experienced the high effort conditions as effortful (positive cost), whereas some participants were insensitive to the effort manipulation (cost near zero) and others were effort seeking (negative cost).

SE = 3.684, t = 4.156, $p < 0.001$), indicating that participants with higher costs performed worse on the MSIT. However the reaction time interference effect did not predict cognitive effort cost (estimate = 4.861, SE = 3.802, t = 1.278, p = 0.202). In this regression the intercept was not significantly different from zero, suggesting that the effort cost measured are performance related (estimate = 4.861, SE = 3.802, t = 1.278, p = 0.202, compared to an intercept-only model estimate = 7.548, SE = 0.323, t = 23.38, $p < 0.001$). We see the same qualitative result using robust regression. This finding suggests participants may adaptively calibrate their effort costs according to their error rates.

We also included individual differences in MSIT error rates in our test of relationships between task performance and self-report surveys (see MSIT performance by participant in Fig. S.1). We did this to compare the predictive validity of the cognitive task performance measures (error rates and reactions times), versus the cognitive effort cost measure.

2.4 Relationship to self-reported motivation and affect.

Next we wished to examine the relationships between our task measures and self-report surveys, both to validate the external validity of our measures and to investigate the broader context associated with effort sensitivity. A practical issue with correlation or regression measures is the high dimensionality of both the survey and task measures. In previous work, we and others have used dimensionality reduction techniques such as factor analysis to summarize key dimensions of survey data prior to regressing them on individual task measures. (Gillan et al., 2016). Here we take that approach a step further by using canonical correlation analysis (CCA), a dimensionality reduction technique that simultaneously performs dimension reduction on both domains of data, so as to identify summaries of each domain that maximally relate to one another (here, the relationship between surveys and task behavior measures Wang et al., 2020).

We performed CCA to test whether task behavior reliably predicts self-reported motivation and affect (with data from Experiment 1, detailed in Methods Section 4.4.2, N = 430 participants who completed the survey and passed attention checks). The dependent variables were all of the self-report sub-scales (see Table 4). The predictor variables were all of the measures of task behavior of interest:

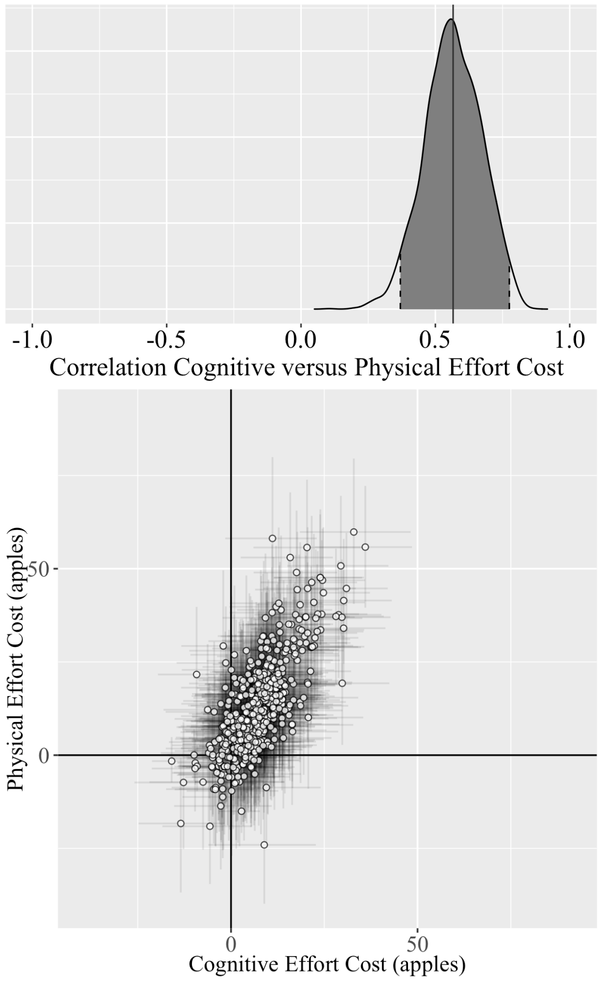

Figure 5: Correlation between cognitive and physical effort. Top panel: posterior distribution of correlation between high effort cost for cognitive and physical effort. Cognitive and physical effort costs are positively correlated (correlation = 0.566, 95% HDI = 0.355 - 0.766). Bottom panel: x-axis: Individual differences in cognitive effort costs, y-axis: Individual differences in physical effort costs. Error bars indicate 80% HDI.

(log transformed) error rate on congruent and incongruent trials, cognitive and physical effort costs, and overall threshold. Including all of the task behavior measures allowed us to explore the unique contribution of (potentially mutually correlated) effort costs, cognitive task performance, and subjective utility to proxies of real-world behavior by testing whether these variables fall on overlapping or separate dimensions. This approach has the benefit of increasing sensitivity (by making use of all of the measures simultaneously), while reducing the risks of multiple comparisons (by treating all of these factors in a single omnibus analysis).

The first question addressed by CCA is the dimensionality of the subspace relating the domains. Previous work (Moutoussis et al., 2021) found many self report measures related to many task measures along only a single "decision acuity" dimension, raising concerns that the correlations were dominated by nonspecific motivational or attentional variation. In the current dataset, CCA revealed three significant dimensions (summarized in , and full result shown in Fig. S.2, Wilks' Lamda (Wilks, 1935), using F-approximation, dimension 1 to 5: stat = 0.797, F-approx = 2.189, df1 = 70, df2 = 3141.555, $p < 0.001$, dimension 2 to 5: stat = 0.877, F-approx = 1.691, df1 = 52, df2 = 2558.280, $p < 0.0016$, dimension 3 to 5: stat = 0.926, F-approx = 1.424, df1 = 36, df2 = 1953.726, $p < 0.0496$). To interpret which task behavior and/or self-report measures contributed most strongly to each of the significant dimensions, we highlighted those that had a coefficient greater than 0.5 along each dimension Table 2. The first loaded most heavily, on the task measure side, on cognitive effort task performance (increased MSIT error rates) and decreased cognitive effort cost, which were associated, on the self-report side, with increased behavioral activation, decreased (self-reported) cognitive function and decreased anxiety.

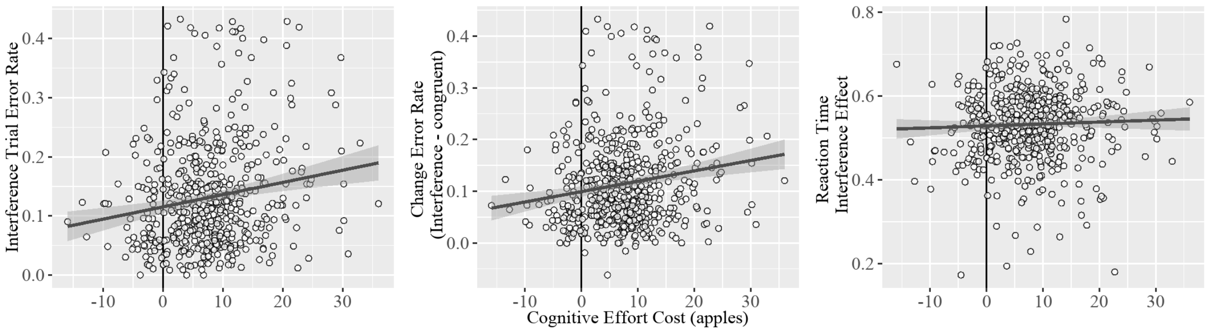

Figure 6: Relationship between task performance and cognitive effort cost. Column 1: error rate on the high effort task (log transformed), column 2: change in error rate from low to high effort condition, column 3: change in reaction time from high to low effort condition. Cognitive effort cost negatively related to error rate measures but not reaction time interference effect.

Canonical dimension (correlation coefficient)	**Task behavior (coefficent)**	**Self-reports (coefficent)**
Dimension 1 (0.30)	↑Congruent Error Rate (0.66) ↑Interference Error Rate (0.60) ↓Cognitive Effort Cost (-0.58)	↑Behavioral Activation (0.77) ↓Cognitive Function (-0.71) ↓Anxiety (-0.72)
Dimension 2 (0.23)	↑Physical Effort Cost (1.30) ↓Cognitive Effort Cost (-0.53)	↓Anxiety (-0.75) ↑Fatigue (0.72)
Dimension 3 (0.18)	↓Cognitive Effort Cost (-1.23) ↑Physical Effort Cost (0.67)	↓Cognitive Function-Abilities (-0.93) ↑Apathy-Social (0.62) ↓Anhedonia (-0.52) ↓Depression (-0.51)

Table 2: Significant canonical dimension (correlation coefficient). For each dimension coefficients larger than 0.5 are displayed in column 2 (for task behavior variables) and in column 3 (for self-reports). Arrows indicate positive or negative coefficients. Dimension 1 was most related to cognitive task performance, whereas dimension 2 was most related to physical effort cost, and dimension 3 most related to cognitive effort cost.

Taken together, this dimension may correspond to carefulness or conscientiousness, and may be closest to "decision acuity" (Moutoussis et al., 2021). The second dimension appeared more closely related, among task measures, to increased physical effort cost, and loaded heavily on this dimension, as well as decreased cognitive effort cost, and was associated, among self-reports, with decreased anxiety, and increased physical fatigue. The second dimension may correspond to physical fatigue and physical effort avoidance, as distinct from cognitive effort avoidance, which is instead most heavily associated with the third dimension. The third dimension was more closely related to decreased cognitive effort cost, and loaded heavily on this dimension, alongside increased physical effort cost, and was associated, for self-report measures, with decreased (self-reported) cognitive function-abilities, increased social apathy, decreased anhedonia, and decreased depression. The third dimension may correspond to cognitive effort avoidance having a role in anhedonia and depression. However the positive relationship between cognitive effort cost and (self-reported) cognitive function-abilities is contrary to our expectations.

Of note, in all three significant dimensions, cognitive and physical effort loaded with opposite signs. This suggests that they reflect distinct aspects of individual differences, associated with separate external measures. Note that this is not inconsistent with the finding that they are nevertheless correlated with each other, overall (Fig. 6): CCA strictly addresses the correlation across, rather than within, domains. These results suggest that the correlation between cognitive and physical effort are not themselves driven by one of the externally self-reported dimensions, such as conscientious performance.

2.5 Validation experiments.

Collateral predictions of the Marginal Value Theorem. In addition to the experiment presented in this article (Experiment 1 (MSIT)), we tested two other versions (Experiments 2 and 3), both to validate the Effort Foraging Task in different groups of participants (see Appendix Section 5.6), and to test its adherence to predictions of the Marginal Value Theorem not tested in Experiment 1. The validation experiments confirm that participants adhere to predictions of the Marginal Value Theorem across manipulations of cognitive effort type, and environment richness.

Effects of different types of cognitive effort. In Experiment 2 (N-Back), we tested a different form of effort as the travel manipulation. Specifically, we compared foraging behavior when the travel task was the 3-Back versus 1-Back level of the N-Back task (Nystrom et al., 2000). As predicted, on average across participants exited trees later in the high (3-Back) relative to low (1-Back) cognitive effort conditions; linear mixed-effects regression estimate for N-Back (3-Back - 1-Back): $= -0.504$ apples, df $=$ 75.981, F $=$ 30.339, $p < 0.001$, physical (smaller-larger) $= -0.448$, df $=$ 75.170, F $=$ 27.151, $p < 0.001$.

Effects of environment richness. In Experiment 3 (Richness), we tested for the effects of reward richness predicted by the Marginal Value Theorem, in which leaner environments (yielding lower overall mean reward rate) should be associated with lower exit thresholds. To test for this effect, we compared two levels of reward richness, by adjusting the mean of a normal distribution used to draw the initial yield of a patch. As predicted, we found that participants lowered their exit thresholds in the lean compared to rich conditions (Richness contrast; sum sq. 0.788, mean sq. 0.788, DenDF $=$ 27.95, F $=$ 10.49, $p < 0.0031$).

3 Discussion

We developed the Effort Foraging Task to quantify the costs of cognitive and physical effort at the level of the individual. Participants played a computer game in which they could forage for virtual apples in a patch with diminishing returns, or abandon that patch for a new (initially) richer patch at the expense of time and effort. Participants completed blocks of the task in which the travel cost was either cognitive and physical effort, each at one of two difficulty levels (high and low effort). We measured their 'exit threshold' as the number of apples the participant could have expected to get on their next harvest on trials in which they decided to travel instead. We found that on average participants lowered their exit threshold (staying longer, accepting diminishing returns) in the high relative to low effort conditions, consistent with the high effort task having a monetary cost. Further analyses in Experiment 1 demonstrated that these cognitive effort costs are correlated with differences in error rates between the easier and harder tasks, suggesting that the costs may at least partially reflect error avoidance. Expected Value of Control model simulations (Musslick et al., 2018) demonstrated the problem of identifiability of effort costs versus other factors that contribute to cognitive effort allocation: skill and reward sensitivity. That is, if someone avoids effort (i.e., restricts allocation of cognitive effort to a demanding task) this could reflect a higher cost of effort, but it could also reflect poorer ability and/or weaker incentives. These individual differences would impact both choices in the foraging task, and performance on the cognitive effort travel task.

We also found that participants' exit thresholds were responsive to all of our other manipulations as well: two different types of cognitive effort (working memory in Experiment 2 vs. interference control in Experiment 1), and environment richness (Experiment 3). Our findings also suggest that the Effort Foraging Task task is suitable for use on crowd sourcing platforms (Prolific, Experiment 1), and for Undergraduates remotely (Experiment 2), yielding results comparable to those found in the laboratory (Experiments 3, as well as previous studies; e.g., Constantino and Daw, 2015).

We investigated the relationship between individual differences in cognitive and physical effort costs. Our design allowed us to measure cognitive and physical effort costs in a common currency, revealing a significant and substantial correlation between these within individuals. This suggests that a common mechanism may compute costs across multiple domains, consistent with research showing overlap in brain areas involved in cognitive and physical effort using human neuroimaging (Chong et al., 2017; Schmidt et al., 2012), and research showing that cognitive fatigue impacts physical effort exertion and fatigue (Giboin & Wolff, 2019; Marcora et al., 2009). Most goals carry multiple types of costs (i.e., cognitive effort, physical effort, time costs), so it is reasonable that cost information would be integrated across domains when evaluating potential goals.

Advantages compared to previous tasks. There are a number of benefits of the Effort Foraging Task as a means of measuring effort costs. First, measure of costs is indirect. Participants are not directly asked if they would rather complete a high effort task or a low effort task. Instead, this is inferred from their exit thresholds. This may mitigate confounds of demand characteristics. We consistently found a subset of participants who were effort-seeking in our task (treating the low effort task as more costly, and overharvesting more in the low compared to high effort conditions), which is not commonly seen in direct tasks (i.e. Treadway et al., 2009; Westbrook et al., 2013).

Direct tasks, such as the Cognitive Effort Discounting Paradigm (Westbrook et al., 2013) may engage real-world economic considerations (e.g., that one should be paid more to work more), which may preclude effort seeking behaviors. This is consistent with the observation that ‘reverse discounting’ is not reported in the Cognitive Effort Discounting Paradigm. A concern common to previous studies is the presentation of two options simultaneously, as this may distort choices or complicate their interpretation (Lopez-Gamundi & Wardle, 2018; Westbrook et al., 2013). Research in intertemporal choice has shown that rats and humans are less impulsive decision makers when they are making sequential choices than simultaneous choices (evaluating a single option and choosing to accept or reject it Carter et al., 2015; Carter & Redish, 2016).

Another strength of the Effort Foraging Task is that participants learned about the environment reward rate and effort costs through experience, and their choices reflected their ongoing evaluation of these quantities. In some previous tasks, choices and the realization of cognitive effort were separated in time (Chong et al., 2017; Westbrook et al., 2013), or choices were hypothetical (see a review in Lopez-Gamundi et al., 2021). In the Effort Foraging Task participants experience the effort immediately after each choice to travel. In addition to being ecologically valid, foraging tasks also lend themselves to formal analysis using the Marginal Value Theorem, which provides a theoretically motivated, and quantitatively rigorous approach to measuring costs.

Relationship between Effort Costs and Self-report Surveys. CCA revealed inter-relationships between Effort Foraging Task variables and self-report proxies of real world motivation and psychiatric symptoms such as depression, anxiety, fatigue, and anhedonia. Three significant dimensions were identified that weighted on multiple cognitive task variables and self-report measures Table 2. This method may be useful in addressing measurement confounds identified by an Expected Value of Control theory simulations (Musslick et al., 2018) which demonstrated the need to measure and account for multiple factors that contribute to motivation for effort (e.g., task automaticity, reward sensitivity) in order to isolate and study a particular component (i.e., effort cost). Multidimensional models like factor analysis and canonical correlation analysis offer a richer view on the interrelationship between multiple variables: in this case, for instance, suggesting that different external factors are most strongly associated with cognitive vs. physical effort costs, vs. task performance (error rates) per se.

The first dimension most reflected cognitive task performance (both congruent and interference trial error rate), the second most reflected physical effort cost, while the third most reflected cognitive effort cost. Higher cognitive effort task error rates and lower cognitive effort cost in the first dimension were associated with increased behavioral activation, decreased (self-reported) cognitive function, and decreased anxiety. Sensibly, we found that individuals who reported better cognitive function in the past week (fewer reported cognitive difficulties on the survey) exhibited better performance in our cognitive effort tasks and lower cognitive effort costs. The association with behavioral activation and anxiety may reflect variability in attentiveness and carefulness to the cognitive effort task. Two aspects of this result speak encouragingly to the external validity of our measures: first that unlike in some previous work (Moutoussis et al., 2021) this attentiveness dimension does not exhaust the relationship between self-report and task-measures. Furthermore, the finding that this dimension is associated with lower cognitive effort costs but, less strongly, larger physical effort costs, suggests that such nonspecific attentiveness is not driving the relationship between cognitive and physical effort costs. All this may be enabled by our relatively careful exclusion of inattentive participants, who can otherwise drive uninformative correlations and obscure more informative relationships (Zorowitz et al., 2021).

The second and third dimensions more closely addressed our primary interest, physical and cognitive effort cost per se, suggesting they have somewhat distinct external correlates. The second dimension was most heavily associated with increased physical effort cost, but also decreased cognitive effort cost. Among the self-report measures it was associated with decreased anxiety, and increased fatigue. The connection between physical effort cost and physical fatigue supports the face validity of the Physical Effort Foraging Task. It could be of clinical relevance if a component of fatigue is due to increased physical effort costs, which could be targeted by interventions which boost motivation.

The third dimension was dominated by decreased cognitive effort cost, but also increased physical effort cost. Among the self report measures, it was most strongly associated with decreased cognitive function-abilities, as well as increased social apathy, decreased anhedonia and decreased depression.

The positive association between cognitive effort cost and cognitive function-abilities is contrary to our predictions, and shows the opposite pattern as the first dimension relationship to cognitive function (asked in terms of difficulties). The relationship between increased depression and anhedonia and higher cognitive effort cost is consistent several other studies reporting increased cognitive effort costs associated with depression (Hershenberg et al., 2016; Marchetti et al., 2018; Patzelt et al., 2019; Tran et al., 2021; Westbrook et al., 2022). We are currently running a clinical study of depression to further decompose the relationship between cognitive and physical effort decisions and specific domains of symptoms, based on recent theoretical work within the Expected Value of Control framework (Grahek et al., 2019).

Cognitive effort seeking. In all of the experiments reported here, we consistently observed a subset of participants who exhibited *negative* cognitive effort costs (i.e., a preference for the high effort option over the low effort option), suggestive of cognitive effort seeking. This is a phenomenon that has been consistently observed in the literature (for a review see Inzlicht et al., 2018), often referred to as the "need for cognition." Although this comprised a minority of participants in our experiments, nevertheless it suggests the need for extensions to existing utility models (i.e., the Expected Value of Control model). Cognitive effort seeking, *prima facie*, indicates positive value assigned to exerting cognitive effort, which may reflect — directly or indirectly — longer term value attached to information-seeking and learning that yield better future performance (e.g., Agrawal et al., 2019; Geana et al., 2016). In a related vein, effort seeking may also have to do with boredom, which may hold a disutility that encourages application of effort (e.g., Agrawal et al., 2022). This is adaptive because effort is valuable so doing nothing carries an opportunity cost. Each of these factors likely comprise their own dimensions of individual variation that were not measured in our experiments. Future research is needed to further investigate the factors that drive effort seeking in this task and others.

Opportunities and future directions. Further work using the Effort Foraging Tasks to measure cognitive and physical effort, and the relationship of these to one another, may be useful in furthering our understanding of the nature of such costs and how they are evaluated, including the testing of alternative accounts (i.e. opportunity cost, processing, and metabolic accounts Baumeister & Heatherton, 1996; Kurzban et al., 2013; Musslick & Cohen, 2021). For example, the Effort Foraging task could be useful for testing opportunity cost accounts, by manipulating whether or not a low effort alternative task is available (e.g., browsing social media instead of completing the effort foraging task for money). Opportunity cost accounts would predict that the cognitive effort cost measured by foraging behavior would be higher during periods in which an alternative was on offer (Kurzban et al., 2013). To test cost of processing accounts, the travel task could involve multi-tasking. By these accounts, participants should treat multi-task sets that recruit more shared representations as more costly than sets that recruit more separated representations (Musslick & Cohen, 2021). Developmental studies using this task may also help understand the interplay between learning, information-seeking, and effort (Chevalier, 2018; Munakata et al., 2012; J. Niebaum & Munakata, 2020; J. C. Niebaum et al., 2019; Snyder & Munakata, 2010).

Future research could also use this task to integrate effort into the study of patch foraging in both human and animal studies. For example, increasing vigor decreases opportunity costs of time while increasing physical effort costs (Niv et al., 2007). Similarly, increasing attention during planning may increase cognitive effort costs while reducing opportunity cost of time due to speeding up goal attainment. There is much to be learned about these tradeoffs that must be managed by the forager. The present study moves in this direction by adding a cognitive or physical effort task requirement to travel in a virtual patch foraging environment with human participants.

Conversely, insights from patch foraging may prove useful in the study of effort-based decision making and cognitive control, such as in task switching paradigms. When deciding whether to persist in pursuing the current task or switch to an another, people may use a threshold rule based on average reward rate (like the Marginal Value Theorem). Therefore the Effort Foraging Task could be used to advance understanding of both foraging and effort-based decision-making, as well as the relationship between them.

Our study also highlights the potential utility of the Effort Foraging Task for studying individual differences in effort costs. For example, to what extent are effort costs trait-like or state-dependent, and how generalizable are they are across different task domains (for example different types of cognitive effort beyond that required for the interference and working memory tasks used here). Intriguingly, research using the Cognitive Effort Discounting Paradigm suggests that cognitive effort motivation may be trait-like, encompassing multiple domains of effort (Crawford et al., 2022). Variations on foraging tasks can be used to study a variety of domains of costs (i.e., risk, ambiguity, and delay preferences), and the inter-relationships between domains of cost within an individual. In a related vein, our study also illustrates the potential value in studying the relationship between empirically measured, quantifiable forms of effort costs (both cognitive and physical) and affective factors relevant to mental health, such as apathy, anhedonia and depression. One promising extension would be to combine this approach with methods of ecological momentary assessment (Krönke et al., 2021; Strobel et al., 2020).

Effort-based decision-making has considerable importance in daily life. Critical questions remain about how to disentangle aspects of motivation for effort, how these aspects are represented in the brain, and the role they play in real-world behaviors. The cognitive computational study of motivation has the potential to help people reach their goals by identifying the mechanisms of motivation and ways to enhance motivation towards what matters most to an individual.

4 Materials and Methods

4.1 Trial structure

On each trial an image of a tree appeared on the screen, representing an immediately available source of reward. Participants could choose to harvest that patch (tree) or travel to a new, replenished patch (Fig. 1). When a tree was harvested it 'shook' and apples were displayed under it (apples were displayed in a single, left justified, row). Reward depleted within a patch such that the more times a tree was harvested the fewer apples it produced. When participants choose to leave the patch they had to "travel" which consisted of completing a cognitively or physically effortful task. Participants had a fixed amount of time to collect apples (money). Therefore they must balance the diminishing returns associated with staying at a patch with the travel costs required to reach a new, replenished patch. To indicate the start of a trial a circle below the tree turned white and participants were able to make their decision. The circle below the tree was brown when participants could not make a decision (apples being displayed, or waiting through harvest delay). If participants took too long to make a decision (1 second deadline) a message "Too slow" appeared, after which they waited the harvest delay (2 seconds total). Participants were instructed that the more "too slow" warnings they saw the fewer apples they would earn.

4.1.1 Block-wise manipulation

Patches were presented block-wise. We manipulated two factors that defined a block; effort type (cognitive and physical) and effort level (low and high). Each block type was tested twice, making 8 blocks total. The total duration of the block was fixed (4 minutes). Participants had a self-paced break between blocks. Participants were instructed that the time in an block was fixed at 4 minutes, and that they had to decide how to spend their time between harvesting and traveling. The cognitive and physical variants of the task were completed separately (i.e. all cognitive effort blocks were completed in sequence, as were all physical effort blocks). The order of cognitive and physical effort variants of the task was counterbalanced across participants. Participants did not know when playing the first effort variant that there would be a second variant upcoming in the experiment. Within blocks of an effort type, each effort level was tested once during the first half and once during the second half. Given that constraint, the effort level was fully counterbalanced, resulting in eight possible block orders. Which of the block orders was used was randomly selected for each participant. Participants were explicitly instructed about which travel task they had to perform in a particular block. In addition they were instructed that throughout the task they could use the background color to know which effort level to perform (light blue for cognitive low effort, light orange for cognitive high effort, light purple for physical low effort, light green for physical high effort).

4.1.2 Task environment

The only difference between blocks was the effort travel task, all other variables of the foraging environment were fixed (Table 3). The time it took to harvest the tree (2 seconds) or travel to a new tree (8.33 seconds) were fixed. The apple yield of the first visit to a patch was drawn from a normal distribution ($\mathcal{N}(15 \text{ apples}, 1)$, maximum = 20 apples). Each following yield was the product of the previous yield and the depletion rate. The depletion rate was drawn every harvest trial from a beta distribution ($\alpha = 14.909$, $\beta = 2.033$) and was on average 0.88 (minimum apple reward for a patch was 0.5 apples). Participants had up to 1 second to respond, or else they were shown the "too slow" message. When

Harvest time	2 seconds
Travel time	8.33 seconds
Cognitive Travel Tasks	MSIT Congruent, Interference
Physical Travel Tasks	Smaller (50%), Larger presses (100% max)
Block duration	4 minutes
Number of blocks per condition	2
Initial reward	$\mathcal{N}(15, 1)$
Decay rate	$\beta(14.13, 2.03)$
Total Number of blocks	8

Table 3: Foraging environment parameters. First column: environment parameter. Second column: Parameter values.

participants harvested the patch, apples appeared on the screen for 1 second. Regardless of the reaction time, the total harvest delay was always 2 seconds long. When the participant decided to exit, the tree moved from the center of the screen leftwards until it went off the screen (415 millisecond animation) then the travel task occupied the screen (7.5 seconds), after which the tree reappeared from the right side of the screen and moved leftwards towards the center of the screen (415 millisecond animation).

4.1.3 Foraging task training

The task began with training the travel task for the first effort cost variant for a particular participant (this could be the cognitive or physical effort task). Next we administered instructions for the foraging task in general (without mentioning the effortful travel requirement), and participants completed a practice block (90 seconds) of the foraging task with no travel task. Then we instructed participants that they would have to complete the effortful travel task when traveling, and they completed two practice blocks (one per effort level, 90 seconds each). When training was complete participants completed the main foraging task for the first travel task type (4 blocks, 4 minutes per block, with self-paced breaks between blocks). After completing all of the blocks of the first travel task, participants began training on the second travel task. Then they were instructed that they would continue to play the foraging task but the travel task had changed. They practiced the foraging task with the second travel task type (one practice block per effort level, 90 seconds each). Finally, they completed the main foraging task for the second travel task type (4 blocks, 4 minutes per block).

4.2 Travel Tasks

4.2.1 Multi-Source Interference Task.

We used the Multi-Source Interference Task as the cognitive effort task. This task includes multiple types of interference effects; Stroop, Flanker, and Simon effects, and is simple to administer with a standard keyboard without the need for participants' to learn novel key mappings (Bush & Shin, 2006). The MSIT trial began with 250 ms fixation cross, then the stimulus appeared for 1000 ms and participants could enter their response. After a total of 1250 ms the trial ended. Participants completed 6 trials per travel for a total of 7.5 seconds of task time. If participants made two errors in a row they saw an attention check (black dot) for 250ms instead of the fixation cross. Participants were instructed to avoid seeing the black dot.

4.2.2 Rapid Key-pressing Task.

Participants performed rapid key-pressing as part of foraging task during travel between trees (7.5 seconds, Fig. 2 right panel). In the task participants rapidly pressed the keyboard with their non-dominant pinky finger. All participants were right handed, and used their left pinky finger to press (the 'a' key). Each press filled a bar that spanned the horizontal extent of the screen. The horizontal bar indicated progress towards the goal number of presses. There were two conditions referred to as the "Larger number of presses" and "Smaller number of presses". Travel time was fixed, so if participants reached the goal presses before the travel duration they waited and saw the message 'Completed!' on screen. If they failed to complete the goal number of presses a black dot appeared on the screen. Participants were instructed to avoid seeing the black dot. To ensure within reason that participants used their non-dominant pinky finger throughout the task, they were required to press 'hold keys' to occupy other fingers. The hold keys were 'w', 'e', and 'f' for the left hand, and 'h' and 'o' for the right hand. To minimize cognitive demands the hold keys were always displayed at the bottom of the screen during the rapid key-pressing task.

Rapid Key-pressing Task training. Key-press training began with a calibration phase (three rounds) to determine the maximum number of presses participants were able to complete in the travel time (7.5 seconds of effort task time). A counter was displayed on the center of the screen showing how many presses a participant had made. The instructions suggested participants were being compared to others, and encouraged them to press as fast as possible, each round they were encouraged to press faster than they had the previous rounds. Then we used a participants' maximum number of presses of all of the rounds and set this as their maximum number. The Larger Number of Presses condition demanded participants complete 100% of their maximum, and the Smaller Number of Presses condition demanded participants complete 50% of their maximum. Participants were only told there was a larger and smaller number, and not what that number was or how it was determined. Then participants practiced a single effort level. Effort level order was counterbalanced. Practice for an effort level began with a single mini-block the duration of the foraging travel time. Then participants had to complete 5 mini-blocks successfully (reaching the required number of presses) to move on. This was meant to establish the expectation that participants would perform well on the travel task, even though there were no incentives or punishments associated with task performance during the foraging task.

4.3 Overview of experiment

The experiment was conducted over an hour session. Participants gave electronic informed consent to participate in the study. All tasks and surveys were presented using the jsPsych library for JavaScript (de Leeuw, 2015), and served with using NivTurk software (Zorowitz & Bennett, 2022) using the Flask software package for Python. Participants began the experiment with self-report measures, followed by the foraging training, the main foraging task, and lastly a debrief survey including demographics.

Participants 678 Prolific participants (18-56 years, mean = 24.5 years ± 6.7, 307 female, 365 male, and 6 prefer not to answer) volunteered for the study. The study was approved by the Princeton University Institutional Review Board and participants were recruited from the Prolific platform for the large online sample. Participants were compensated with $8.33 for one hour a performance bonus up to $4 (Prolific bonus mean = $3.52, standard deviation = 0.78, range = $0.35 - 4). The total number of apples harvested in the Effort Foraging Task were converted into real money at the end of the experiment, with each apple being worth fractions of a cent (0.009 cents per apple). The conversion factor was set using pilot data, such that the best performing participant (earned the most apples) would make the maximum bonus. To accommodate both the physical effort task (completed with the non-dominant pinky finger) and the foraging task within standard keyboard layout, all participants were right-handed. Participants completed foraging decisions with their right hand and effort travel tasks with their left hand.

Self-reports Motivated by recent theoretical work (Grahek et al., 2019) we sought to test whether cognitive effort costs were predictive of symptoms of depression. We created a comprehensive battery of self-report surveys measuring many aspects of cognitive and physical amotivation (i.e., apathy, anhedonia, depression, Le Heron, Apps., et al., 2018; Le Heron et al., 2019). At the start of the experiment participants completed the self-report battery (Table 4). The first 98 participants were not given a battery, leaving 458 participants who completed the self-reports and were not excluded based on their foraging behavior. Of these, 28 participants failed at least one attention check item embedded in the surveys and were excluded for self-report based analyses (see Appendix Section 5.3).

The first was the Need for Cognition scale (Cacioppo & Petty, 1982; Cacioppo et al., 1984). This self-report scale measures the extent to which individuals are prone towards engaging in cognitively effortful activities. Previous studies have shown that Need for Cognition negatively correlates with individual differences in effort avoidance (Westbrook et al., 2013) so we predicted a negative relationship with cognitive effort cost. This relationship would serve as evidence for convergent validity for the Effort Foraging Task. Next was the Behavioral Inhibition, Behavioral Activation Scales (BIS/BAS Carver & White, 1994) which measures an individuals' sensitivity to the behavioral approach system and behavioral avoidance system. We used the abbreviated form of the BIS/BAS to reduce the total number of items (Pagliaccio et al., 2016). This measure was useful for several purposes, firstly, we used it as a measure of reward sensitivity, which was a factor we wanted to capture as it is a potential confound. Research has

Scale name	**N items**	**Abbreviation**	**Time scale**
Patient Health Questionnaire-9	9	PHQ-9	last 2 weeks
Generalized Anxiety Disorder-7	7	GAD7	last 2 weeks
Apathy Motivation Index	14	AMI	last 2 weeks
Snaith-Hamilton Pleasure Scale	13	SHAPS	last few days
PROMIS Cognitive Function Short Form 4a	4	PROMIS-CF	last 7 days
PROMIS Cognitive Function Abilities Short Form 4a	4	PROMIS-CFA	last 7 days
PROMIS Fatigue Short Form 4a	4	PROMIS-F	last 7 days
PROMIS General Self-Efficacy	4	PROMIS-SEFF	last 7 days
Need for Cognition	18	NFC	trait
Adult Temperament Questionnaire-Effortful Control	19	ATQ-EF	trait
Behavioral Inhibition/Activation Scales - Abbreviated	12	BIS/BAS	trait
Total	108		

Table 4: Self-report battery. Self-reports were completed in the order listed in this table. Scales asking about similar timescales were grouped together.

shown the BAS is predictive of striatal activation in anticipation of rewards (Costumero et al., 2016). Secondly, research on depression shows that Behavioral Activation scores tend to be lower and Behavioral Inhibition scores tend to be higher in more depressed participants (Alloy et al., 2008; Kasch et al., 2002; McFarland et al., 2006; Pinto-Meza et al., 2006; Quilty et al., 2014). Third was the Apathy Motivation Index (Ang et al., 2017) which measures apathy and motivation in the behavioral, social, and emotional domains and was designed to be suitable for use in the general population. Lastly we administered the Snaith–Hamilton Pleasure Scale (SHAPS, Snaith et al., 1995) which measures anhedonia by asking about responses to common domains of pleasure. We administered the Patient Health Questionnaire-9 (PHQ-9, Kroenke et al., 2001), a common measure of depression severity, and also the Generalized Anxiety Disorder-7, a common measure of anxiety severity (Spitzer et al., 2006).

We administered four scales from the Patient-Reported Outcomes Measurement Information System (PROMIS, Cella et al., 2007); the Cognitive Function Short Form 4a, the Cognitive Function Abilities Short Form 4a, Fatigue Short Form 4a, and General Self-Efficacy. We were interested in cognitive symptoms of depression such as slowed thinking and reduced concentration. To measure this we used the PROMIS Cognitive Function Short Form 4a which measures subjective cognitive functioning. For this sub-scale higher scores indicate fewer complaints about recent cognitive function. We also used the complimentary PROMIS Cognitive Function Abilities Short Form 4a, for which higher scores indicate better subjective cognitive function. We were also interested in physical fatigue symptoms of depression which we hypothesized would be correlated with greater physical effort costs, we measured this using the PROMIS Fatigue Short Form 4a. The theoretical work by Grahek and colleagues 2018 also identified self-efficacy as a potentially relevant factor in cognitive control decision-making in depression. By this account more depressed participants may be less likely to predict that exerting effort will lead to a positive outcome. We made a first attempt at measuring this with the PROMIS General Self-Efficacy sub-scale, however the ideal way to measure this factor would be with another cognitive task (see Frömer et al., 2021). We also added a self-report measure of cognitive control capacity, which we hoped to contrast with cognitive control motivation measures. We used the Adult Temperament Questionnaire - Effortful control subscale (Evans & Rothbart, 2007) which has been related to depression in a previous study (Marchetti et al., 2018).

4.4 Analysis methods

4.4.1 Hierarchical Bayesian Model Methods.

We created a model based on the Marginal Value Theorem (Charnov, 1976) to predict participants choices to harvest or exit a patch. First we computed known reward rate values of the foraging environment per effort condition per participant: total rewards harvested, number of harvest periods, and total times travelled (see foraging environment parameters in Table S2). Then, we solved for the unknown component of average reward rate; the cost of travel (c). We estimated the cost of the high effort task ($c_{high\,effort}$) for an individual by predicting harvest versus exit decisions using a hierarchical Bayesian logistic model (Eq. (3) and Eq. (5)). For each foraging trial, model compares the expected reward on the next harvest (R_e, defined as the average of the previous harvest and the product of the previous harvest with the mean depletion rate (0.88)) against the overall average reward rate for a block type (ρ), using a softmax function (with inverse temperature parameter, β) to make a choice (harvest or exit). The cost of travel in high effort blocks ($c_{high\,effort}$) was expressed as the marginal increase in cost of travel ($c_{low\,effort} + c_{high\,effort}$) from low to high effort. Defining this cost as a difference measure controls for any additional biases individual participants may have which are common to both conditions (i.e., consistently high exit thresholds for some participants and low thresholds for others). We used ($c_{high\,effort}$) as the dependent measure of the effort cost for an individual.

For each effort level (low and high) and effort type (cognitive and physical) we predicted choices to stay or exit a patch:

$$P(stay_{low\,effort}) = \frac{1}{1 + \exp\left(\beta(R_e - \rho_{low\,effort})\right)} \quad , \tag{2}$$

where,

$$\rho_{low\,effort} = \frac{\sum r - \sum c_{low\,effort}}{T_{low\,effort}} \tag{3}$$

and,

$$P(stay_{high\,effort}) = \frac{1}{1 + \exp\left(\beta(R_e - \rho_{high\,effort})\right)}, \tag{4}$$

where,

$$\rho_{high\,effort} = \frac{\sum r - \sum c_{low\,effort} + c_{high\,effort}}{T_{high\,effort}} \tag{5}$$

There were five parameters in the model, the inverse temperature (β), the cognitive low ($c_{cog\,low\,effort}$) and high effort costs ($c_{cog\,high\,effort}$), and the physical low ($c_{phys\,low\,effort}$ and high effort costs ($c_{phys\,high\,effort}$). The model included a full covariance matrix of the parameters (5-by-5 matrix) which consists of a correlation matrix and a scale (standard deviation) matrix. Parameters were drawn from a multi-variate Gaussian distribution. We used the covariance matrix to directly estimate the correlation between individual differences in high cognitive and physical effort costs. For both models the prior distributions were $c_{low\,effort} \sim \mathcal{N}(0, 40)$, $c_{high\,effort} \sim \mathcal{N}(0, 30)$, $\beta \sim \mathcal{N}(0, 0.5)$. The prior on the correlation matrix was unbiased as to the presence or absence of a correlation (LKJ Correlation Distribution prior = 1, (Lewandowski et al., 2009)). Individual participant parameters and their group-level distributions were estimated using Markov Chain Monte Carlo sampling, implemented in Stan with the CmdStanR package (4,000 samples, 2,000 warm-up samples, across 4 chains, Stan Development Team, 2021).

4.4.2 Canonical correlation analysis.

To leverage the strength of our data in having many detailed individual differences measures of theoretically related constructs we used Canonical Correlation analysis to preform a many to many correlation (we used the cc function from the CCA package because it can handle missing data, González & Déjean, 2021, in the R language). The task measures included were; cognitive effort cost, interference and congruent trial error rate (transformed as log(2-correct)), physical effort cost, and overall exit threshold (estimated in log apples over all blocks by participant using linear mixed-effects regression). The self-reports were the Need for Cognition, Executive Function (Adult-temperament questionnaire), PROMIS Cognitive Function and Cognitive Function-Abilities, Behavioral Inhibition and Behavioral Activation, the Apathy Motivation Index sub-scales, anhedonia (Snaith Hamilton Pleasure Scale), physical fatigue (PROMIS Fatigue), Self-efficacy (PROMIS), anxiety (Generalized Anxiety Disorder-7), and depression (Patient Health Questionnaire-9).

It is possible that using task parameter estimates fitted using a correlation matrix might have biased the CCA. We re-ran the analysis with a conservative model that had no covariance matrix between parameters. Everything in the model was the same except that participant parameters were drawn from a normal distribution with the mean of the group-level parameter and standard deviation $N(0.5, 0.5)$ for the inverse temperature, $N(40, 40)$ for the cognitive and physical low effort costs, and $N(30, 30)$ for the cognitive and physical high effort costs. We repeated the canonical correlation analysis with the parameter estimates of the model without the correlation matrix.

Initial Sample Size	678
Final Sample Size	537
Number of outlier participants	
Did not complete experiment	5
Foraging trial response deadline missed ($>$ 10% trials)	30
MSIT Congruent error rate ($>$ 2SD)	10
MSIT Interference error rate ($>$ 2SD)	51
Smaller number of presses uncompleted (%,$>$2SD)	12
Larger number of presses uncompleted (%,$>$2SD)	29
Number of exit trials per condition ($<$2SD)	11
Change exit threshold cognitive ($<$2SD | $>$2SD)	30
Change exit threshold physical ($<$2SD | $>$2SD)	31

Table 5: Effort Foraging Task Behavior Based Exclusion Methods. Column 1: basis of exclusion, column 2: numbers for Experiment. Number of participants outliers by exclusion criteria by experiment. Participants could be excluded on multiple grounds, therefore the number of outlier participants are listed but total don't add up to the number of excluded participants in the top two rows.

4.4.3 Exclusion criteria.

Participants completed the study on their own outside of the laboratory. To ensure data quality we used task behavior to constrain our sample to participants who completed the experiment in earnest. The exclusion criteria were; not completing the experiment, missing the response deadline on a large number of harvesting trials, poor cognitive or physical travel task performance, too few exit trials in a condition, outlier in change in exit threshold from low to high cognitive and physical effort conditions. Participants with very large shifts in thresholds produced strong outliers in our Marginal Value Theorem model, furthermore, many of the excluded participants had very few exit trials in one condition (for example harvesting each tree many times to avoid travel). Table 5 shows the number of participants who were outliers as well as the overall number of participants excluded per study. In greater detail, a participant was excluded if they missed a large number of harvest trials (if they did not respond after 1 second deadline in greater than 10% of all harvest trials). Participants were excluded if they performed poorly on any of the travel tasks (using the metrics MSIT congruent and interference trial error rate, percent smaller presses completed, percent larger presses completed). For each travel task we computed the group mean and standard deviation, and excluded participants who were 2 standard deviations below the group mean performance.

Acknowledgements Thanks to Sara Constantino for providing base task code for the in-laboratory experiments and Evan Russek for providing base task code for the remote experiments. Thanks to Allison Burton, Temitope Oshinowo, and Connor Lawhead for collecting the in laboratory data. Thanks to Elizabeth Tong, Temitope Oshinowo, and Jeremy Lee for adapting and developing online task code and for online data collection. Thanks to Kevin Lloyd for providing the base code for the optimal threshold simulations. Thanks to Sebastien Bouret for helpful comments regarding the relationship between cognitive and physical effort costs. Thanks Yoel Sanchez Araujo, Sam Zorowitz, Ben Singer and Dave Turner for assistance with cluster computing for model fitting. Thanks to Lindsay Hunter for feedback on the manuscript.

This work was supported by the John Templeton Foundation, the National Science Foundation Graduate Research Fellowship Program, and the National Institutes of Health Grant T32MH065214. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

5 Appendix

5.1 MSIT Performance Results

There were individual differences in MSIT performance Fig. S.1. Participants bonus earnings were not influenced by their performance. In the main experiment we did not set a performance criterion because that would have complicated the interpretation of the foraging behavior (we would have to estimate not just costs but, for example, subjective efficacy estimates per participant). However in the training established the expectation that participants had to try to be accurate while performing cognitive task by tasking participants with completing a certain number of mini-blocks with high accuracy.

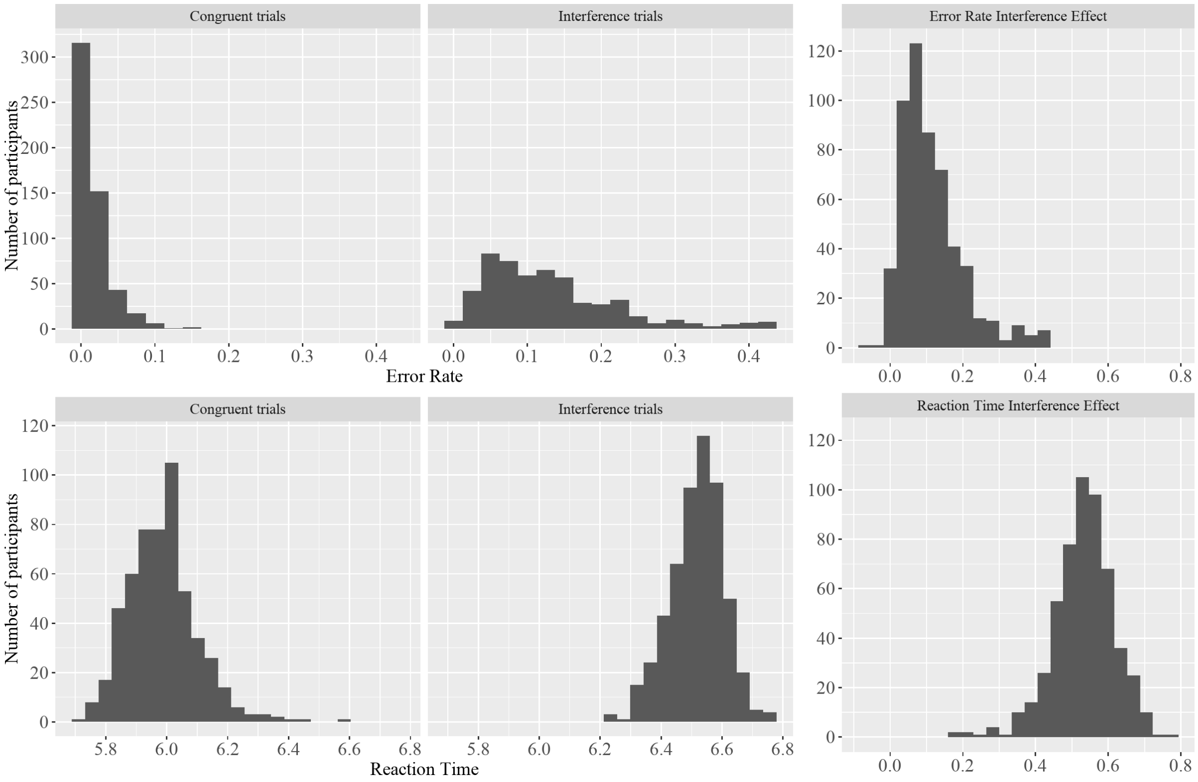

Figure S.1: Individual differences in MSIT performance. Histogram of individual differences in MSIT performance for non-excluded participants in Experiment 1. Top row: error rate (computed as log(2-correct)), bottom row: reaction time. Column 1: congruent trials, column 2: interference trials, column 3: interference effect (interference minus congruent).

5.2 Canonical correlation results

For ease of interpretation we created a table to display variables that were most associated with each dimension (using a coefficient threshold of 0.5). Complete results for the canonical correlation are in Fig. S.2.

5.3 Self-report attention checks

Infrequent attention check items were embedded in the self-report surveys to ensure participants were reading the items (following Zorowitz et al., 2021). 28 participants were excluded from self-report analyses (i.e., CCA) because they failed either of the attention check items embedded in the Apathy Motivation Index, and Generalized Anxiety Disorder-7 self-reports (see items in Fig. S.3). We did not use the items embedded in PROMIS-Cognitive Function or the Patient Health Questionnaire-9 self-reports because response patterns indicated ambiguity in the questions.

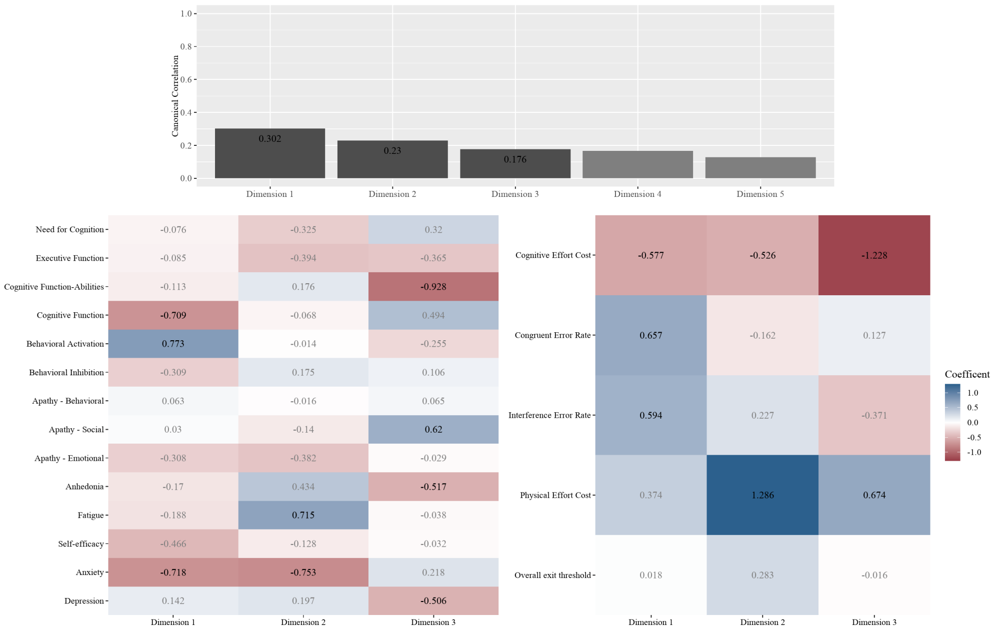

Figure S.2: Canonical correlation results. Top panel: canonical correlations by dimension. Bottom left panel: dimension coefficients for task parameters (X coefficients). Bottom right panel: dimension coefficients for self-reports (Y coefficients). Coefficients with absolute value larger than 0.5 shown in black. Only coefficients from significant dimensions are displayed.

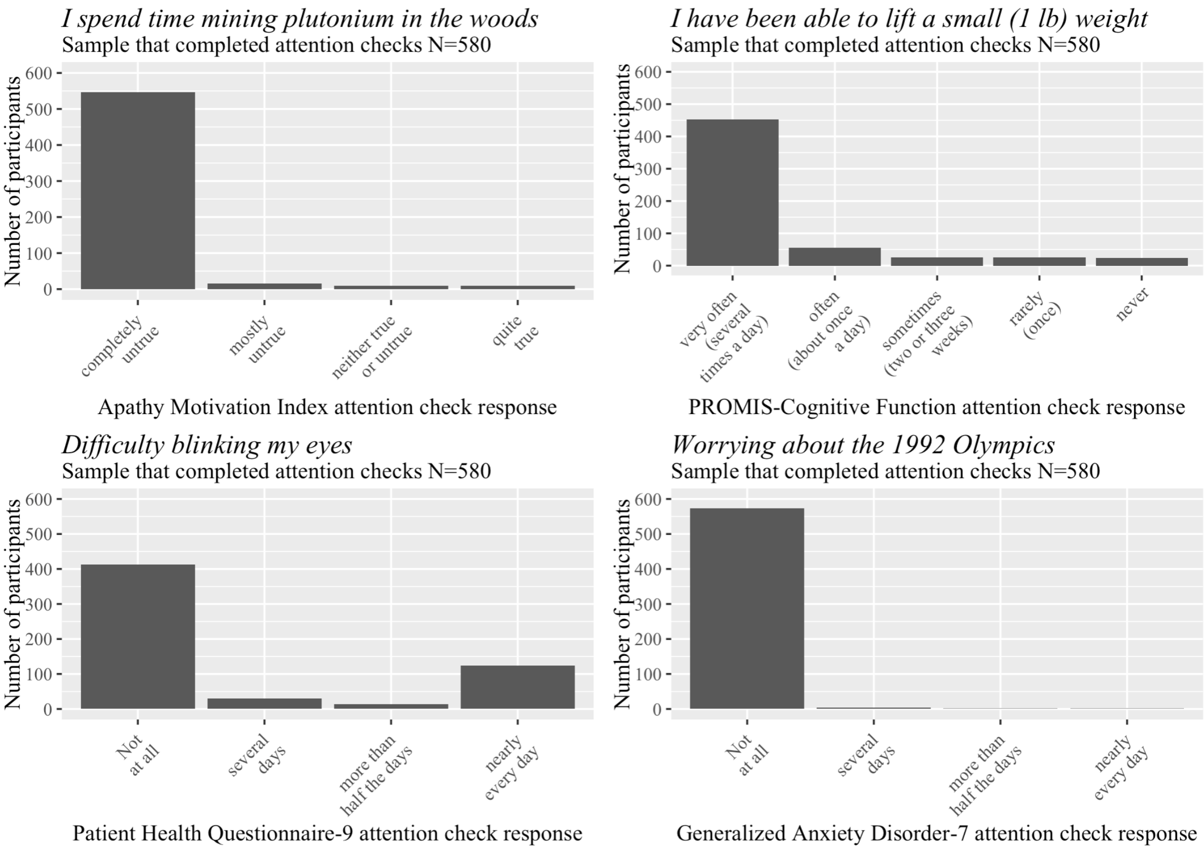

Figure S.3: Attention check items results. Correct answer was left-most responses. Participants were excluded if they were inattentive on the Apathy Motivation Index, and Generalized Anxiety Disorder-7 items.

	Experiment 1 (MSIT)	**Experiment 2 (N-Back)**	**Experiment 3 (Rich condition)**
Best threshold	6.78 apples	4.65 apples	6.17 apples
Reward rate	3.39	2.32	2.37
Number of harvests μ	6.77	9.95	9
Number of harvests SD	1.69	2.01	2.56

Table S1: Best exit threshold policy in simulated data. Columns: Experiments. Row 1: Best threshold from simulation. Row 2: Reward rate achieved with best threshold. Rows 3-4 How many harvests it took to reach the best threshold.

5.4 Simulation to find best threshold.

We simulated the best foraging threshold by creating a foraging environment with an agent with a fixed exit threshold and observing the resulting reward rate. We used a policy iteration algorithm to find the maximal reward rate for a given foraging environment. The foraging environment was defined by the following parameters from our experiments; the harvest time (2 seconds), travel time (8.33 seconds), the distribution of initial rewards to a tree $N(15, 1)$ distribution of the decay function (beta distribution, $\beta(14.90873, 2.033008)$). We assumed the agent knew the mean depletion rate (0.88 multiplied by the previous reward) and used this value to predict the expected reward on the current trial. If the predicted reward was less than or equal to the agent's threshold it exited the patch $R_e \leq \rho$, otherwise it harvested the patch which yielded reward. We simulated 840 'seconds' of foraging time for all experiments (though the result should be robust to duration). The simulation outputs were the 'best threshold' (threshold that yielded the highest reward rate, results vary slightly by simulation run), the resulting 'best reward rate', as well as the mean and standard deviation number of harvests to reach that exit threshold.

The agents' threshold parameter was initialized at 4 apples. For an iteration i, the threshold was set as the mean reward rate observed in iteration i-1, this allowed the threshold to gradually improve in terms of reward rate between iterations. The simulation stopped and the best threshold was determined based on the stopping threshold of a 0.001 apple per second improvement in reward rate on iteration i compared iteration i-1 (with a maximum of 200 iterations). Best exit threshold policy in simulated data (not including effort costs) was 6.78 apples, the reward rate achieved with best threshold was 3.39 apples per second, and on average it took 6.77 ± 1.69 harvests to reach the best threshold (Table S1).

5.5 Overall threshold results

We found that the mean exit threshold across all conditions in Experiment 1 was 6.30 apples (SE = 0.11, df = 615.95, t = 56.09, $p < 0.001$). The group average was close to the best threshold identified by simulation (6.78 apples), however individuals varied widely. Using linear regression we estimated the mean exit threshold across all conditions ("overall threshold") per participant and included these estimates as an additional dimension of individual differences in the task (when testing the relationship between task behavior and surveys). The overall exit threshold may be a relevant individual difference in representations of subjective reward rate (Fig. S.5), a benefit of this task is that it can simultaneously measure exit thresholds and effort costs. Striatal dopamine is hypothesized to represent average reward rate in foraging settings (Constantino et al., 2017; Le Heron et al., 2020). Consistent with this exit thresholds are lower (more over-harvesting) in individuals with Parkinson's (Constantino et al., 2017; Le Heron, Plant, et al., 2018), when participants are chronically or acutely stressed (Lenow et al., 2017), and in individuals with opioid dependence (Raio et al., 2022).

5.6 Validation Experiments

We developed two cognitive effort variants of the effort foraging task. In both versions we used a cognitive and a physical effort manipulation. Experiment 2 (N-Back) was developed in an undergraduate population. Experiment 1 (MSIT) was an abbreviated version developed for a large-scale online study. This tested the generalizability of the task in terms of population as well as in the type of cognitive effort (working memory versus inhibition).

5.6.1 Experiment 2 (N-Back)

Participants completed the study on their own outside of the laboratory. Participants were excluded if they performed poorly on any of the travel tasks (using the metrics MSIT congruent accuracy, MSIT

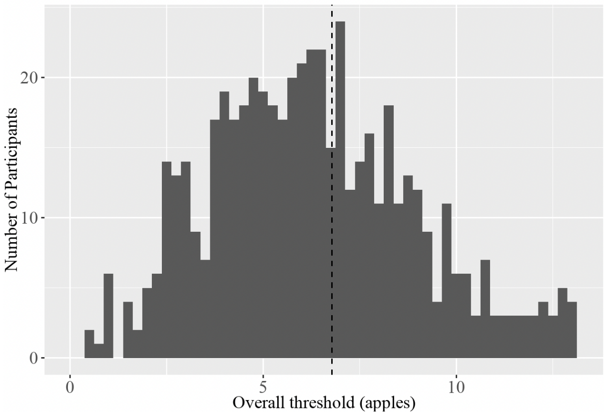

Figure S.3: Individual differences in overall threshold. Histogram of individual differences in mean foraging exit threshold estimated using mixed-effects regression Experiment 1 (MSIT). Some individuals over-harvest (exit threshold below best threshold, dotted line, 6.78 apples) while others under-harvest (exit threshold above best threshold).

Parameter	**Experiment 1 (MSIT)**	**Experiment 2 (N-Back)**	**Experiment 3 (Richness)**
Harvest time	2 seconds		
Travel time	8.33 seconds	20 seconds	
Cognitive Travel Tasks	MSIT Congruent, Interference	1-Back, 3-Back	
Physical Travel Tasks	Larger (100% max), Smaller (50%) number of presses)		
Block duration	4 minutes	7 minutes	
Number blocks condition	2		
Initial reward	$\mathcal{N}(15, 1)$		$\mathcal{N}(15, 1)$, $\mathcal{N}(20, 1)$
Decay rate	$\beta(14.13,\ 2.03)$		
Total Number of blocks	8		8

Table S2: Foraging environment parameters comparison chart. First column: environment parameter. Second column: Experiment 1 (MSIT). Third column: Experiment 2 (N-Back).

Experiment	**Experiment 2 (N-Back)**
Initial Sample Size	116
Final Sample Size	81
Number of outlier participants	
Did not complete experiment	9
Foraging trial response deadline missed (> 10% trials)	7
1-Back accuracy (D', < 2SD)	5
3-Back accuracy (D', < 2SD)	5
Smaller number of presses uncompleted (%,>2SD)	9
Larger number of presses uncompleted (%,>2SD)	5
Number of exit trials per condition (<2SD)	12
Change exit threshold cognitive (<2SD | >2SD)	5
Change exit threshold physical (<2SD | >2SD)	6

Table S3: Experiment 2 (N-Back) Exclusion Methods. Column 1: basis of exclusion, column 2: numbers for Experiment 2 (N-Back). Number of participants outliers by exclusion criteria by experiment. Participants could be excluded on multiple grounds, therefore the number of outlier participants are listed but total won't add up to the number of excluded participants in the top two rows.

interference accuracy, 1-Back D', 3-Back D', percent smaller presses completed, percent larger presses completed).

Methods 116 Undergraduate students volunteered for a two and a half hour self-guided remote experiment (Experiment 2 (N-Back), 18-27 years, mean = 20 years ± 1.5, 70 female, 42 male, 4 prefer not to answer). The study was approved by the Princeton University Institutional Review Board and participants were recruited from a pool maintained by the Princeton Psychology Department. Undergraduate students were compensated with 2.5 psychology course credit hours and a performance bonus up to $10 in the form of an Amazon gift card (bonus M = $7.68, SD = 0.61, range = $4.41 - 8.35). The conversion of apples to money was 0.11 cents per apple.

N-Back working memory task The N-Back task was performed as part of foraging task during travel between trees. In the N-Back task letters are displayed on screen in a sequence. Participants judge whether the stimulus that is currently on the screen matches the stimulus they saw a number of screens back (N-Back). On every trial participants responded via keypress whether the letter was a match ("s" key) or non-match ("d" key) to the letter on the previous screen (1-Back case) or three screens before (3-Back case). A trial began with a fixation cross (for 250 milliseconds) followed by the letter on screen (for 500 milliseconds) followed by a blank screen (for 950 milliseconds, total trial duration = 1.7 seconds). During a foraging travel bog 10 letters were presented, of which, 2 or 3 were targets (letter matches letter N-Back) and 2 or 3 were lures (matches current letter but not in position N-Back). The number of targets and lures were selected randomly each time an N-Back stimulus sequence was generated. We only used consonants to prevent participants from using mnemonics (letters were: 'B', 'C', 'D', 'F', 'G', 'H', 'J', 'K', 'M', 'N', 'P', 'Q', 'R', 'S', 'T', 'V', 'W', 'X', 'Y', 'Z') and half of the letters were presented in upper case and the other half lower case to prevent participants using iconic memory. (Cohen et al., 1994)

N-Back working memory task training We trained the N-Back task extensively to try to bring participants to highest possible levels of performance and minimize automaticity differences (in which some participants would have more experience with the N-Back or similar tasks, making the task less effortful for them compared to someone with little experience). Participants had to reach a certain performance criterion to move on from training. After being instructed on the task participants began practice for one of the effort levels (counterbalanced). First they completed two extended blocks (50 trials with a self-paced break up to 45 seconds between) with feedback about error type (types of feedback: "non-match", "missed match", "no response", displayed in red font for 800 ms after the trial). Then they performed one extended block without any feedback (50 trials).

We tasked participants with completing a certain number of mini-blocks with high accuracy to begin the foraging task. We did so to establish the expectation that participants had to exert effort when they chose to travel while foraging. A mini-block was a success when they saw no error feedback (large

Parameter	Mean	Lower bound	Upper bound
Inverse Temperature	0.425	0.383	0.469
Cognitive Low Effort Travel Cost	114.909	105.088	125.274
Cognitive Effort Cost	16.776	10.356	23.441
Physical Low Effort Travel Cost	124.295	112.371	136.350
Physical Effort Cost	15.943	8.915	23.007

Table S4: Experiment 2 (N-Back) Parameter posterior distribution values. Table includes the mean of the group-level posterior distribution and the upper and lower bounds (95% HDI).

black dot), and they were told they were moving on to the next mini-block. The error feedback was displayed when participants made two consecutive errors (including omission errors). If they did see one or more error feedback symbols they had to repeat that mini-block. They had to successfully complete 8 mini-blocks of the 1-Back task, and 12 mini-blocks of the 3-Back task. This training also ensured that participants could adequately perform the task. Participants had self-paced breaks in between all mini-blocks (up to 60 seconds).

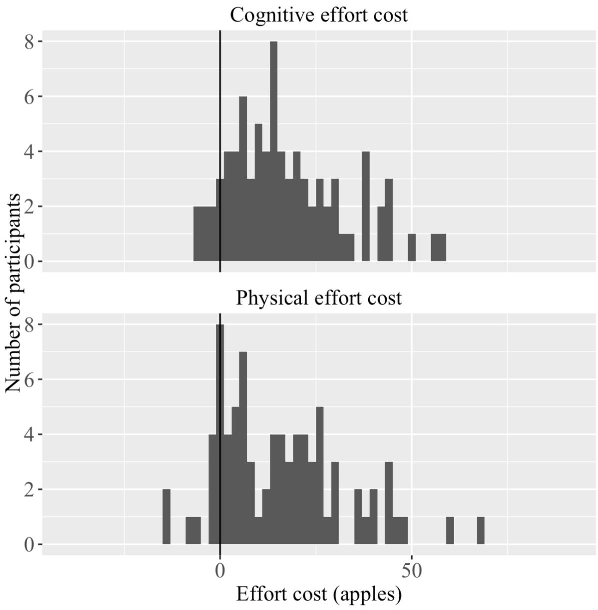

Figure S.4: Individual differences in effort costs Experiment 2. There were individual differences in 3-Back travel cost. Some participants were insensitive to the manipulation (3-Back cost near zero).

Explicit awareness of effort avoidance In Experiment 2 (N-Back), we asked participants if, and how, the required travel task changed their decision to travel to a new tree. For the cognitive (N-Back) variant, of those who completed the debrief survey (N = 113), 36% of participants (N = 41) reported changing their behavior based on the travel task (saying in their own words that they avoided the high cognitive effort [3-Back] task and stayed longer at a tree), whereas 64% of participants (N = 72) explicitly stated that the travel task did not change their decisions. This supports the idea that, for the majority of participants, the task is an indirect measure (i.e., participants whose behavior was influenced by the travel task were not aware of doing so).

Overall thresholds by experiment. We computed the group mean overall exit thresholds separately in Experiments 1 and 2 using mixed effect linear regression using only an intercept term. We compared these observed group overall thresholds to best thresholds from simulations. The group level mean overall threshold was 6.30 apples in Experiment 1 (SE = 0.12, df = 535.87, t = 54.36 , $p < 0.001$) which was close to the best policy in simulation of 6.78 apples (see Fig. S.5, and Table S1). The group level mean overall threshold was 4.02 apples in Experiment 2 (SE = 0.22, df = 80.22, t = 18.47, $p < 0.001$) which was similar to the best policy in simulation of 4.65 apples.

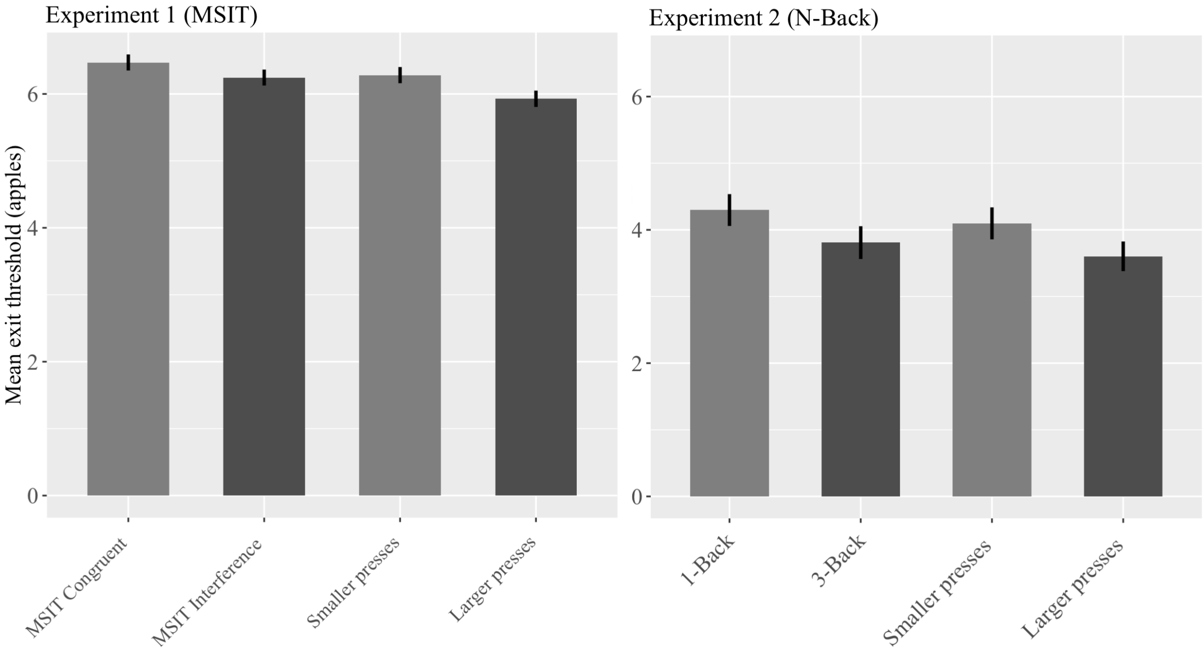

Figure S.5: Group level exit thresholds by Experiment. Left plot: Experiment 1, right plot: Experiment 1. x-axis: Foraging conditions indicated by the travel task required, y-axis: Group-level mean exit threshold (apples), error bars indicate SEM. Effort level indicated by bar color (light gray = low effort, dark gray = high effort). Group average thresholds near best threshold from simulation (best threshold respect to reward rate was 6.78 apples in Experiment 1 (MSIT) and 4.65 apples in Experiment 2 (N-Back)).

Relationship between cognitive and physical effort costs In the larger online sample we found a significant positive correlation between cognitive and physical effort costs (Experiment 1 (MSIT), N = 537, correlation = 0.55). In the smaller sample we did not find conclusive evidence for or against the correlation (Experiment 2 (N-Back), N = 81, wide highest density interval). Encouragingly the credible interval for the Experiment 2 (N-Back) model overlapped with the credible interval in Experiment 1 (MSIT). We speculated this may have to do with the Experiment 2 (N-Back) sample being less representative of the general population. We found that the Undergraduate students had higher Need for Cognition Scores compared to the Prolific sample. Princeton University students may have been admitted for being exceptionally cognitive effort seeking ("intellectuals") or for being exceptionally physical effort seeking ("athletes"). Indeed Princeton University has a large percentage of student athletes, although we did not collect data on how many participants were student athletes. Perhaps in these students there is a decoupling of effort-based decision making in the cognitive and physical domain that is not typical of the general population. Consistent with this account we observed many more participants who were indifferent to one type of effort (effort cost near zero) but had high effort costs for the other domain (see spreading along the x- and y-axes in the Experiment 2 (N-Back) but not Experiment 1 (MSIT) sample Fig. S.6). Experiment 1 (MSIT) participants were also likely more representative of the general population in that there was a wider age range and many more participants.

We directly fit the correlation between cognitive and physical high effort costs via a covariance matrix of fixed-effects parameters. We did not find a reliable correlation in Experiment 2 (N-Back) (mean correlation = 0.048, 95% HDI = -0.369 - 0.462). One explanation is that we were under powered to detect a correlation (N = 81 in Experiment 2 (N-Back)). Indeed the highest density interval is very wide in Experiment 2 (N-Back) model, and the 80% highest density interval does overlap with the posterior distribution in Experiment 1 (MSIT) model.

5.6.2 Experiment 3 (Richness)

Methods In Experiment 3 (Richness) we conducted a study manipulating the tree richness as a benchmark of how participants adjust their exit threshold in response to reward rate (richness was not manipulated in Experiments 1 and 2). We compared two levels of reward richness by adjusting the mean

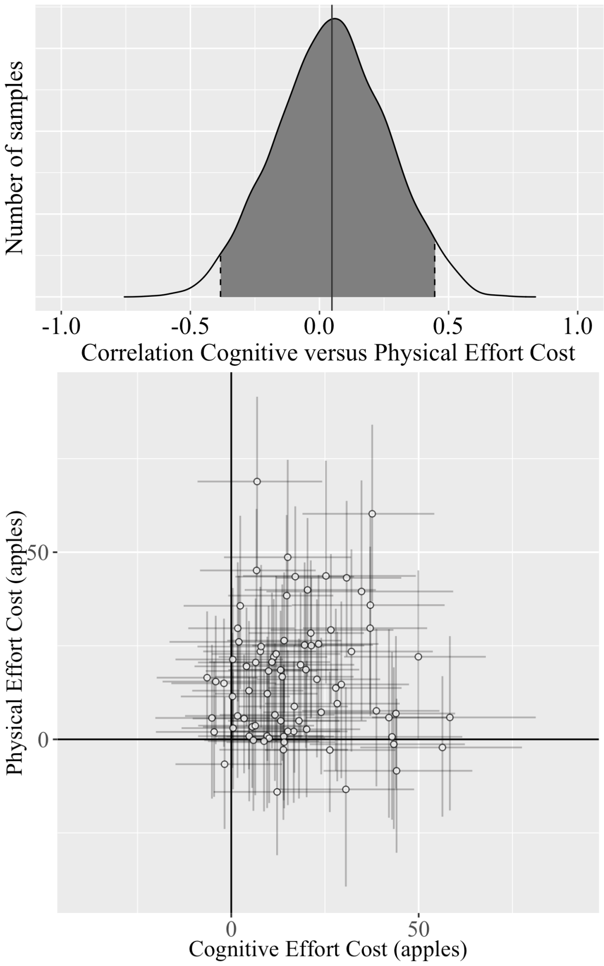

Figure S.6: Experiment 2 (N-Back) Cognitive versus Physical Effort Correlation. Row 1: posterior distribution of correlation between high effort cost for cognitive and physical effort. Cognitive and physical effort costs are not correlated in Experiment 2 sample (95% HDI overlapping with zero) wide confidence interval suggests sample is under-powered. Row 2: x-axis: Individual differences in Cognitive Effort Costs, y-axis: Individual differences in Physical Effort Costs. Error bars indicate 80% HDI.

of a normal distribution used to draw the initial reward paid out by a tree. In the 'scarce condition' the initial reward mean was 15 apples $N(15, 1)$ and in the 'rich condition' the initial reward mean was 20 apples $N(20, 1)$. We tested all combinations of the effort and richness orchard types and counterbalanced block order within effort type. We predicted participants would lower their threshold (exit later) in the scarce condition because reward rate is lower in the scarce compared to the rich condition. This would confirm that participants still adhere to predictions of the Marginal Value Theorem even in our novel experiment context where effort was added to the travel.

The richness manipulation was conducted during piloting studies of physical effort version of the Effort Foraging Task. There were several differences between the pilot studies (Experiment 3) and the main experiments (1 and 2). Pilot studies were conducted in the laboratory (rather than remotely). For pilot studies we pre-screened participants to have relatively low Need for Cognition (we did not do so in the main experiments). The pre-screen survey was completed online no later than 24 hours before the study. Participants gave written consent to complete the pre-screen. To avoid explicit cueing of the objective of the study we administered two foil self-report scales following the Need for Cognition scale; the Individualism and Collectivism Scale (Triandis & Gelfand, 1998) and the Ambiguity Tolerance Scale (Mac Donald, 1970). Participants with Need for Cognition scores less than or equal to 70 points (out of 90 possible points) were invited to the study. Participants again gave written consent to participate in the study.

43 participants volunteered for Experiment 3 (Richness) (24 female, 19 male, 18-34 years old, mean age = 21.5 years ± 3.7). Experiment 3 (Richness) includes two pilot studies in which richness was

manipulated. In the 'button-pressing rate' version, participants had to maintain a fixed rate (smaller number of presses per second vs. larger number of presses per second). In the 'button-press count' version, participants had to complete a smaller number or larger number of their maximum calibrated presses (this was the same physical effort requirement as in the main experiments). The cognitive effort requirement in both versions was the N-Back task (same as in Experiment 2 (N-Back)). There were 21 participants in button-pressing rate version and 19 participants in button-pressing count version. Three participants were excluded due to poor button pressing performance (> 2SD uncompleted presses). In both versions participants completed 4 N-Back blocks followed by 4 rapid button pressing blocks. Orchard duration was 5 minutes in the button-press rate version, and 7 minutes in the button-press count version. Harvest and travel time were the same for Experiments 1 (N-Back) and 3 (Richness).

To test whether participants responded as predicted to the richness manipulation, we fitted a mixed-effects linear regression model to exit thresholds (using the lme4 pacakge in the R language, Bates et al., 2022). The model predicted exit threshold (expected (log) apples) by orchard type separately fit for all conditions (for cognitive high and low effort, and physical high and low effort, and scarce and rich orchards) for all participants. Then we computed a multi degrees-of-freedom test on the linear mixed-effects model (using contestMD function of the lmerTest package (Kuznetsova et al., 2020)). The contrast tested whether the mean-value parameters are significantly different in the scarce compared to the rich condition (collapsing over all of the different travel tasks).

References

Agrawal, M., Mattar, M., Daw, N., & Cohen, J. D. (2019). Rational arbitration of hippocampal replay. *undefined*. Retrieved November 1, 2022, from https://www.semanticscholar.org/paper/Rational-Arbitration-of-Hippocampal-Replay-Agrawal-Mattar/b42f4eae895d2c7c6dd634492c7ea9cf433f5e2c

Agrawal, M., Mattar, M. G., Cohen, J. D., & Daw, N. D. (2022). The temporal dynamics of opportunity costs: A normative account of cognitive fatigue and boredom [Place: US Publisher: American Psychological Association]. *Psychological Review*, *129*, 564–585. https://doi.org/10.1037/rev0000309

Alloy, L. B., Abramson, L. Y., Walshaw, P. D., Cogswell, A., Grandin, L. D., Hughes, M. E., Iacoviello, B. M., Whitehouse, W. G., Urosevic, S., Nusslock, R., & Hogan, M. E. (2008). Behavioral approach system and behavioral inhibition system sensitivities and bipolar spectrum disorders: Prospective prediction of bipolar mood episodes. *Bipolar Disorders*, *10*(2), 310–322. https://doi.org/10.1111/j.1399-5618.2007.00547.x

Ang, Y.-S., Lockwood, P., Apps, M. A. J., Muhammed, K., & Husain, M. (2017). Distinct subtypes of apathy revealed by the apathy motivation index. *PloS One*, *12*(1), e0169938. https://doi.org/10.1371/journal.pone.0169938

Bates, D., Maechler, M., Bolker [aut, B., cre, Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., & simulate.formula), P. N. K. (c. o. (2022, February 5). *Lme4: Linear mixed-effects models using 'eigen' and s4* (Version 1.1-28). Retrieved March 19, 2022, from https://CRAN.R-project.org/package=lme4

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview [Place: US Publisher: Lawrence Erlbaum]. *Psychological Inquiry*, *7*(1), 1–15. https://doi.org/10.1207/s15327965pli0701_1

Borderies, N., Bornert, P., Gilardeau, S., & Bouret, S. (2020). Pharmacological evidence for the implication of noradrenaline in effort. *PLoS Biology*, *18*(10), e3000793. https://doi.org/10.1371/journal.pbio.3000793

Bornert, P., & Bouret, S. (2021). Locus coeruleus neurons encode the subjective difficulty of triggering and executing actions [Publisher: Public Library of Science]. *PLOS Biology*, *19*(12), e3001487. https://doi.org/10.1371/journal.pbio.3001487

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113. https://doi.org/10.1146/annurev-psych-010814-015044

Bush, G., & Shin, L. M. (2006). The multi-source interference task: An fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network. *Nature Protocols*, *1*(1), 308–313. https://doi.org/10.1038/nprot.2006.48

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition [Place: US Publisher: American Psychological Association]. *Journal of Personality and Social Psychology*, *42*(1), 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition [Publisher: Routledge _eprint: https://doi.org/10.1207/s15327752jpa4803_13]. *Journal of Personality Assessment*, *48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Carter, E. C., Pedersen, E. J., & McCullough, M. E. (2015). Reassessing intertemporal choice: Human decision-making is more optimal in a foraging task than in a self-control task. *Frontiers in Psychology*, *6*, 95. https://doi.org/10.3389/fpsyg.2015.00095

Carter, E. C., & Redish, A. D. (2016). Rats value time differently on equivalent foraging and delay-discounting tasks [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology: General*, *145*(9), 1093–1101. https://doi.org/10.1037/xge0000196

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales [Place: US Publisher: American Psychological Association]. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J. F., Bruce, B., & Rose, M. (2007). The patient-reported outcomes measurement information system (PROMIS). *Medical care*, *45*(5), S3–S11. https://doi.org/10.1097/01.mlr.0000258615.42478.55

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Chevalier, N. (2018). Willing to think hard? the subjective value of cognitive effort in children [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cdev.12805]. *Child Development*, *89*(4), 1283–1295. https://doi.org/10.1111/cdev.12805

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of Neuroscience*, *29*(39), 12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009

Chong, T., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS biology*, *15*(2), e1002598. https://doi.org/10.1371/journal.pbio.1002598

Chong, T., Bonnelle, V., & Husain, M. (2016, January 1). Chapter 4 - quantifying motivation with effort-based decision-making paradigms in health and disease. In B. Studer & S. Knecht (Eds.), *Progress in brain research* (pp. 71–100). Elsevier. https://doi.org/10.1016/bs.pbr.2016.05.002

Cohen, J. D., Forman, S. D., Braver, T. S., Casey, B. J., Servan-Schreiber, D., & Noll, D. C. (1994). Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Human Brain Mapping*, *1*(4), 293–304. https://doi.org/10.1002/hbm.460010407

Constantino, S. M., Dalrymple, J., Gilbert, R. W., Varanese, S., Di Rocco, A., & Daw, N. D. (2017, August 8). *A neural mechanism for the opportunity cost of time* (preprint). Neuroscience. https://doi.org/10.1101/173443

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective & Behavioral Neuroscience*, *15*(4), 837–853. https://doi.org/10.3758/s13415-015-0350-y

Costumero, V., Barrós-Loscertales, A., Fuentes, P., Rosell-Negre, P., Bustamante, J. C., & Ávila, C. (2016). BAS-drive trait modulates dorsomedial striatum activity during reward response-outcome associations. *Brain Imaging and Behavior*, *10*(3), 869–879. https://doi.org/10.1007/s11682-015-9466-5

Crawford, J. L., Eisenstein, S. A., Peelle, J. E., & Braver, T. S. (2022). Domain-general cognitive motivation: Evidence from economic decision-making – final registered report. *Cognitive Research: Principles and Implications*, *7*(1), 23. https://doi.org/10.1186/s41235-022-00363-z

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Evans, D. E., & Rothbart, M. K. (2007). Developing a model for adult temperament. *Journal of Research in Personality*, *41*(4), 868–888. https://doi.org/10.1016/j.jrp.2006.11.002

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

Geana, A., Wilson, R., Daw, N. D., & Cohen, J. (2016). Information-seeking, learning and the marginal value theorem: A normative approach to adaptive exploration. *CogSci*.

Giboin, L.-S., & Wolff, W. (2019). The effect of ego depletion or mental fatigue on subsequent physical endurance performance: A meta-analysis. *Performance Enhancement & Health*, *7*(1), 100150. https://doi.org/10.1016/j.peh.2019.100150

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, *5*, e11305. https://doi.org/10.7554/eLife.11305

González, I., & Déjean, S. (2021, March 1). *CCA: Canonical correlation analysis* (Version 1.2.1). Retrieved February 23, 2022, from https://CRAN.R-project.org/package=CCA

Grahek, I., Everaert, J., Krebs, R. M., & Koster, E. H. W. (2018). Cognitive control in depression: Toward clinical models informed by cognitive neuroscience [Publisher: SAGE Publications Inc]. *Clinical Psychological Science*, *6*(4), 464–480. https://doi.org/10.1177/2167702618758969

Grahek, I., Shenhav, A., Musslick, S., Krebs, R. M., & Koster, E. H. W. (2019). Motivation and cognitive control in depression. *Neuroscience & Biobehavioral Reviews*, *102*, 371–381. https://doi.org/10.1016/j.neubiorev.2019.04.011

Hayden, B. Y. (2018). Economic choice: The foraging perspective. *Current Opinion in Behavioral Sciences*, *24*, 1–6. https://doi.org/10.1016/j.cobeha.2017.12.002

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–939. https://doi.org/10.1038/nn.2856

Hershenberg, R., Satterthwaite, T. D., Daldal, A., Katchmar, N., Moore, T. M., Kable, J. W., & Wolf, D. H. (2016). Diminished effort on a progressive ratio task in both unipolar and bipolar depression. *Journal of Affective Disorders*, *196*, 97–100. https://doi.org/10.1016/j.jad.2016.02.003

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

Kane, G. A., James, M. H., Shenhav, A., Daw, N. D., Cohen, J. D., & Aston-Jones, G. (2021, September 26). *Rat anterior cingulate cortex continuously signals decision variables in a patch foraging task* [Section: New Results Type: article]. bioRxiv. https://doi.org/10.1101/2021.06.07.447464

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression [Place: US Publisher: American Psychological Association]. *Journal of Abnormal Psychology*, *111*(4), 589–597. https://doi.org/10.1037/0021-843X.111.4.589

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural mechanisms of foraging. *Science (New York, N.Y.)*, *336*(6077), 95–98. https://doi.org/10.1126/science.1216930

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology. General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Krönke, K.-M., Mohr, H., Wolff, M., Kräplin, A., Smolka, M. N., Bühringer, G., Ruge, H., & Goschke, T. (2021). Real-life self-control is predicted by parietal activity during preference decision making: A brain decoding analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *21*(5), 936–947. https://doi.org/10.3758/s13415-021-00913-w

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *The Behavioral and brain sciences*, *36*(6), 10.1017/S0140525X12003196. https://doi.org/10.1017/S0140525X12003196

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & Jensen, S. P. (2020, October 23). *lmerTest: Tests in linear mixed effects models* (Version 3.1-3). Retrieved February 25, 2022, from https://CRAN.R-project.org/package=lmerTest

Le Heron, C., Apps., M. A. J., & Husain, M. (2018). The anatomy of apathy: A neurocognitive framework for amotivated behaviour. *Neuropsychologia*, *118*, 54–67. https://doi.org/10.1016/j.neuropsychologia.2017.07.003

Le Heron, C., Holroyd, C. B., Salamone, J., & Husain, M. (2019). Brain mechanisms underlying apathy. *Journal of Neurology, Neurosurgery, and Psychiatry*, *90*(3), 302–312. https://doi.org/10.1136/jnnp-2018-318265

Le Heron, C., Kolling, N., Plant, O., Kienast, A., Janska, R., Ang, Y.-S., Fallon, S., Husain, M., & Apps, M. A. J. (2020). Dopamine modulates dynamic decision-making during foraging [Place: US Publisher: Society for Neuroscience]. *The Journal of Neuroscience*, *40*(27), 5273–5282. https://doi.org/10.1523/JNEUROSCI.2586-19.2020

Le Heron, C., Plant, O., Manohar, S., Ang, Y.-S., Jackson, M., Lennox, G., Hu, M. T., & Husain, M. (2018). Distinct effects of apathy and dopamine on effort-based decision-making in parkinson's disease. *Brain*, *141*(5), 1455–1469. https://doi.org/10.1093/brain/awy110

Lenow, J. K., Constantino, S. M., Daw, N. D., & Phelps, E. A. (2017). Chronic and acute stress promote overexploitation in serial decision making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(23), 5681–5689. https://doi.org/10.1523/JNEUROSCI.3618-16.2017

Levy, D. J., & Glimcher, P. W. (2011). Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *The Journal of Neuroscience*, *31*(41), 14693–14707. https://doi.org/10.1523/JNEUROSCI.2218-11.2011

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current opinion in neurobiology*, *22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Lopez-Gamundi, P., & Wardle, M. C. (2018). The cognitive effort expenditure for rewards task (c-EEfRT): A novel measure of willingness to expend cognitive effort. *Psychological Assessment*, *30*(9), 1237–1248. https://doi.org/10.1037/pas0000563

Lopez-Gamundi, P., Yao, Y.-W., Chong, T. T.-J., Heekeren, H. R., Mas-Herrero, E., & Marco-Pallarés, J. (2021). The neural basis of effort valuation: A meta-analysis of functional magnetic resonance imaging studies. *Neuroscience and Biobehavioral Reviews*, *131*, 1275–1287. https://doi.org/10.1016/j.neubiorev.2021.10.024

Mac Donald, A. P. (1970). Revised scale for ambiguity tolerance: Reliability and validity [Publisher: SAGE Publications Inc]. *Psychological Reports*, *26*(3), 791–798. https://doi.org/10.2466/pr0.1970.26.3.791

Marchetti, I., Shumake, J., Grahek, I., & Koster, E. H. W. (2018). Temperamental factors in remitted depression: The role of effortful control and attentional mechanisms. *Journal of Affective Disorders*, *235*, 499–505. https://doi.org/10.1016/j.jad.2018.04.064

Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, *106*(3), 857–864. https://doi.org/10.1152/japplphysiol.91324.2008

Marzilli Ericson, K. M., White, J. M., Laibson, D., & Cohen, J. D. (2015). Money earlier or later? simple heuristics explain intertemporal choices better than delay discounting does [Publisher: SAGE Publications Inc]. *Psychological Science*, *26*(6), 826–833. https://doi.org/10.1177/0956797615572232

McFarland, B. R., Shankman, S. A., Tenke, C. E., Bruder, G. E., & Klein, D. N. (2006). Behavioral activation system deficits predict the six-month course of depression. *Journal of Affective Disorders*, *91*(2), 229–234. https://doi.org/10.1016/j.jad.2006.01.012

Moutoussis, M., Garzón, B., Neufeld, S., Bach, D. R., Rigoli, F., Goodyer, I., Bullmore, E., NSPN Consortium, Guitart-Masip, M., & Dolan, R. J. (2021). Decision-making ability, psychopathology, and brain connectivity. *Neuron*, *109*(12), 2025–2040.e7. https://doi.org/10.1016/j.neuron.2021.04.019

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current directions in psychological science*, *21*(2), 71–77. https://doi.org/10.1177/0963721412436807

Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control [Publisher: Elsevier]. *Trends in Cognitive Sciences*, *25*(9), 757–775. https://doi.org/10.1016/j.tics.2021.06.001

Musslick, S., Cohen, J. D., & Shenhav, A. (2018). Estimating the costs of cognitive control from task performance: Theoretical validation and potential pitfalls, 6.

Niebaum, J., & Munakata, Y. (2020). Deciding what to do: Developments in children's spontaneous monitoring of cognitive demands [Place: United Kingdom Publisher: Wiley-Blackwell Publishing Ltd.]. *Child Development Perspectives*, *14*(4), 202–207. https://doi.org/10.1111/cdep.12383

Niebaum, J. C., Chevalier, N., Guild, R. M., & Munakata, Y. (2019). Adaptive control and the avoidance of cognitive control demands across development. *Neuropsychologia*, *123*, 152–158. https://doi.org/10.1016/j.neuropsychologia.2018.04.029

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

Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage*, *11*(5), 424–446. https://doi.org/10.1006/nimg.2000.0572

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *Journal of Cognitive Neuroscience*, *23*(11), 3419–3432. https://doi.org/10.1162/jocn_a_00011

Pagliaccio, D., Luking, K. R., Anokhin, A. P., Gotlib, I. H., Hayden, E. P., Olino, T. M., Peng, C.-Z., Hajcak, G., & Barch, D. M. (2016). Revising the BIS/BAS scale to study development: Measurement invariance and normative effects of age and sex from childhood through adulthood [Place: US Publisher: American Psychological Association]. *Psychological Assessment*, *28*(4), 429–442. https://doi.org/10.1037/pas0000186

Patzelt, E. H., Kool, W., Millner, A. J., & Gershman, S. J. (2019). The transdiagnostic structure of mental effort avoidance [Number: 1 Publisher: Nature Publishing Group]. *Scientific Reports*, *9*(1), 1689. https://doi.org/10.1038/s41598-018-37802-1

Pinto-Meza, A., Caseras, X., Soler, J., Puigdemont, D., Pérez, V., & Torrubia, R. (2006). Behavioural inhibition and behavioural activation systems in current and recovered major depression participants [Place: Netherlands Publisher: Elsevier Science]. *Personality and Individual Differences*, *40*(2), 215–226. https://doi.org/10.1016/j.paid.2005.06.021

Quilty, L. C., Mackew, L., & Bagby, R. M. (2014). Distinct profiles of behavioral inhibition and activation system sensitivity in unipolar vs. bipolar mood disorders. *Psychiatry Research*, *219*(1), 228–231. https://doi.org/10.1016/j.psychres.2014.05.007

Raio, C. M., Biernacki, K., Kapoor, A., Wengler, K., Bonagura, D., Xue, J., Constantino, S. M., Horga, G., & Konova, A. B. (2022, March 27). Suboptimal foraging decisions and involvement of the ventral tegmental area in human opioid addiction [Pages: 2022.03.24.485654 Section: New Results]. https://doi.org/10.1101/2022.03.24.485654

Salamone, J. D., Correa, M., Yang, J.-H., Rotolo, R., & Presby, R. (2018). Dopamine, effort-based choice, and behavioral economics: Basic and translational research. *Frontiers in Behavioral Neuroscience*, *12*. Retrieved March 6, 2022, from https://www.frontiersin.org/article/10.3389/fnbeh.2018.00052

Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., & Pessiglione, M. (2012). Neural mechanisms underlying motivation of mental versus physical effort [Publisher: Public Library of Science]. *PLOS Biology*, *10*(2), e1001266. https://doi.org/10.1371/journal.pbio.1001266

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

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 [_eprint: https://doi.org/10.1146/annurev-neuro-072116-031526]. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the snaith-hamilton pleasure scale. *The British Journal of Psychiatry: The Journal of Mental Science*, *167*(1), 99–103. https://doi.org/10.1192/bjp.167.1.99

Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition*, *116*(2), 155–167. https://doi.org/10.1016/j.cognition.2010.04.007

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Stan. (2021). *CmdStanR: The r interface to CmdStan — cmdstanr-package*. Retrieved November 2, 2022, from https://mc-stan.org/cmdstanr/reference/cmdstanr-package.html

Strobel, A., Wieder, G., Paulus, P. C., Ott, F., Pannasch, S., Kiebel, S. J., & Kührt, C. (2020). Dispositional cognitive effort investment and behavioral demand avoidance: Are they related? [Publisher: Public Library of Science]. *PLOS ONE*, *15*(10), e0239817. https://doi.org/10.1371/journal.pone.0239817

Toro-Serey, C., Kane, G. A., & McGuire, J. T. (2021). Choices favoring cognitive effort in a foraging environment decrease when multiple forms of effort and delay are interleaved. *Cognitive, Affective, & Behavioral Neuroscience*. https://doi.org/10.3758/s13415-021-00972-z

Tran, T., Hagen, A. E. F., Hollenstein, T., & Bowie, C. R. (2020). Physical- and cognitive-effort-based decision-making in depression: Relationships to symptoms and functioning: [Publisher: SAGE PublicationsSage CA: Los Angeles, CA]. *Clinical Psychological Science*. https://doi.org/10.1177/2167702620949236

Tran, T., Hagen, A. E. F., Hollenstein, T., & Bowie, C. R. (2021). Physical- and cognitive-effort-based decision-making in depression: Relationships to symptoms and functioning [Publisher: SAGE Publications Inc]. *Clinical Psychological Science*, *9*(1), 53–67. https://doi.org/10.1177/2167702620949236

Treadway, M. T., Buckholtz, J. W., Cowan, R. L., Woodward, N. D., Li, R., Ansari, M. S., Baldwin, R. M., Schwartzman, A. N., Kessler, R. M., & Zald, D. H. (2012). Dopaminergic mechanisms of individual differences in human effort-based decision-making. *The Journal of Neuroscience*, *32*(18), 6170–6176. https://doi.org/10.1523/JNEUROSCI.6459-11.2012

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 [Publisher: Public Library of Science]. *PLOS ONE*, *4*(8), e6598. https://doi.org/10.1371/journal.pone.0006598

Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism [Place: US Publisher: American Psychological Association]. *Journal of Personality and Social Psychology*, *74*(1), 118–128. https://doi.org/10.1037/0022-3514.74.1.118

Walton, M. E., Rudebeck, P. H., Bannerman, D. M., & Rushworth, M. F. S. (2007). Calculating the cost of acting in frontal cortex. *Annals of the New York Academy of Sciences*, *1104*, 340–356. https://doi.org/10.1196/annals.1390.009

Wang, H.-T., Smallwood, J., Mourao-Miranda, J., Xia, C. H., Satterthwaite, T. D., Bassett, D. S., & Bzdok, D. (2020). Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. *NeuroImage*, *216*, 116745. https://doi.org/10.1016/j.neuroimage.2020.116745

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

Westbrook, A., Yang, X., Bylsma, L. M., Daches, S., George, C. J., Seidman, A. J., Jennings, J. R., & Kovacs, M. (2022). Economic choice and heart rate fractal scaling indicate that cognitive effort is reduced by depression and boosted by sad mood. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2022.07.008

Wilks, S. S. (1935). On the independence of k sets of normally distributed statistical variables [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *3*(3), 309–326. https://doi.org/10.2307/1905324

Zorowitz, S., & Bennett, D. (2022). *NivTurk* (Version (v1.2-prolific)). Zenodo. https://doi.org/10.5281/zenodo.6609218

Zorowitz, S., Niv, Y., & Bennett, D. (2021, April 12). Inattentive responding can induce spurious associations between task behavior and symptom measures. https://doi.org/10.31234/osf.io/rynhk