medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Major depression symptom severity associations with willingness to exert effort and patch foraging strategy

Laura A. Bustamante, PhD[1,2] Deanna M. Barch, PhD [1]
Johanne Solis[3] Temitope Oshinowo[2] Ivan Grahek, PhD [4]
Anna B. Konova, PhD[3] Nathaniel D. Daw, PhD[2]
Jonathan D. Cohen, MD, PhD [2]

[1]Washington University in St. Louis
[2]Princeton University
[3]Rutgers University
[4]Brown University
Corresponding author:
Laura Ana Bustamante
bustamante@wustl.edu

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Abstract

Individuals with major depressive disorder (MDD) can experience reduced motivation and cognitive function, leading to challenges with goal-directed behavior. When selecting goals, people maximize 'expected value' by selecting actions that maximize potential reward while minimizing associated costs, including effort 'costs' and the opportunity cost of time. In MDD, differential weighing of costs and benefits are theorized mechanisms underlying changes in goal-directed cognition and may contribute to symptom heterogeneity. We used the Effort Foraging Task to quantify cognitive and physical effort costs, and patch leaving thresholds in low effort conditions (hypothesized to reflect perceived opportunity cost of time) and investigated their shared versus distinct relationships to clinical features in participants with MDD (N=52, 43 in-episode) and comparisons (N=27). Contrary to our predictions, none of the decision-making measures differed with MDD diagnosis. However, each of the measures were related to symptom severity, over and above effects of ability (i.e., performance). Greater anxiety symptoms were selectively associated with *lower* cognitive effort cost (i.e. greater willingness to exert effort). Anhedonia symptoms were associated with increased physical effort costs. Finally, greater physical anergia was related to decreased patch leaving thresholds. Markers of effort-based decision-making may inform understanding of MDD heterogeneity. Increased willingness to exert cognitive effort may contribute to anxiety symptoms such as rumination and worry. The association of decreased leaving thresholds with symptom severity is consistent with reward rate-based accounts of reduced vigor in MDD. Future research should address subtypes of depression with or without anxiety, which may relate differentially to cognitive effort decisions.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Introduction

Goal-directed behavior in major depressive disorder (MDD)

Individuals with major depressive disorder (MDD) can experience challenges with goal-directed behavior, including reduced motivation due to symptoms such as apathy, anergia, and anhedonia, as well as reduced cognitive function. Decisions about which goals to pursue and which actions to take to achieve them, can be understood in terms of costs and benefits. People maximize 'expected value' by selecting actions that maximize potential reward while minimizing associated costs. Effort-based decision making involves minimizing cognitive and physical effort costs (56, 57, 51, 55, 67), as well as the opportunity cost of time (often also emphasized in value-based decision-making) (16).

Under effort- and value-based decision-making disruption accounts of MDD symptoms, changes in goal-directed behavior in depression can come from multiple causes, for example, differences in representing either the benefits or the costs of potential actions. The present study focuses on cognitive and physical effort costs, as well as opportunity costs (i.e., reward rate), to understand how differences in these components of goal-directed behavior relate to clinical features of MDD.

Both cognitive and physical effort-based decision-making have been reported to differ in MDD, though findings have been mixed. MDD has been associated with decreased willingness to exert cognitive effort relative to comparison groups in some studies (1, 65, 69) though not in others (4, 62). Willingness to exert physical effort has been found to be decreased in MDD relative to comparison groups in some studies (7, 14, 63, 65, 68, 70, 72), though not in others (12, 58, 62, 68, 71).

Dissociating between cognitive and physical effort costs

Both cognitive and physical effort-based decision-making appear to be associated with MDD features and may underlie certain MDD symptoms, though findings have been mixed. Importantly, MDD is highly heterogeneous in terms of variation in symptom domains and severity across individuals. This may contribute to inconsistent findings with respect to diagnostic group differences and associations with clinical features. MDD presentation encompasses many different symptoms, and decision-making mechanisms have many different components (including reward sensitivity, effort costs, task ability). Gaining traction on mechanistically informed treatments will require precise computational measures of decision-making to tease apart their specific relationships to precise symptom measures.

Initial studies measuring both effort types within-participants found differential relationships between cognitive and physical effort decisions and symptoms (62, 65), suggesting potential applications to characterizing MDD heterogeneity. Studies measuring each effort type separately have reported decreased willingness to exert physical effort associated with symptom severity (i.e., anhedonia) (58, 62, 70), while others reported no relationship to symptom severity (e.g., not related to depression, anhedonia, apathy, (1, 4, 30, 65). For cognitive effort, some studies report associations with symptom severity (i.e., global functioning) (62, 69) while others do not (e.g., not related to depression, anhedonia, apathy) (1, 4, 30, 65). It remains unclear which symptoms map onto which component decision processes, and how shared or distinct these mappings are between cognitive and physical effort.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Hypothesized symptom relationships

Multiple symptom MDD domains have been proposed to relate to value- and effort-based decision-making. Subjective reward rate, reflecting the opportunity cost of time, is proposed to drive the vigor of actions, represented via midbrain dopamine tone (44). Drawing on this, Huys, Daw, and Dayan (2015) (33) proposed that physical anergia and psychomotor slowing symptoms of depression may be caused by reduced subjective reward rate representations. Relatedly, reduced willingness to exert physical effort to obtain rewards has been proposed as a mechanism underlying anhedonia and apathy symptoms (32, 49, 17). Grahek and colleagues (2019) (27) proposed that changes in motivational processes, resulting in reduced willingness to exert cognitive effort, may underlie, in part, reduced cognitive function associated with MDD, challenging the standard assumption that reduced cognitive function reflects reduced cognitive control capacity (59, 53, 41). This is of particular importance because reduced cognitive function in MDD contributes to disability (34) and often does not improve with otherwise effective anti-depressant treatments (28, 54). By the cognitive effort-based decision-making account, interventions to improve cognitive function would focus on boosting motivation and target willingness to engage control, rather than cognitive control ability (e.g., computerized cognitive training) as suggested by the reduced capacity account.

The variable prevalence of these symptoms across studies may contribute to mixed findings. For example, reduced motivation may be minimal or absent in some individuals with MDD (43, 2). In addition, certain symptom domains of depression may show a differential relationship to effort relative to others. Anxiety (the most common MDD comorbidity) (35) symptoms such as rumination and worry may require cognitive effort (e.g., sampling for replay and planning (5) and anxiety has been related to increased effortful model-based planning (25). Anxiety has also been linked to increased cognitive effort exertion to maintain performance in the face of increased attentional demands posed by threat-related stimuli (20). Additionally, social anxiety may be associated with enhanced motivation to exert cognitive effort in social contexts (such as a psychology experiment) (31). It therefore may be important to account for anxiety heterogeneity and relate cognitive effort-based decision making to specific depression symptom expression profiles (or subtypes), as some have suggested (38, 23).

Experiment overview

The goal of the current study was to quantify multiple components of effort-based decision-making and decompose to their contributions to MDD symptom expression profiles. Each component measured (cognitive and physical effort cost, cognitive task ability, subjective reward rate) has been proposed as an underlying mechanism of specific MDD symptoms. To test these accounts, we had participants who met diagnostic criteria for MDD (most in-episode) and demographically matched comparison participants with no psychiatric diagnoses complete the cognitive and physical Effort Foraging Task (10). To our knowledge, all previous MDD studies used explicit tasks in which participants choose between low-effort/low-reward and high-effort/high-reward options. We hypothesized that the Effort Foraging Task, which measures effort avoidance more implicitly by inferring the cost of effort from foraging behavior, would be less contaminated by demand characteristics that bring about changes in what participants value about effort (e.g., try to please the experimenter) (46). Reduced control over demand characteristics in explicit tasks, in turn, may contribute to mixed findings. While we hypothesized the Effort Foraging Task would yield more valid findings regarding the relationship between MDD and willingness to exert effort, it could also reasonably be the

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

case that this task may tap into distinct aspects of effort-based decision making than the tasks used in the extant literature (e.g., implicit versus explicit processing).

In the Effort Foraging Task, participants choose between harvesting a depleting patch, or traveling to a new patch, which is costly in time and effort. Analyses focused on 'exit thresholds', the reward value at which the participant decided to exit the current patch. The exit threshold reveals the point of equivalence in the tradeoff between the cost of harvesting with diminishing rewards and the cost of traveling to a new patch, and this is captured by a foraging-theory model (the Marginal Value Theorem, MVT) (13). According to the MVT, exit thresholds should reflect subjective average reward rate, and this comports with human behavior (16, 15, 37). Based on this patch-leaving behavior, a foraging-theory-based computational model quantified individual differences in the 'cost' of effort. The longer a participant delayed leaving the patch in high versus low effort conditions (i.e., relatively lower the exit threshold) the larger their inferred effort avoidance. Exit thresholds in low effort blocks were used to assess overall foraging strategy. Lastly, we assessed effortful travel task ability using accuracy and reaction times. We aimed to tease apart the influences of each of these effort decision-making components on clinical features of MDD by examining (i) diagnostic group differences, (ii) associations with overall depression severity (using the HAMD total score), and (iii) associations with symptom severity.

Based on previous findings and theoretical work we predicted cognitive effort costs would be increased in the MDD group and related to cognitive function symptoms (i.e., subjective cognitive complaints relative to baseline) (26, 27). Based on previous findings with explicit tasks, we predicted physical effort cost would be increased in the MDD group and related to anhedonia (58, 62, 70). Following Huys et al. (2015) (33) we hypothesized that the MDD group would differ in foraging strategy, exhibiting lower exit overall thresholds, and that this would relate to physical anergia/slowing.

Materials and Methods

Study overview

Participants

97 participants volunteered for the study and gave informed consent as approved by the Rutgers University Institutional Review Board (67 MDD, mean=26.9 years, SD=11.1, 18-61; 30 comparison, mean=27.1 years, SD=9.64, 19-59, further details in SI section 1). Groups were matched on key demographic variables (SI section 1.2). We oversampled MDD participants to maximize power to detect continuous symptom relationships to task behavior within this group. The comparison sample size was then selected to be adequate to detect group differences in behavior. Power analysis indicated that we could detect a medium effect size for group differences and symptom relationships with 80% power (SI section 2). All participants completed a detailed clinical assessment in session 1, but 7 MDD and 3 comparison participants opted not to return. Clinical symptom ratings and self-report analysis included data from the 60 MDD and 27 comparison participants who returned for the second (task) session. 50 MDD participants were currently depressed (43 with task data), 6 were in partial remission and 4 were in full remission (all with task data, details in SI section 3). 32 MDD participants used psychotropic medication while 28 did not.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Clinician ratings and self-reports

The Structured Clinical Interview for DSM-5 (22) confirmed assignment of MDD and/or comorbid anxiety diagnosis in the MDD group (and absence of exclusionary diagnoses in both groups, SI section 1), as well as whether participants were currently depressed (or in partial or full remission). Depression symptoms were assessed using the Hamilton Depression Rating Scale (HAMD) (29), Brief Psychiatric Rating Scale (47), MGH Cognitive and Physical Functioning Questionnaire (21), Patient Heath Questionnaire-9 (PHQ-9) (36), Generalized Anxiety Disorder-7 (60), Snaith–Hamilton Pleasure Scale (43), Apathy Motivation Index (2), Adult Temperament Questionnaire Effortful Control subscale (19), and Need for Cognition scale (11) (SI section 4, Table S1).

Symptom severity was measured using confirmatory factor analysis to combine clinician-rated and self-report measures of the following symptoms: anhedonia, anxiety, appetite symptoms, behavioral apathy, emotional apathy, social apathy, cognitive function symptoms, depressed mood/suicidality, and physical anergia/slowing, as well as for trait effortful control and need for cognition (see SI section 5,Table S2, and Table S3). Items were assigned based on what each scale was validated to measure. Assigned items were z-scored and averaged to compute a symptom severity score in the MDD group only. Items with an inter-item correlation below 0.2 were eliminated (multilevel package, item.total function, (9)). Internal consistency was computed using Cronbach's alpha (Table S2, ltm package, cronbach.alpha function, (52)). Factors with alpha$<$0.6 were excluded from further analysis (i.e., emotional apathy, appetite symptoms). The resulting items were then applied to compute confirmatory factor scores for (1) MDD only, which was the focus of our analysis, and (2) all participants to test generalizability of effects.

Effort Foraging Task

In the Effort Foraging Task participants harvested apples in virtual orchards (Figure S1, as described in Experiment 2 of Bustamante et al. (2023)) (10). On each foraging trial the participant visits a 'patch' which can be harvested to yield rewards (apples, converted to a monetary bonus). The marginal return decreases with each successive harvest. At any point the participant can travel to a new patch, which has replenished rewards, but it takes time and effort to travel there (Figure S2). The cognitive effort manipulation was the N-Back working memory task (1- and 3-Back levels, (45), also used in (62, 69)). The physical effort manipulation was rapid key pressing (also used in (64, 63, 70, 71, 68, 62, 7)) with the non-dominant pinky finger (50% or 100% of an individually calibrated maximum). Patches were presented blockwise in counterbalanced order (SI section 6). Blocks varied only in their (explicitly instructed) effort requirement (environment specifications in Table S4). Reaching a new patch was not dependent on performance. Participants had to reach a performance criterion during training to begin foraging.

We followed a subset of exclusions validated in Bustamante et al. (2023) that most impede estimates of effort costs (SI section 7). Participants were excluded if they missed the response deadline on many foraging trials (1 MDD participant excluded missed 49.5%) or had very few exit trials (1 MDD participant was excluded for 1 high effort physical and 3 high effort cognitive exits, 1 MDD participant was excluded from cognitive effort analyses for 2 cognitive high effort exits). The final sample included in behavioral analyses was 52 MDD participants (53 MDD participants for physical effort) and 27 comparison participants. We confirmed there were no diagnostic group differences among participants included in task-based analyses.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Marginal Value Theorem (MVT) model

The MVT predicts a forager should leave a patch when the instantaneous reward rate falls below the long-run average (13, 16). Travel costs were estimated using a hierarchical Bayesian logistic model. For each trial (t), the model compared the expected reward on the next harvest ($R_{e,t}$, Eq. (1)) against the condition-specific exit threshold ($\rho_{\text{condition}}$.

$$R_{e,t} = \frac{R_{e,t-1} + R_{e,t-2} \cdot \kappa}{2} \tag{1}$$

Using the difference of these to determine whether to harvest (1) or exit (0) via a softmax function (with inverse temperature, β, Eq. (2)).

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

The cost of travel in high effort blocks ($c_{\text{high effort}}$) was expressed as the marginal increase in cost of travel ($c_{\text{low effort}} + c_{\text{high effort}}$) from low effort, to control for any biases common to both conditions (e.g., variation in overall exit thresholds). Exit thresholds (ρ) were taken as fixed per-condition, determined by the total rewards ($\sum r$), total amount of time (number of harvest periods, T=condition duration/harvest time) and total travel costs ($\sum c$, sum over total times travelled in a condition) across all blocks of a condition.

where,

$$\rho_{\text{condition}} = \frac{\sum r - \sum c_{\text{condition}}}{T_{\text{condition}}} \tag{3}$$

Individual and group-level parameters were estimated using Markov Chain Monte Carlo sampling (using cmdstanr) (61). We used trace plots, $\hat{\text{R}}$ diagnostic statistics, and posterior predictive checks to assess model fit (SI section 8). We compared goodness of fit between diagnostic groups with an unpaired t-test on participants' log posterior likelihoods (SI section 9). We compared participants' overall exit thresholds to the best threshold found by simulation (SI section 10).

Analysis overview

All subsequent analyses used point estimates (mean) from participant-level effort costs and applied them in frequentist tests to control for potential confounding variables and conduct multiple-comparison corrections. All analyses were conducted in the R language (many using the stats package) (50). The HAMD total score was used to assess major depressive episode severity in the past week (herein 'overall depression'). We verified that results matched using self-reported depression (PHQ-9) due to concerns with HAMD validity (39, 24, 3). Our focus was on MDD symptom severity (z-scored for all analyses). To ensure remitted status was not driving key effects, we repeated all analyses zooming in on the current depressed group (excluding remitted participants). To test if effects were generalizable across the sample, we repeated analyses zooming out to all participants. We verified that no key results differed when controlling for psychotropic medication use (binary variable).

Diagnostic group differences

We tested for diagnostic group differences in cognitive or physical effort costs, controlling for high-effort task performance (3-Back D' or % larger number of presses completed), years of education, age, and BMI (body mass index, for physical effort) using linear regression.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

To confirm hierarchical shrinkage did not bias results, we also fit group differences directly within the MVT model (SI section 8).

Symptom associations with effort costs

Within the MDD group, we tested overall depression severity effects on cognitive or physical effort costs using linear regression, controlling for years of education, age, BMI (for physical effort) and high-effort task performance. Therefore, observed symptom associations are over and above effects of travel task ability. Next, we decomposed overall depression effects on effort costs into specific symptoms in a series of regression models. Because of mutual correlations between symptoms, we used multiple comparisons correction within a series of symptom models (FDR, 7 tests for each effort cost). We conducted a comparison of correlations to confirm specificity of observed relationships to effort type (cocor package) (18, 40). Lastly, we used Canonical Correlation Analysis (CCA) as an omnibus analysis of the relationships between task measures and symptoms, traits, and demographic variables (SI section 11).

Additional task measures

Participants may have differed in their ability to complete the required effort, which could confound decision-making differences and/or relate to symptoms. This is especially relevant for the cognitive task, which was not calibrated to individual ability. We addressed this by controlling for performance in symptom regression and CCA analyses. We also examined whether performance was associated with (i) diagnostic group, (ii) overall depression, and (iii) symptom domains (SI section 12).

Some depression symptoms are theorized to arise from reduced subjective reward rate representations (33), which decrease vigor (44). We tested whether overall exit thresholds (from low effort conditions, which were least confounded by effort sensitivity) were associated with (i) diagnostic group, (ii) overall depression, or (iii) symptom domains (SI section 12.2).

Results

Demographic and clinical characteristics

Diagnostic groups were matched on gender, race, age, parental education, household income, and childhood income (also within participants included in behavioral analyses,Table S5). The comparison group had more years of education than the MDD group (Table S5), so it was included as a covariate in all analyses. Depression severity varied widely in the MDD group (Figure S3). The MDD group scored higher on all symptom domains except emotional apathy. Need for cognition did not differ between groups, while effortful control was higher in comparisons (Figure S4).

Sensitivity to effort manipulations

On average, participants avoided effort (group-level posterior effort costs non-overlapping with zero, Table S6). The model converged ($\hat{R}<1.032$ for all parameters, Figure S5) and the observed probability of harvesting, across participants and trials, fell within the posterior predictive distribution ($p>0.768$). Simulated data recapitulated the empirical group-level change in exit threshold and overall exit threshold (Figure S6). We found no conclusive evidence for

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

or against a correlation between cognitive and physical effort costs (mean=0.048, 95% HDI=-0.369, 0.462, Table S6, discussed in the context of Experiment 2, Bustamante et al. (2023), in SI section 13).

Effort costs relationships to clinical features

Diagnostic group differences

We predicted effort costs by diagnostic group, controlling for high-effort task performance, education, age, and BMI (for physical effort). There were no group differences in either effort cost (Figure 1, cognitive: $p>0.66$, physical: $p>0.48$), even when controlling for psychotropic medication use (Table S7), and excluding remitted MDD participants (cognitive: $p>0.96$, physical: $p>0.67$). To ensure shrinkage in the hierarchical model did not obscure a group difference, we directly fitted diagnostic group differences for all MVT model parameters and found no differences, even when excluding remitted MDD participants (Table S8). There were no group differences in the model-agnostic measure of effort sensitivity (SI section 14, Figure S7). We computed the log posterior likelihoods per participant and found no significant difference between diagnostic groups, suggesting comparable goodness of fit (t=-0.59, df=53.54, $p>0.56$).

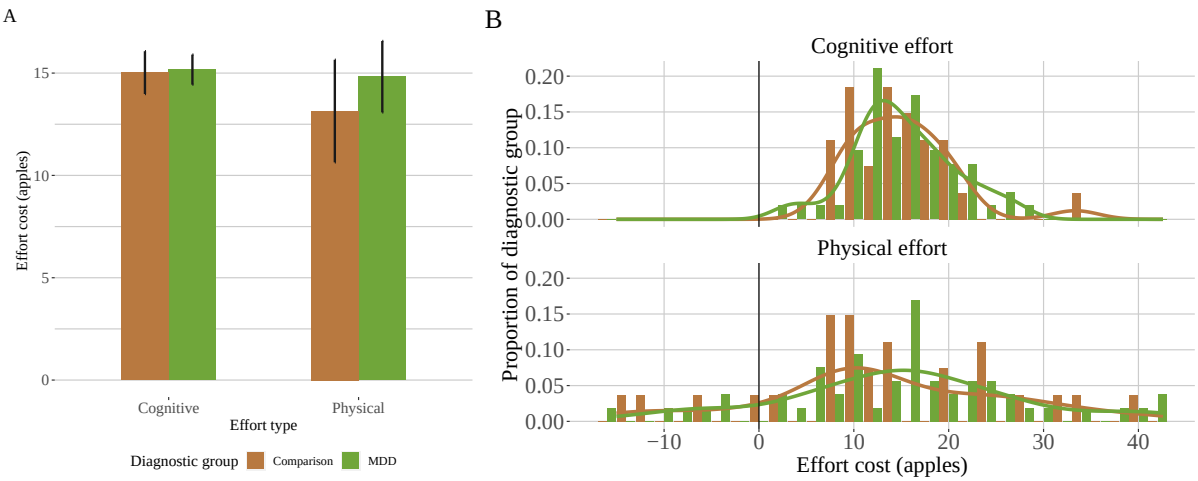

Figure 1: Effort cost by diagnostic group and effort type. Left panel: mean and standard error of the mean of individual differences in effort cost (y-axis) by effort type (x-axis). Right panel: individual differences histograms, x-axis indicates effort cost (larger values indicate more effort avoidance), y-axis indicates proportion of diagnostic group.

Overall depression severity

Surprisingly, overall depression severity was associated with *decreased* cognitive effort cost ($p<0.005$, Figure 2, Table 1). We found no reliable association with physical effort costs (Table 1) and the correlation magnitudes between the two types of effort were significantly difference (z=-2.25, p=0.024). Results were maintained when using self-reported depression (PHQ-9) as the overall severity measure (Table S9), when restricting the analyses to current MDD participants (Table S10), and when controlling for medication use (Table S7). Next, we identified which symptom domains contributed to the cognitive effort cost relationship to overall depression, and whether physical effort cost was related to any symptom domain.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Symptom specific relationships

We fitted regression models to estimate symptom domain relationships to effort costs while controlling for high-effort travel task performance, age, education, and BMI (for physical effort). Anxiety was related to *decreased* cognitive effort cost (Figure 2, Table 1), and this pattern was maintained in the current MDD group, across all participants (Table S10), and when controlling for medication use (Table S7). Decreased cognitive effort cost was associated with anhedonia, behavioral apathy, cognitive function symptoms, depressed mood/suicidality, and physical anergia/slowing in the MDD group. The result was unchanged when controlling for medication (Table S7). In the current MDD group and all participants the associations with anxiety, behavioral apathy, and depressed mood/suicidality were maintained, whereas anhedonia and cognitive function symptoms were not (Table S10). Physical anergia/slowing was maintained in the current MDD group but not in all participants. Because symptoms are mutually correlated, it is unclear if these are a cluster of symptoms that relate to cognitive effort costs or if a subset drove the relationship (see Figure S8 for heatmap of symptoms and task correlations).

Given our inclusion of participants with comorbid anxiety disorders and prior literature relating anxiety to increased effortful model-based strategy, we hypothesized that anxiety symptoms drove the overall depression association with cognitive effort cost. Anxiety was among the top symptoms predicting cognitive effort cost. We tested whether other symptoms were related to cognitive effort cost when controlling for anxiety and found no reliable relationships (although the behavioral apathy effect remained significant before multiple comparisons correction, Table 1).

We examined symptom associations with physical effort costs. Within the MDD group, anhedonia was associated with *increased* physical effort costs (Table 1). This effect was maintained (1) when controlling for medication (though not after FDR correction, Table S7), (2) in all participants and (3) the current MDD group only (Table S10). There was a significant difference in the correlation magnitudes of cognitive and physical effort cost with anxiety ($z=-2.47$, $p=0.014$) and anhedonia ($z=-3.25$, $p=0.001$) within the MDD group.

Canonical correlation analysis (CCA)

We used an omnibus analysis, CCA, to address the non-independence of symptom domains, while including all relevant covariates. Within the MDD group, we examined the relationship between task measures (i.e., N-Back accuracy, effort costs, and overall exit threshold) and symptom measures (simultaneously accounting for age, years of education, BMI, need for cognition, and self-reported effortful control). There was one significant dimension out of 5 (Dimension 1 correlation=0.688, Wilks statistic=0.16, F-approx=1.61, df1=60, df2=205, $p<0.008$). Dimension 1 captured many patterns from the symptom regression models, even when accounting for more individual differences (Table 2, Figure S9). For Dimension 1 task behavior coefficients, cognitive effort cost and 3-Back D' were positive, and physical effort cost and 1-Back D' were negative. For Dimension 1 symptom coefficients, anxiety, need for cognition, behavioral apathy, education, and social apathy were all negative. This pattern suggests a cluster of symptoms (i.e., anxiety, behavioral and social apathy) positively associated with physical effort cost, but negatively with cognitive effort cost. Including need for cognition, education, BMI, and N-Back accuracy maximized the correlation between task and symptom severity.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

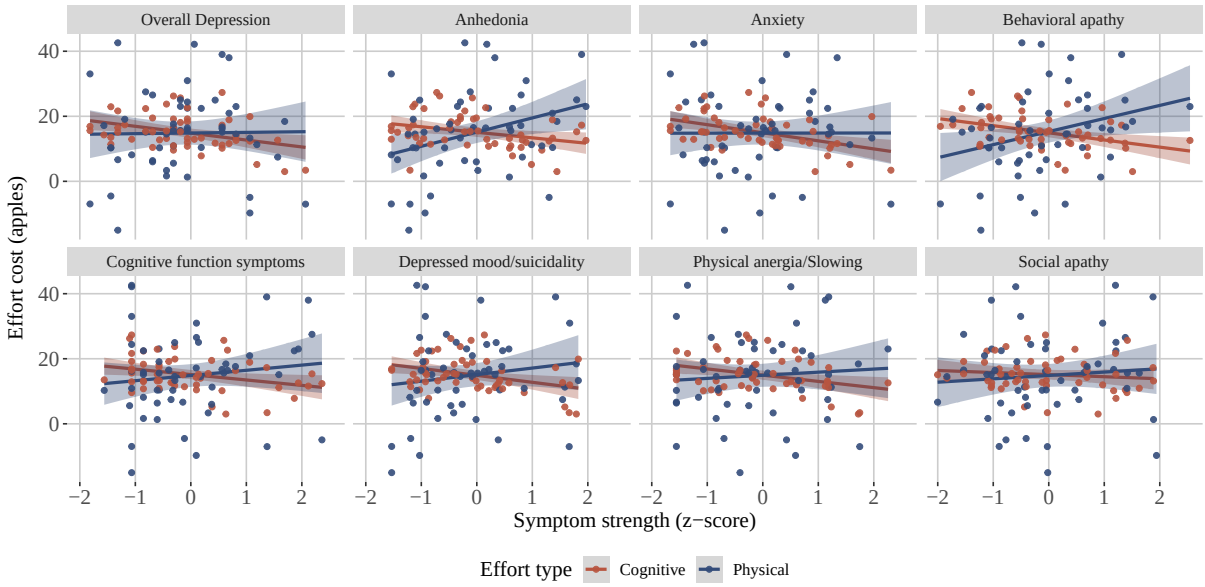

Figure 2: Effort costs relationships to individual MDD symptom domains. Blue indicates cognitive effort and red indicates physical effort. A: y-axes: effort costs from MVT model, x-axes: symptom severity (z-scores) for overall depression (Hamilton Rating Scale Total), anhedonia, anxiety, behavioral apathy, cognitive function symptoms, depressed mood/suicidality, physical anergia/slowing, and social apathy (MDD group only).

Additional task measures

Travel task performance

All the effort cost symptom analyses are over and above any effects related to travel task ability (i.e., task performance), which was controlled for. We examined diagnostic group differences and performance associations to symptoms directly (SI section 16). There were minimal diagnostic group differences in performance; however, the MDD group completed a lower percentage of keypresses in the high physical effort condition and responded faster on average on the cognitive task (Figure S10, Table S11). Neither cognitive nor physical performance was reliably related to overall depression. While anxiety symptoms were associated with cognitive effort costs, they were not associated with cognitive task performance (Figure S11, SI section 16). Anhedonia symptoms were related to a lower percentage of completed keypresses in the low (but not high) physical effort condition ($p<0.010$). Cognitive task performance did not predict cognitive effort cost, nor did physical performance predict physical effort cost, suggesting effort decisions and execution are dissociable in this task (Figure S11).

Overall exit threshold

We did not observe diagnostic group differences in overall exit thresholds in a model controlling for age and education ($p>0.269$). However, when medication was a covariate, medication use was associated with lower exit thresholds, and MDD group membership was associated with higher thresholds (Table S7). Overall exit thresholds were lower in participants with greater overall depression, and this association was significant within the current depressed group (Figure 3, Table 3). Lower overall exit thresholds were also associated with higher anx-

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Symptom	Estimate	SE	*t*	*p*	$p_{adjusted}$
A. Cognitive effort cost, MDD					
Overall depression	-0.31	0.10	-2.96	0.005*	
Anhedonia	-2.75	1.19	-2.32	0.025	0.035*
Anxiety	-3.72	0.97	-3.84	0.000	0.003*
Behavioral apathy	-3.92	1.28	-3.06	0.004	0.008*
Social apathy	-1.07	1.14	-0.94	0.351	0.351
Cognitive function symptoms	-2.10	0.97	-2.16	0.036	0.042*
Depressed mood/suicidality	-3.64	1.14	-3.19	0.003	0.008*
Physical anergia/slowing	-3.44	1.16	-2.96	0.005	0.008*
B. Physical effort cost, MDD					
Overall depression	0.15	0.27	0.58	0.565	
Anhedonia	8.38	2.90	2.89	0.006	0.042*
Anxiety	0.60	2.72	0.22	0.826	0.826
Behavioral apathy	7.59	3.25	2.33	0.024	0.084
Social apathy	1.76	2.76	0.64	0.527	0.615
Cognitive function symptoms	4.23	2.64	1.60	0.117	0.273
Depressed mood/suicidality	3.76	2.90	1.30	0.201	0.352
Physical anergia/slowing	2.95	3.24	0.91	0.368	0.515
C. Cognitive effort cost, MDD, controlling for anxiety					
Anhedonia	-1.82	1.11	-1.64	0.108	0.325
Behavioral apathy	-2.82	1.22	-2.31	0.025	0.152
Social apathy	-1.08	1.00	-1.07	0.288	0.346
Cognitive function symptoms	-0.01	1.14	-0.01	0.989	0.989
Depressed mood/suicidality	-1.71	1.38	-1.24	0.222	0.333
Physical anergia/slowing	-1.64	1.30	-1.26	0.214	0.333

Table 1: Symptom effort cost regressions (MDD group only). (A) Predicting cognitive effort cost from overall depression severity, and each symptom domain, controlling for cognitive task performance (3-Back D') years of education and age. (B) Predicting physical effort cost from overall depression severity, and each symptom domain, controlling for physical task performance (% larger number of presses completed), BMI, years of education and age. (C) Predicting cognitive effort cost from each symptom domain, controlling for anxiety, cognitive task performance (3-Back D') and age (* indicates $p<0.05$, FDR correction within symptom models).

iety, depressed mood/suicidality, and physical anergia/slowing. These results remained when controlling for medication (and adding anhedonia, Table S7). These symptom specific relationships were not maintained in the current depressed group (though anergia/slowing and depressed mood/suicidality were significant before FDR correction, Table S12) nor in all participants. Overall, this pattern is consistent with theories of reduced subjective reward rate representation associated with anergia and psychomotor slowing in depression (33).

Discussion

This cross-sectional study compared cognitive and physical effort-based decision-making using computational-model-derived parameters from the Effort Foraging Task (10), within-

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Task behavior (coefficient)	**Symptom** (coefficient)
↑Cognitive Effort Cost (0.88)	↓Anxiety (-0.44)
↓Physical Effort Cost (-0.61)	↓Need for cognition (-0.39)
↑3-Back D' (0.43)	↓Behavioral apathy (-0.39)
↓1-Back D' (-0.31)	↓Years education (-0.33)
	↑BMI (0.32)
	↓Social apathy (-0.27)

Table 2: Canonical Correlation Analysis Result Summary. Canonical Dimension 1 was the only significant dimension (correlation=0.713, p<*0.008*, MDD group only). Canonical coefficients maximize the correlation between task behavior (column 1) and symptoms (column 2, accounting for age, years of education, BMI, need for cognition and effortful control). Dimension 1 coefficients displayed in descending order by absolute value (arrows correspond to positive or negative coefficients). Coefficients below absolute value 0.25 are not displayed. Consistent with the symptom regression model results, CCA shows that cognitive effort cost was negatively correlated with symptom severity (including anxiety, and behavioral apathy), while physical effort cost was positively correlated with symptom severity (including anhedonia).

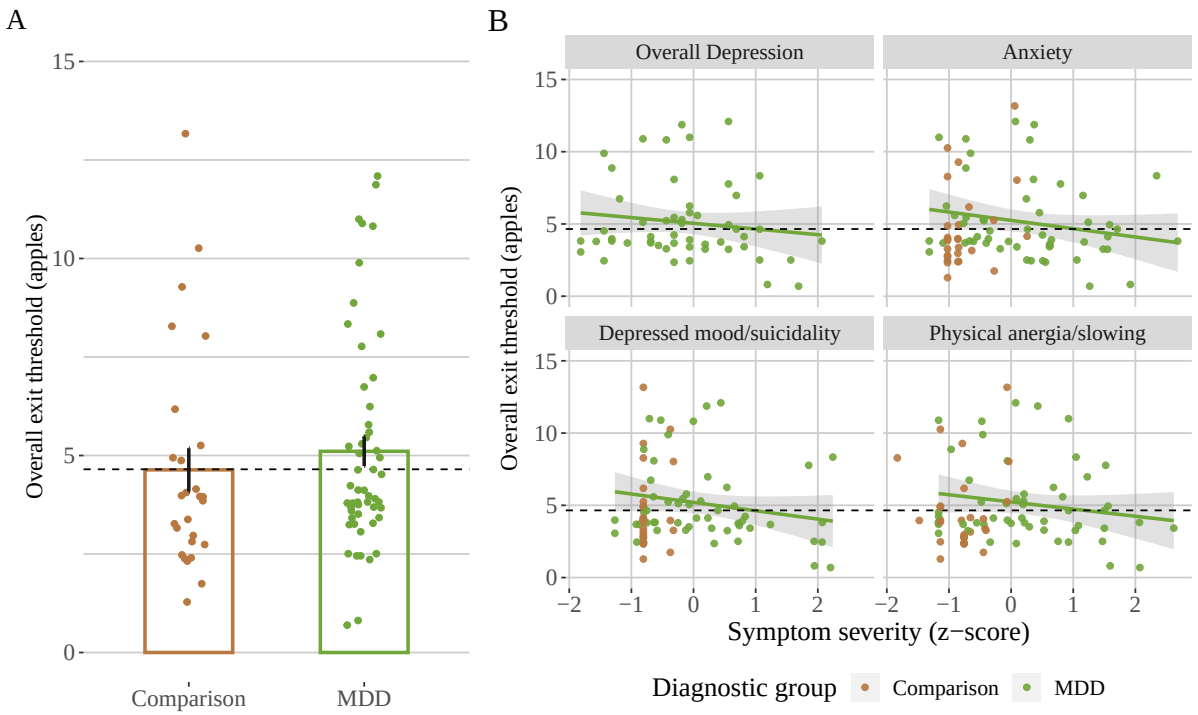

Figure 3: Relationship of individual MDD symptom domains with overall exit threshold (MDD group only). A: No diagnostic group differences (x-axis) in overall threshold threshold (y-axis, apples, estimated from low effort conditions). Bar indicates group means, error bars indicate standard error of the mean, points indicate mean overall exit threshold per participant. B: Lower overall exit threshold (y-axes) were significantly related to symptom severity (x-axes, z-scores) for overall depression (Hamilton Rating Scale Total), anxiety, depressed mood/suicidality, and physical anergia/slowing. Dashed line indicates best threshold policy, linear regression line for MDD group only, comparison participants included for completeness.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Symptom	Estimate	SE	*t*	*p*	$p_{adjusted}$
Overall exit threshold, MDD					
Overall depression	-0.03	0.01	-2.21	0.032*	
Anhedonia	-0.24	0.13	-1.82	0.075	0.106
Anxiety	-0.28	0.11	-2.43	0.019	0.044*
Behavioral apathy	-0.32	0.15	-2.19	0.033	0.058
Social apathy	-0.09	0.12	-0.70	0.490	0.490
Cognitive function symptoms	-0.16	0.11	-1.48	0.147	0.171
Depressed mood/suicidality	-0.34	0.12	-2.73	0.009	0.044*
Physical anergia/slowing	-0.31	0.13	-2.43	0.019	0.044*

Table 3: Symptom overall exit threshold regressions (MDD group only). Predicting individual differences in overall exit thresholds (log, from low effort conditions) by symptom severity, controlling for age and years of education (* indicates $p<0.05$, FDR correction within symptom models).

participant, in a heterogeneous group of participants with MDD and non-psychiatric comparisons. We found novel and important dissociable relations between symptom dimensions of MDD and cognitive versus physical effort. Our results corroborate several computational theories of depression and support breaking depression down into symptom domains and to examining components of effort- and value-based decision-making within-participants.

Diagnostic group differences

We predicted MDD would be associated with increased effort avoidance for both cognitive (1, 65, 69, 30) and physical (7, 14, 63, 65, 70, 72) effort. Contrary to our hypotheses, we did not observe significant group differences in effort costs. This aligns with null results in other studies of cognitive (4, 62) and physical effort avoidance (12, 58, 62, 68, 71). There were minimal task *performance* differences, except the MDD group had faster N-Back reaction times and completed fewer keypresses in the high physical effort condition. Travel task performance and effort costs were not correlated in this sample, suggesting they are dissociable in this task.

Symptom associations with effort costs

Greater overall depression severity in the MDD group was associated with greater willingness to exert cognitive effort, with no such relationship for physical effort. The comparison of correlations indicated specificity with respect to effort type. Anxiety symptom severity mostly accounted for the cognitive effort cost association (no other symptom domain was related to cognitive effort cost when controlling for anxiety), whereas physical effort cost was not reliably related to anxiety. The comparison of correlations indicated effort type specificity. To our knowledge no effort-based decision-making studies have reported decreased cognitive effort avoidance associated with MDD, suggesting unaccounted anxiety variation might contribute to inconsistent findings.

The negative association between anxiety and cognitive effort cost is consistent with reports of increased model-based planning associated with anxiety in unselected samples (25, 31). Increased cognitive effort exertion in anxiety might act as a compensatory mechanism to maintain performance amid threat-related attentional demands (20). Clinically, increased will-

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

ingness to exert cognitive effort may contribute to anxiety symptoms such as rumination and worry through increased planning and replay (5). Higher effort tasks might reduce anxious thoughts due to increased cognitive load (presumably via distraction). This is consistent with research showing reduced momentary anxiety during a high relative to low cognitive effort task (66).

Despite aiming to minimize demand characteristics, the social context of the experiment may have motivated anxious participants to exert effort (rather than expressing underlying preferences). An online study with the Effort Foraging Task found anxiety was associated with increased cognitive effort cost, opposite to present findings (10). However, the study was conducted in a large, unselected sample with self-reported symptoms, complicating translation to this clinical sample (another online study also did not report cognitive effort avoidance relationships to anxiety and depression) (48).

Anhedonia was significantly associated with increased physical effort costs within the MDD group, the current MDD group, and all participants (consistent with some reports, (58, 62, 70), but not others, (6, 12, 14, 65, 68, 71, 72)). Willingness to exert cognitive effort showed the opposite relationship to anhedonia and a comparison of correlations demonstrated differential relationships by effort type.

These findings support measuring both cognitive and physical effort decision-making function markers, which may inform heterogeneity or subtypes of MDD. Two lines of evidence suggest observed relationships between behavior and symptoms were driven by motivation rather than ability. First, all analyses relating effort costs to symptom severity accounted for ability (i.e., high effort performance) such that reported effects are over and above effects related to ability. Second, we found no reliable direct relationships between ability and MDD symptoms.

The Effort Foraging Task was developed to measure effort preferences more implicitly to increase validity. Therefore, methodological differences from previous (explicit) effort tasks may have contributed to the identification of symptom relationships, which has been mixed in other studies. On the other hand, implicit and explicit decisions may reflect unique effort-based decision-making aspects that differentially relate to MDD. It remains a question whether results from this task are more valid or if they tap into a different dimension of effort-based decision-making than the explicit task literature. If this were the case, it would enhance the novel contribution but also limit generalizability of the findings. Nevertheless, this work opens a new avenue for understanding how effort-based decision making relates to depression and other psychiatric symptoms, and how variations in task and modeling approaches affect such relationships.

Canonical correlation analysis

We used CCA as an omnibus analysis to examine linear relationships between the task measures and symptoms, traits, and demographics (age and BMI). There was one significant dimension. For the task coefficients, cognitive effort cost and 3-Back D' were positive, physical effort cost and 1-Back D' were negative. For the symptom coefficients anxiety, need for cognition, behavioral apathy, years of education, and social apathy were all negative, while BMI was positive. This pattern corroborates the regression findings, suggesting cognitive effort cost is negatively and physical effort cost is positively associated with symptom severity.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Subjective reward rate in depression

The MVT predicts that decisions to leave a patch reveal perceived environmental quality (i.e., opportunity cost of time), and this is supported by evidence in humans (e.g., lower thresholds under acute and chronic stress, and in persons with Parkinson's) (16, 37, 15). Following Huys et al. (2015) (33), we hypothesized MDD would be associated with lower exit thresholds, reflecting reduced average reward expectations. Few studies have linked psychiatric symptoms to sensitivity to opportunity cost of time, though one found an association with self-reported apathy (42). Overall exit thresholds did not differ by group (though a group effect emerged when accounting for psychotropic medication use) but were decreased in MDD participants with greater overall depression severity (also when excluding remitted participants), as well as anxiety, depressed mood/suicidality, and physical anergia/slowing symptoms (though these patterns were not maintained in the current MDD group nor across all participants). This suggests a difference in value-based decision-making associated with the severity of MDD symptoms, such that environments may be subjectively represented as less rewarding, reducing goal-directed behavior and vigor (33, 44).

Limitations

These results leave open the question of whether observed symptom associations generalize to other psychiatric conditions or other effort-based decision measures (e.g., tasks, ecological momentary assessment). To determine if the association between anxiety and cognitive effort cost is specific to MDD, future studies could include participants with primary clinical anxiety disorders. Another limitation is that sample size of the remitted depressed group did not allow for comparison with other groups. Heterogeneity in psychotropic medication use is another limitation, given neurotransmitter effects on aspects of cognition measured in the task. However, key effects were robust to excluding remitted participants and controlling for psychotropic medication use. The cross-sectional design limits understanding causality between symptoms and task behavior. Longitudinal designs could distinguish state versus trait influences on cognitive control and effort-based decision making and their interaction with symptoms.

Clinical implications

Ultimately insights from this research may inform interventions to increase willingness to exert effort for individuals experiencing challenges with goal-directed behavior due to psychiatric disability. Therapies that use cognitive restructuring to target physical effort perception might be effective for addressing anhedonia symptoms. Therapies to address physical anergia may target subjective reward rate, possibly through pharmacological dopamine manipulations (44). For applications to anxiety, the causal direction of the association with cognitive effort cost should be established. Does reduced cognitive effort cost cause anxiety symptoms, or does anxiety cause the pattern observed (i.e., benefits of distraction)? The tendency to exert cognitive effort could be leveraged as a strength in treating anxious depression (e.g., positive fantasizing, more cognitively effortful therapies, novel therapeutic applications using distraction) (8).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Availability of Data and Materials

The analysis code and data will be openly available at the Open Science Framework upon publication.

Acknowledgements

Thanks to Nicole Antkiewicz, Valerie Lilley, Yongjjing Ren, and Daniel Oliver for help with data collection. Thanks to Sam Zorowitz for technical assistance with remote task administration. Thanks to Sean Devine for help with the Hierachal Bayesian model. This work was supported by the New Jersey Alliance for Clinical and Translational Science (UL1TR003017), the New Jersey Health Foundation (NJHF#PC158-23), and NIH Awards T32MH065214, and T32DA007261.

Conflict of interest

The authors declare no conflict of interest.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

References

[1] Yuen-Siang Ang, Steven E. Gelda, and Diego A. Pizzagalli. Cognitive effort-based decision-making in major depressive disorder. *Psychological Medicine*, 53(9):4228–4235, July 2023.

[2] Yuen-Siang Ang, Patricia Lockwood, Matthew A. J. Apps, Kinan Muhammed, and Masud Husain. Distinct Subtypes of Apathy Revealed by the Apathy Motivation Index. *PloS One*, 12(1):e0169938, 2017.

[3] R. Michael Bagby, Andrew G. Ryder, Deborah R. Schuller, and Margarita B. Marshall. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *The American Journal of Psychiatry*, 161(12):2163–2177, December 2004.

[4] Deanna M. Barch, Adam J. Culbreth, Dror Ben Zeev, Andrew Campbell, Subigya Nepal, and Erin K. Moran. Dissociation of Cognitive Effort-Based Decision Making and its Associations with Symptoms, Cognition, and Everyday Life Function Across Schizophrenia, Bipolar Disorder, and Depression. *Biological Psychiatry*, April 2023.

[5] Rachel L. Bedder, Sashank Pisupati, and Yael Niv. Modelling Rumination as a State-Inference Process, May 2023.

[6] I. M. Berwian, H. Walter, E. Seifritz, and Q. J. M. Huys. Predicting relapse after antidepressant withdrawal – a systematic review. *Psychological Medicine*, 47(3):426–437, February 2017. Publisher: Cambridge University Press.

[7] Isabel M. Berwian, Julia G. Wenzel, Anne G. E. Collins, Erich Seifritz, Klaas E. Stephan, Henrik Walter, and Quentin J. M. Huys. Computational Mechanisms of Effort and Reward Decisions in Patients With Depression and Their Association With Relapse After Antidepressant Discontinuation. *JAMA psychiatry*, 77(5):513–522, May 2020.

[8] Marlijn E. Besten, Marie-José van Tol, Jacolien van Rij, and Marieke K. van Vugt. The impact of mood-induction on maladaptive thinking in the vulnerability for depression. *Journal of Behavior Therapy and Experimental Psychiatry*, 81:101888, December 2023.

[9] Paul Bliese, Gilad Chen, Patrick Downes, Donald Schepker, and Jonas Lang. multilevel: Multilevel Functions, March 2022.

[10] Laura A. Bustamante, Temitope Oshinowo, Jeremy R. Lee, Elizabeth Tong, Allison R. Burton, Amitai Shenhav, Jonathan D. Cohen, and Nathaniel D. Daw. Effort Foraging Task reveals positive correlation between individual differences in the cost of cognitive and physical effort in humans. *Proceedings of the National Academy of Sciences*, 120(50):e2221510120, December 2023. Publisher: Proceedings of the National Academy of Sciences.

[11] J. T. Cacioppo, R. E. Petty, and C. F. Kao. The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3):306–307, June 1984.

[12] Flurin Cathomas, Federica Klaus, Karoline Guetter, Erich Seifritz, Matthias N Hartmann-Riemer, Philippe N Tobler, Stefan Kaiser, and Mariia Kaliuzhna. Associations Between Negative Symptoms and Effort Discounting in Patients With Schizophrenia and Major Depressive Disorder. *Schizophrenia Bulletin Open*, 2(1):sgab022, January 2021.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[13] Eric L. Charnov. Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, 9(2):129–136, April 1976.

[14] Marie-Laure Cléry-Melin, Liane Schmidt, Gilles Lafargue, Nicolas Baup, Philippe Fossati, and Mathias Pessiglione. Why don't you try harder? An investigation of effort production in major depression. *PloS One*, 6(8):e23178, 2011.

[15] Sara M. Constantino, Jessica Dalrymple, Rebecca W. Gilbert, Sara Varanese, Alessandro Di Rocco, and Nathaniel D. Daw. A Neural Mechanism for the Opportunity Cost of Time. preprint, Neuroscience, August 2017.

[16] Sara M. Constantino and Nathaniel D. Daw. Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective & Behavioral Neuroscience*, 15(4):837–853, December 2015.

[17] Jessica A Cooper, Amanda R Arulpragasam, and Michael T Treadway. Anhedonia in depression: biological mechanisms and computational models. *Current Opinion in Behavioral Sciences*, 22:128–135, August 2018.

[18] Birk Diedenhofen and Jochen Musch. cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLoS ONE*, 10(4):e0121945, April 2015.

[19] David E. Evans and Mary K. Rothbart. Developing a model for adult temperament. *Journal of Research in Personality*, 41(4):868–888, August 2007.

[20] Michael W. Eysenck, Nazanin Derakshan, Rita Santos, and Manuel G. Calvo. Anxiety and cognitive performance: Attentional control theory. *Emotion*, 7(2):336–353, 2007. Place: US Publisher: American Psychological Association.

[21] Maurizio Fava, Dan V. Iosifescu, Paola Pedrelli, and Lee Baer. Reliability and validity of the Massachusetts general hospital cognitive and physical functioning questionnaire. *Psychotherapy and Psychosomatics*, 78(2):91–97, 2009.

[22] Michael B. First. Structured Clinical Interview for the DSM (SCID). In *The Encyclopedia of Clinical Psychology*, pages 1–6. John Wiley & Sons, Ltd, 2015. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118625392.wbecp351.

[23] Christopher Gagne, Ondrej Zika, Peter Dayan, and Sonia J. Bishop. Impaired adaptation of learning to contingency volatility in internalizing psychopathology, December 2020. Publisher: eLife Sciences Publications Limited.

[24] R. D. Gibbons, D. C. Clark, and D. J. Kupfer. Exactly what does the Hamilton Depression Rating Scale measure? *Journal of Psychiatric Research*, 27(3):259–273, 1993.

[25] Claire M. Gillan, Michal Kosinski, Robert Whelan, Elizabeth A. Phelps, and Nathaniel D. Daw. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, 5:e11305, March 2016.

[26] Ivan Grahek, Jonas Everaert, Ruth M. Krebs, and Ernst H. W. Koster. Cognitive Control in Depression: Toward Clinical Models Informed by Cognitive Neuroscience. *Clinical Psychological Science*, 6(4):464–480, July 2018. Publisher: SAGE Publications Inc.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[27] Ivan Grahek, Amitai Shenhav, Sebastian Musslick, Ruth M. Krebs, and Ernst H. W. Koster. Motivation and cognitive control in depression. *Neuroscience & Biobehavioral Reviews*, 102:371–381, July 2019.

[28] DC Halahakoon and JP Roiser. Cognitive impairment in depression and its (non-)response to antidepressant treatment. *Evidence-based mental health*, 19(4), November 2016. Publisher: Evid Based Ment Health.

[29] M. Hamilton. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23:56–62, February 1960.

[30] Rachel Hershenberg, Theodore D. Satterthwaite, Aylin Daldal, Natalie Katchmar, Tyler M. Moore, Joseph W. Kable, and Daniel H. Wolf. Diminished effort on a progressive ratio task in both unipolar and bipolar depression. *Journal of Affective Disorders*, 196:97–100, May 2016.

[31] Lindsay E. Hunter, Aaron M. Bornstein, and Catherine A. Hartley. A common deliberative process underlies model-based planning and patient intertemporal choice. Technical report, bioRxiv, December 2018. Section: New Results Type: article.

[32] Masud Husain and Jonathan P. Roiser. Neuroscience of apathy and anhedonia: a transdiagnostic approach. *Nature Reviews Neuroscience*, 19(8):470–484, August 2018. Number: 8 Publisher: Nature Publishing Group.

[33] Quentin J. M. Huys, Nathaniel D. Daw, and Peter Dayan. Depression: a decision-theoretic analysis. *Annual Review of Neuroscience*, 38:1–23, July 2015.

[34] Judith Jaeger, Stefanie Berns, Sarah Uzelac, and Sara Davis-Conway. Neurocognitive deficits and disability in major depressive disorder. *Psychiatry Research*, 145(1):39–48, November 2006.

[35] R. C. Kessler, C. B. Nelson, K. A. McGonagle, J. Liu, M. Swartz, and D. G. Blazer. Comorbidity of DSM-III-R major depressive disorder in the general population: results from the US National Comorbidity Survey. *The British Journal of Psychiatry. Supplement*, (30):17–30, June 1996.

[36] K. Kroenke, R. L. Spitzer, and J. B. Williams. The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9):606–613, September 2001.

[37] Jennifer K. Lenow, Sara M. Constantino, Nathaniel D. Daw, and Elizabeth A. Phelps. 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, June 2017.

[38] Charles J. Lynch, Faith M. Gunning, and Conor Liston. Causes and Consequences of Diagnostic Heterogeneity in Depression: Paths to Discovering Novel Biological Depression Subtypes. *Biological Psychiatry*, 88(1):83–94, July 2020.

[39] Simeng Ma, Jun Yang, Bingxiang Yang, Lijun Kang, Peilin Wang, Nan Zhang, Wei Wang, Xiaofen Zong, Ying Wang, Hanping Bai, Qingshan Guo, Lihua Yao, Li Fang, and Zhongchun Liu. The Patient Health Questionnaire-9 vs. the Hamilton Rating

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Scale for Depression in Assessing Major Depressive Disorder. *Frontiers in Psychiatry*, 12:747139, 2021.

[40] Xiao-li Meng, Robert Rosenthal, and Donald B. Rubin. Comparing correlated correlation coefficients. *Psychological Bulletin*, 111:172–175, 1992. Place: US Publisher: American Psychological Association.

[41] Mark J. Millan, Yves Agid, Martin Brüne, Edward T. Bullmore, Cameron S. Carter, Nicola S. Clayton, Richard Connor, Sabrina Davis, Bill Deakin, Robert J. DeRubeis, Bruno Dubois, Mark A. Geyer, Guy M. Goodwin, Philip Gorwood, Thérèse M. Jay, Marian Joëls, Isabelle M. Mansuy, Andreas Meyer-Lindenberg, Declan Murphy, Edmund Rolls, Bernd Saletu, Michael Spedding, John Sweeney, Miles Whittington, and Larry J. Young. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nature Reviews. Drug Discovery*, 11(2):141–168, February 2012.

[42] Akshay Nair, Ritwik K. Niyogi, Fei Shang, Sarah J. Tabrizi, Geraint Rees, and Robb B. Rutledge. Opportunity cost determines free-operant action initiation latency and predicts apathy. *Psychological Medicine*, 53(5):1850–1859, April 2023.

[43] Paul A. Nakonezny, Thomas J. Carmody, David W. Morris, Benji T. Kurian, and Madhukar H. Trivedi. Psychometric evaluation of the Snaith-Hamilton pleasure scale in adult outpatients with major depressive disorder. *International Clinical Psychopharmacology*, 25(6):328–333, November 2010.

[44] Yael Niv, Nathaniel D. Daw, Daphna Joel, and Peter Dayan. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, 191(3):507–520, April 2007.

[45] L. E. Nystrom, T. S. Braver, F. W. Sabb, M. R. Delgado, D. C. Noll, and J. D. Cohen. Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage*, 11(5 Pt 1):424–446, May 2000.

[46] Martin T. Orne. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11):776–783, 1962. Place: US Publisher: American Psychological Association.

[47] John E. Overall and Donald R. Gorham. The Brief Psychiatric Rating Scale. *Psychological Reports*, 10(3):799–812, June 1962. Publisher: SAGE Publications Inc.

[48] Edward H. Patzelt, Wouter Kool, Alexander J. Millner, and Samuel J. Gershman. The transdiagnostic structure of mental effort avoidance. *Scientific Reports*, 9(1):1689, February 2019. Number: 1 Publisher: Nature Publishing Group.

[49] Mathias Pessiglione, Fabien Vinckier, Sébastien Bouret, Jean Daunizeau, and Raphaël Le Bouc. Why not try harder? Computational approach to motivation deficits in neuropsychiatric diseases. *Brain*, 141(3):629–650, March 2018.

[50] RCoreTeam. R: A language and environment for statistical computing, 2015.

[51] Lionel Rigoux and Emmanuel Guigon. A model of reward- and effort-based optimal decision making and motor control. *PLoS computational biology*, 8(10):e1002716, 2012.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[52] Dimitris Rizopoulos. **ltm** : An *R* Package for Latent Variable Modeling and Item Response Theory Analyses. *Journal of Statistical Software*, 17(5), 2006.

[53] P. L. Rock, J. P. Roiser, W. J. Riedel, and A. D. Blackwell. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*, 44(10):2029–2040, July 2014.

[54] Joshua D Rosenblat, Ron Kakar, and Roger S McIntyre. The Cognitive Effects of Antidepressants in Major Depressive Disorder: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. *International Journal of Neuropsychopharmacology*, 19(2):pyv082, February 2016.

[55] John D. Salamone, Merce Correa, Jen-Hau Yang, Renee Rotolo, and Rose Presby. Dopamine, Effort-Based Choice, and Behavioral Economics: Basic and Translational Research. *Frontiers in Behavioral Neuroscience*, 12, 2018.

[56] Amitai Shenhav, Matthew M. Botvinick, and Jonathan D. Cohen. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, 79(2):217–240, July 2013.

[57] Amitai Shenhav, Sebastian Musslick, Falk Lieder, Wouter Kool, Thomas L. Griffiths, Jonathan D. Cohen, and Matthew M. Botvinick. Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, 40(1):99–124, 2017. _eprint: https://doi.org/10.1146/annurev-neuro-072116-031526.

[58] Lindsey Sherdell, Christian E. Waugh, and Ian H. Gotlib. Anticipatory pleasure predicts motivation for reward in major depression. *Journal of Abnormal Psychology*, 121(1):51–60, February 2012.

[59] Hannah R. Snyder. Major Depressive Disorder is Associated with Broad Impairments on Neuropsychological Measures of Executive Function: A Meta-Analysis and Review. *Psychological bulletin*, 139(1):81–132, January 2013.

[60] Robert L. Spitzer, Kurt Kroenke, Janet B. W. Williams, and Bernd Löwe. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10):1092–1097, May 2006.

[61] Stan. CmdStanR: the R interface to CmdStan — cmdstanr-package, 2021.

[62] Tanya Tran, Amanda E. F. Hagen, Tom Hollenstein, and Christopher R. Bowie. Physical- and Cognitive-Effort-Based Decision-Making in Depression: Relationships to Symptoms and Functioning. *Clinical Psychological Science*, 9(1):53–67, January 2021. Publisher: SAGE Publications Inc.

[63] Michael T. Treadway, Nicholas A. Bossaller, Richard C. Shelton, and David H. Zald. Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *Journal of Abnormal Psychology*, 121(3):553–558, August 2012.

[64] Michael T. Treadway, Joshua W. Buckholtz, Ashley N. Schwartzman, Warren E. Lambert, and David H. Zald. Worth the 'EEfRT'? The Effort Expenditure for Rewards Task as an Objective Measure of Motivation and Anhedonia. *PLOS ONE*, 4(8):e6598, August 2009. Publisher: Public Library of Science.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[65] Fabien Vinckier, Claire Jaffre, Claire Gauthier, Sarah Smajda, Pierre Abdel-Ahad, Raphaël Le Bouc, Jean Daunizeau, Mylène Fefeu, Nicolas Borderies, Marion Plaze, Raphaël Gaillard, and Mathias Pessiglione. Elevated Effort Cost Identified by Computational Modeling as a Distinctive Feature Explaining Multiple Behaviors in Patients With Depression. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 7(11):1158–1169, November 2022.

[66] Katherine Vytal, Brian Cornwell, Nicole Arkin, and Christian Grillon. Describing the interplay between anxiety and cognition: From impaired performance under low cognitive load to reduced anxiety under high load. *Psychophysiology*, 49(6):842–852, 2012. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8986.2012.01358.x.

[67] Mark E. Walton, Peter H. Rudebeck, David M. Bannerman, and Matthew F. S. Rushworth. Calculating the cost of acting in frontal cortex. *Annals of the New York Academy of Sciences*, 1104:340–356, May 2007.

[68] Yan-yu Wang, Yi Wang, Jia Huang, Xi-he Sun, Xi-zhen Wang, Shu-xian Zhang, Guohui Zhu, Simon S. Y. Lui, Eric F. C. Cheung, Hong-wei Sun, and Raymond C. K. Chan. Shared and distinct reward neural mechanisms among patients with schizophrenia, major depressive disorder, and bipolar disorder: an effort-based functional imaging study. *European Archives of Psychiatry and Clinical Neuroscience*, 272(5):859–871, August 2022.

[69] Andrew Westbrook, Xiao Yang, Lauren M. Bylsma, Shimrit Daches, Charles J. George, Andrew J. Seidman, J. Richard Jennings, and Maria Kovacs. 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*, August 2022.

[70] Xin-Hua Yang, Jia Huang, Cui-Ying Zhu, Ye-Fei Wang, Eric F. C. Cheung, Raymond C. K. Chan, and Guang-Rong Xie. Motivational deficits in effort-based decision making in individuals with subsyndromal depression, first-episode and remitted depression patients. *Psychiatry Research*, 220(3):874–882, December 2014.

[71] Xinhua Yang, Jia Huang, Phillippa Harrision, Matthew E. Roser, Kai Tian, Dongfang Wang, and Guangya Liu. Motivational differences in unipolar and bipolar depression, manic bipolar, acute and stable phase schizophrenia. *Journal of Affective Disorders*, 283:254–261, March 2021.

[72] Ying-Min Zou, Ke Ni, Yan-Yu Wang, En-Qing Yu, Simon S. Y. Lui, Fu-Chun Zhou, Han-Xue Yang, Alex S. Cohen, Gregory P. Strauss, Eric F. C. Cheung, and Raymond C. K. Chan. Effort-cost computation in a transdiagnostic psychiatric sample: Differences among patients with schizophrenia, bipolar disorder, and major depressive disorder. *PsyCh Journal*, 9(2):210–222, April 2020.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Supplementary Information

Contents

- SI section 1 Study overview
- SI section 1.2 Diagnostic group matching
- SI section 2 Power analysis
- SI section 3 Major depressive episode criteria
- SI section 4 Clinician ratings and self reports
- Table S1 Clinical interview and self-report battery
- SI section 5 Symptom confirmatory factor analysis
- Table S2 Confirmatory symptom factors alpha and diagnostic group differences
- Table S3 Items for confirmatory symptom factors
- Figure S1 Effort Foraging Task Diagram
- Figure S2 Travel task methods
- SI section 7 Foraging behavior analysis exclusions
- Table S4 Foraging environment and best threshold
- SI section 8 MVT model additional methods
- SI section 9 MVT model evaluation
- SI section 10 Simulation to find best threshold
- SI section 11 Canonical Correlation Analysis (CCA) methods
- SI section 12 Additional task measures methods
- SI section 12.2 Overall exit threshold
- Table S5 Demographic factors by diagnostic group
- Figure S3 MDD sample characteristics
- Figure S4 Symptom severity by diagnostic group
- Table S6 MVT group-level posterior distributions
- Figure S5 Model diagnostics
- Figure S6 Posterior predictive check
- SI section 13 Relationship between cognitive and physical effort costs
- Table S7 Symptom regressions, control for psychotropic medication use, MDD

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

- Table S8 Diagnostic group difference MVT model
- SI section 14 Model-agnostic sensitivity to effort manipulation
- Figure S7 Model-agnostic change in exit threshold by diagnostic group
- Table S9 Self-reported overall depression regression results
- Table S10 Symptoms effort cost regressions, current MDD and all participants
- Figure S8 Task behavior symptom heatmap
- Figure S9 Canonical Correlation Analysis
- SI section 16 Additional task measures results
- Figure S10 Travel task performance by diagnostic group
- Table S11 Diagnostic group differences in travel task performance
- Figure S11 Performance relationship to effort costs and symptoms
- Table S12 Overall exit threshold relationship to symptoms, current MDD and all participants

1 Study overview

The study was conducted at the Rutgers-Princeton Center for Computational Neuro-Psychiatry by trained clinical researchers. The study was approved by the Rutgers University Institutional Review Board. Participants were recruited in outpatient clinics at Rutgers University, as well as via Google Ads. We recruited participants with MDD confirmed using the Structured Clinical Interview for DSM-5 (9, SCID-5) and no co-occurring psychiatric conditions except for anxiety disorders, which were permitted. MDD participants in remission ("no significant symptoms during the past 2 months") and in partial remission ("symptoms are present but full criteria are not met, for a period less than 2 months") were permitted. We also recruited a demographically matched comparison group without any psychiatric diagnosis. The study was administered remotely via secure web-based software (Zoom, except for 5 MDD participants who came an in-person session) and participants were compensated $20 per hour and an Effort Foraging Task bonus of up to $10. On the first session a clinical interviewer administered the SCID-5 and participants completed self-report surveys. On the second session participants completed the Effort Foraging Task. In a third session participants completed tasks assessing reward sensitivity and cognitive control recruitment in response to efficacy manipulations, which will be reported separately.

1.1 Inclusion and exclusion criteria

The study inclusion criteria were 1) between the ages of 18-65, 2) has the capacity to provide informed consent, and 3) is fluent in English, 4) score of 6 or higher on the Wechsler Test of Adult Reading (equivalent to standardized score of 53 - 60 for ages 18-64, (24)), 5) (MDD

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

only) Meets DSM-5 criteria for MDD as confirmed by the SCID-5, (MDD only) if the participant is treated with anti-depressant medication, they are on stable treatment with this medication (i.e. no change in medication type, or substantial change in dose, for at least 4 weeks prior to participating in the study). The study exclusion criteria were 1) history of traumatic brain injury or head injury, 2) Intellectual disability or pervasive developmental disorder, 3) neurological disease, 4) has met DSM-5 criteria for a substance-use disorder within the last 6 months per the SCID-5 (with the exception of nicotine dependence, which was permitted), 5) received electroconvulsive therapy within the last 8 weeks, 6) left-handedness (due to keyboard set up for the Effort Foraging Task), 7) (comparisons only) meets DSM-5 criteria for any psychiatric diagnosis as confirmed by the SCID-5, 8) (comparisons only) current use of any psychotropic medication.

1.2 Diagnostic group matching

We used Pearson's Chi-squared test to compare categorical diagnostic group differences (gender, race, ethnicity, total household income, occupational status, relationship status, alcohol frequency, alcohol amount, caffeine amount, tobacco use), and Welch two sample unpaired t-test to compare continuous diagnostic group differences (age, childhood income, years of education for self and parents), using the chisq.test and t.test functions of the stats package in the R language, (20)). If diagnostic groups were not matched on a demographic variable, we included it as a covariate in our group difference and symptom relationships to task behavior analyses.

2 Power analysis

We conducted a power analysis with the current sample size (using the pwr package (4)). For the symptom regressions, power analysis indicated we could detect a medium effect size with 80% power with the current sample of 52 MDD participants (F^2=0.254, Cohen's F^2 'medium' effect size between 0.15 and 0.35, (5)). For the diagnostic group differences, power analysis indicated we could detect a medium effect size with 80% power with the current sample of 52 MDD and 27 comparison participants (D=0.673, Cohen's D 'medium' effect size between 0.5 and 0.8, (5)).

3 Major depressive episode criteria

Participants were assigned to the MDD diagnostic group based on the SCID-5, which considers the lifetime history for the MDD diagnosis, and symptoms in the past two months to establish whether a participant is 1) currently depressed, meeting criteria for a major depressive episode any time in the past month, 2) in partial remission, either experiencing some symptoms but not meeting full criteria for a major depressive episode, or there is a period lasting less than two months without significant symptoms, or 3) in full remission, during the past two months no significant symptoms. All participants in the MDD group have a lifetime history of depression, but varied in how many symptoms they experienced in the past week, as well as in the past two months. This dynamic variation in MDD symptom expression, if unaccounted for, may contribute to mixed results surrounding diagnostic group differences in effort-based decision making. This motivated us to focus on individual differences in symptom severity at the time of the study, rather than solely diagnostic group differences. The

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

present study utilized the HAMD, which measures symptoms in the past week, along with the BPRS which measures symptoms in the past 2-3 days, as well as self-reports which asked about either the past few days, two weeks, or month. Therefore, currently depressed participants may have low scores on the HAMD (or BPRS) if they report fewer symptoms in the past week relative to the past month (during which they did meet full criteria for a major depressive episode). Low scores on the HAMD (or BPRS) would be expected for fully or partially remitted MDD participants, although scores could be higher for these participants if they are experiencing some symptoms but did not meet full criteria for a major depressive episode.

4 Clinician ratings and self reports

The Structured Clinical Interview for DSM-5 (SCID-5) (9) was used to confirm assignment of MDD, co-morbid anxiety, and comparison groups, and that participants met study diagnostic inclusion and exclusion criteria. The clinical interviewer used responses in the SCID-5 to rate the severity of different symptoms in the past week via the the Hamilton Depression Rating Scale (HAMD) (12) and to determine whether MDD participants were currently depressed, or in partial or full remission. This was followed by the semi-structured interview the Brief Psychiatric Rating Scale (19) to assess current psychiatric-symptom severity. Next participants completed self-report surveys to measure; cognitive function symptoms and physical anergia with the Massachusetts General Hospital Cognitive and Physical Functioning Questionnaire (8), depression symptoms with the Patient Heath Questionnaire-9 (13), anxiety symptoms with the Generalized Anxiety Disorder-7 (21), anhedonia symptoms with the Snaith–Hamilton Pleasure Scale (17), behavioral, emotional, and social apathy with the Apathy Motivation Index (1). We also measured trait executive function with the Adult Temperament Questionnaire Effortful Control subscale (7), and trait cognitive control seeking with the Need for Cognition scale (3).

Scale name	N items	Abbreviation	Time scale
Hamilton Depression Rating Scale	21	HAM-D	Past week
Brief Psychiatric Rating Scale	24	BPRS	Last 2 weeks
MGH Cognitive and Physical Functioning Questionnaire	7	MGH-CPFQ	Last month
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
Adult Temperament Questionnaire-Effortful Control	19	ATQ-EC	Trait
Need for Cognition	18	NFC	Trait
Total Self-report	87		

Table S1: Clinical interview and self-report battery. Clinical interview and self-reports were completed in session 1 in the order listed in this table. Scales asking about similar timescales were grouped together.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

5 Symptom confirmatory factor analysis

We performed confirmatory factor analysis and assigned items from clinician ratings and self-report into the following domains (using measures listed in Table S1, exact items in Table S3): anhedonia, appetite symptoms, anxiety, behavioral apathy, cognitive function symptoms, depressed mood, effortful control (trait), emotional apathy, need for cognition (trait), physical anergia/slowing, social apathy. Assigned items were z-scored and averaged to compute a symptom score in (1) the MDD group only and (2) all participants. Confirmatory factor scores for comparison participants did not include ratings from clinician measures (i.e., HAMD, and BPRS) and instead were the average z-score of all self-report items in a factor.

We tested whether diagnostic groups significantly differed in symptom intensity and cognitive control trait measures using t-tests (Table S2). For left skewed distributions (comparisons clustered on very low symptom scores for anxiety, cognitive function symptoms, depressed mood, and physical anergia/slowing) we confirmed the diagnostic group differences results were maintained using the non-parametric Wilcoxon rank sum test with continuity correction (wilcox.test function of the stats package in the R language, (20)).

Symptom domain	**alpha(95% CI)**	**N**	**MDD**	**Comp.**	***t***	**df**	***p***
Anhedonia	0.89(0.84-0.92)	15	0.132	-0.294	3.07	55.46	0.003
Anxiety	0.90(0.86-0.93)	10	0.197	-0.626	7.386	83.523	<0.001
Appetite symptoms	0.52(0.20-0.69)	3	0.091	-0.609	5.501	76.74	<0.001
Behavioral apathy	0.72(0.53-0.82)	8	0.186	-0.554	6.124	58.79	<0.001
Emotional apathy	0.54(0.30-0.68)	5	-0.049	0.137	-1.22	44.08	0.230
Social apathy	0.71(0.55-0.81)	5	0.149	-0.331	3.303	58.024	0.002
Cognitive function symp.	0.84(0.75-0.89)	5	0.257	-0.572	6.880	80.293	<0.001
Depressed mood	0.86(0.81-0.90)	8	0.106	-0.626	7.508	66.346	<0.001
Effortful control (trait)	0.84(0.76-0.88)	17	-0.177	0.393	-6.29	80.5	<0.001
Need for cognition (trait)	0.90(0.85-0.93)	17	-0.033	0.074	-0.794	58.151	0.430
Physical anergia/slowing	0.76(0.64-0.84)	7	0.157	-0.611	7.616	84.89	<0.001

Table S2: Confirmatory symptom factors alpha and diagnostic group differences. Column 1: symptom domain, column 2: Cronbach's alpha (95% confidence interval), column 3: number of items, column 4-8: MDD mean, Comparison mean, and t-statistic, degrees of freedom, p-value. Emotional apathy and appetite symptoms had low Cronbach's alpha scores and were not included in further analysis.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Category	Items	Scale
Anhedonia	Little interest or pleasure in doing things	PHQ-9
Anhedonia	I would enjoy my favorite television or radio program	SHAPS
Anhedonia	I would enjoy being with family or close friends	SHAPS
Anhedonia	I would find pleasure in my hobbies and pastimes	SHAPS
Anhedonia	I would be able to enjoy my favorite meal	SHAPS
Anhedonia	I would enjoy a warm bath or refreshing shower	SHAPS
Anhedonia	I would find pleasure in the scent of flowers or the smell of a fresh sea breeze or freshly baked bread	SHAPS
Anhedonia	I would enjoy seeing other people's smiling faces	SHAPS
Anhedonia	I would enjoy looking smart when I have made an effort with my appearance	SHAPS
Anhedonia	I would enjoy reading a book, magazine or newspaper	SHAPS
Anhedonia	I would enjoy a cup of tea or coffee or my favorite drink	SHAPS
Anhedonia	I would find pleasure in small things; e.g., bright sunny day, a telephone call from a friend	SHAPS
Anhedonia	I would be able to enjoy a beautiful landscape or view	SHAPS
Anhedonia	I would get pleasure from helping others	SHAPS
Anhedonia	I would feel pleasure when I receive praise from other people	SHAPS
Anxiety	Anxiety psychic	HAM-D
Anxiety	Anxiety - somatic	HAM-D
Anxiety	Anxiety	BPRS
Anxiety	Feeling nervous, anxious, or on edge	GAD-7
Anxiety	Not being able to stop or control worrying	GAD-7
Anxiety	Worrying too much about different things	GAD-7
Anxiety	Trouble relaxing	GAD-7
Anxiety	Being so restless that it's hard to sit still	GAD-7
Anxiety	Becoming easily annoyed or irritable	GAD-7
Anxiety	Feeling afraid as if something awful might happen	GAD-7
Behavioral apathy	How has your motivation/interest/enthusiasm been over the past month?	MGH-CPFQ
Behavioral apathy	Work and interests	HAM-D
Behavioral apathy	Self-neglect	BPRS
Behavioral apathy	I make decisions firmly and without hesitation.	AMI
Behavioral apathy	When I decide to do something, I am able to make an effort easily.	AMI
Behavioral apathy	I get things done when they need to be done, without requiring reminders from others.	AMI
Behavioral apathy	When I decide to do something, I am motivated to see it through to the end.	AMI
Behavioral apathy	When I have something I need to do, I do it straightaway so it is out of the way.	AMI
Emotional apathy	Emotional withdrawal	BPRS

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Category	Items	Scale
Emotional apathy	I feel sad or upset when I hear bad news.	AMI
Emotional apathy	Based on the last two weeks, I would say I care deeply about how my loved ones think of me.	AMI
Emotional apathy	I feel bad when I hear an acquaintance has an accident or illness.	AMI
Emotional apathy	If I realize I have been unpleasant to someone, I will feel terribly guilty afterwards.	AMI
Social apathy	I start conversations with random people.	AMI
Social apathy	I enjoy doing things with people I have just met.	AMI
Social apathy	I suggest activities for me and my friends to do.	AMI
Social apathy	I go out with friends on a weekly basis.	AMI
Social apathy	I start conversations without being prompted.	AMI
Appetite	Somatic symptoms - gastro-intestinal	HAM-D
Appetite	Weight loss	HAM-D
Appetite	Poor appetite or overeating	PHQ-9
Effortful control (trait)	I am often late for appointments.	ATQ-EC
Effortful control (trait)	I often make plans that I do not follow through with.	ATQ-EC
Effortful control (trait)	I can keep performing a task even when I would rather not do it.	ATQ-EC
Effortful control (trait)	I can make myself work on a difficult task even when I don't feel like trying.	ATQ-EC
Effortful control (trait)	If I think of something that needs to be done, I usually get right to work on it.	ATQ-EC
Effortful control (trait)	I usually finish doing things before they are actually due (for example, paying bills, finishing homework, etc.).	ATQ-EC
Effortful control (trait)	When I am afraid of how a situation might turn out, I usually avoid dealing with it.	ATQ-EC
Effortful control (trait)	It's often hard for me to alternate between two different tasks.	ATQ-EC
Effortful control (trait)	When I am trying to focus my attention, I am easily distracted.	ATQ-EC
Effortful control (trait)	When interrupted or distracted, I usually can easily shift my attention back to whatever I was doing before.	ATQ-EC
Effortful control (trait)	It is very hard for me to focus my attention when I am distressed.	ATQ-EC
Effortful control (trait)	When I am happy and excited about an upcoming event, I have a hard time focusing my attention on tasks that require concentration.	ATQ-EC
Effortful control (trait)	Even when I feel energized, I can usually sit still without much trouble if it's necessary.	ATQ-EC
Effortful control (trait)	I can easily resist talking out of turn, even when I'm excited and want to express an idea.	ATQ-EC
Effortful control (trait)	I usually have trouble resisting my cravings for food drink, etc.	ATQ-EC

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Category	Items	Scale
Effortful control (trait)	When I'm excited about something, it's usually hard for me to resist jumping right into it before I've considered the possible consequences.	ATQ-EC
Effortful control (trait)	When I see an attractive item in a store, it's usually very hard for me to resist buying it.	ATQ-EC
Need for cognition (trait)	I would prefer complex to simple problems.	NFC
Need for cognition (trait)	I like to have the responsibility of handling a situation that requires a lot of thinking.	NFC
Need for cognition (trait)	Thinking is not my idea of fun.	NFC
Need for cognition (trait)	I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.	NFC
Need for cognition (trait)	I try to anticipate and avoid situations where there is a likely chance I will have to think in depth about something	NFC
Need for cognition (trait)	I find satisfaction in deliberating hard and for long hours.	NFC
Need for cognition (trait)	I only think as hard as I have to.	NFC
Need for cognition (trait)	I prefer to think about small daily projects to long term ones.	NFC
Need for cognition (trait)	I like tasks that require little thought once I've learned them.	NFC
Need for cognition (trait)	The idea of relying on thought to make my way to the top appeals to me.	NFC
Need for cognition (trait)	I really enjoy a task that involves coming up with new solutions to problems.	NFC
Need for cognition (trait)	Learning new ways to think doesn't excite me very much.	NFC
Need for cognition (trait)	I prefer my life to be filled with puzzles I must solve.	NFC
Need for cognition (trait)	The notion of thinking abstractly is appealing to me.	NFC
Need for cognition (trait)	I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.	NFC
Need for cognition (trait)	I feel relief rather than satisfaction after completing a task that requires a lot of mental effort.	NFC
Need for cognition (trait)	It's enough for me that something gets the job done; I don't care how or why it works.	NFC
Cognitive function symptoms	Distractibility (speech and actions interrupted by stimuli unrelated to the interview)	BPRS
Cognitive function symptoms	How has your ability to focus/sustain attention been over the past month?	MGH-CPFQ

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Category	Items	Scale
Cognitive function symptoms	How has your ability to remember/recall information been over the past month?	MGH-CPFQ
Cognitive function symptoms	How has your ability to find words been over the past month?	MGH-CPFQ
Cognitive function symptoms	How has your sharpness/mental acuity been over the past month?	MGH-CPFQ
Cognitive function symptoms	Trouble concentrating on things, such as reading the newspaper or watching television	PHQ-9
Depressed mood	Depressed mood	HAM-D
Depressed mood	Feelings of guilt	HAM-D
Depressed mood	Suicide	HAM-D
Depressed mood	Depression	BPRS
Depressed mood	Suicidality	BPRS
Depressed mood	Guilt	BPRS
Depressed mood	Feeling down, depressed, or hopeless	PHQ-9
Depressed mood	Feeling bad about yourself - or that you are a failure or have let yourself or your family down	PHQ-9
Depressed mood	Thoughts that you would be better off dead, or of hurting yourself	PHQ-9
Physical anergia/slowing	Somatic symptoms - general	HAM-D
Physical anergia/slowing	Psychomotor retardation	HAM-D
Physical anergia/slowing	Motor retardation (slowed or reduced movements or speech)	BPRS
Physical anergia/slowing	How has your wakefulness/alertness been over the past month?	MGH-CPFQ
Physical anergia/slowing	How has your energy been over the past month?	MGH-CPFQ
Physical anergia/slowing	Feeling tired or having little energy	PHQ-9
Physical anergia/slowing	Moving or speaking so slowly that other people could have noticed. Or the opposite - being so figety or restless that you have been moving around a lot more than usual.	PHQ-9

Table S3: Items for confirmatory symptom factors. Column 1: symptom domain, column 2: items, column 3: measurement scale (abbreviations in Table S1).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

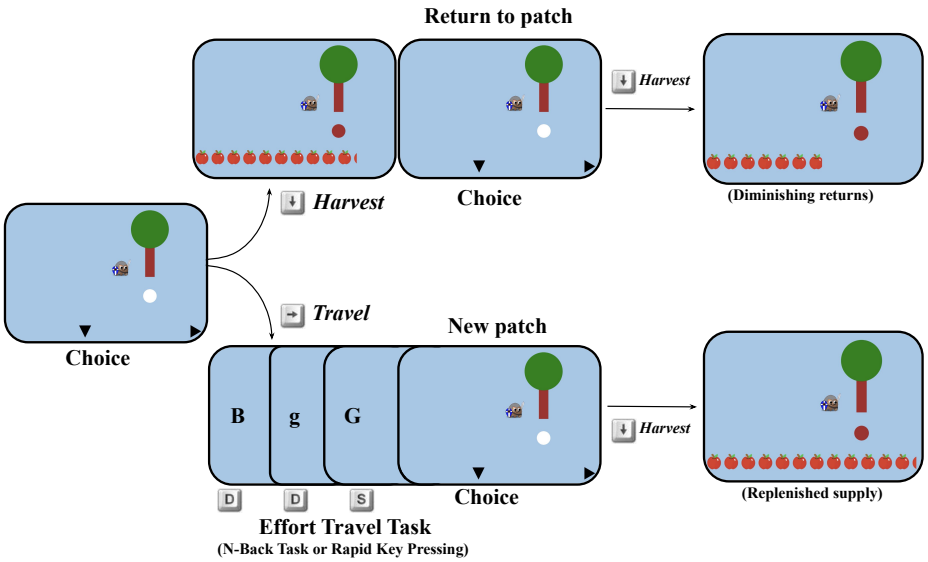

Figure S1: Effort Foraging Task Diagram. On each trial participants chose to harvest a virtual patch (apple tree) using the down arrow key, or travel to a new patch. Harvesting a patch yielded diminishing returns, whereas traveling to a new patch cost time and effort. Travel tasks were either the 1-Back or 3-Back levels of the N-Back task, or a smaller or larger number of rapid keypresses. Adapted from Figure 1 of Bustamante et al., 2023 (2).

6 Effort Foraging Task counterbalancing

The order of cognitive and physical effort variants was counterbalanced across participants. 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 orders. Participants were assigned a block order using latin squares within each diagnostic group.

7 Foraging behavior analysis exclusions

Of the 60 MDD and 27 comparison participants, 1 MDD participant did not complete the Effort Foraging Task due to technical difficulties with their keyboard. All other participants completed the task, however technical difficulties with the experiment server caused 4 missing data files from the MDD group. We followed a subset of exclusions validated in Bustamante et al. (2023) that most interfere with estimating effort costs. First, participants were excluded if they had very few exit trials within an effort type, making their data under-powered for estimating exit thresholds, and overly deterministic for logistic regression, which are the basis of the effort cost measures (2*SD below the mean, <8.82 trials). As a result 1 MDD participant was excluded from analysis for the whole task (1 exit in high effort physical and 3 exits in high effort cognitive condition) and 1 MDD participant was excluded from the cognitive effort analyses (2 exits for the cognitive high effort condition). Second, participants were excluded from the task if they missed the response deadline on many foraging trials (2*SD above the mean, >15.05%, 1 MDD participant excluded who missed 49.5% of trials) which

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

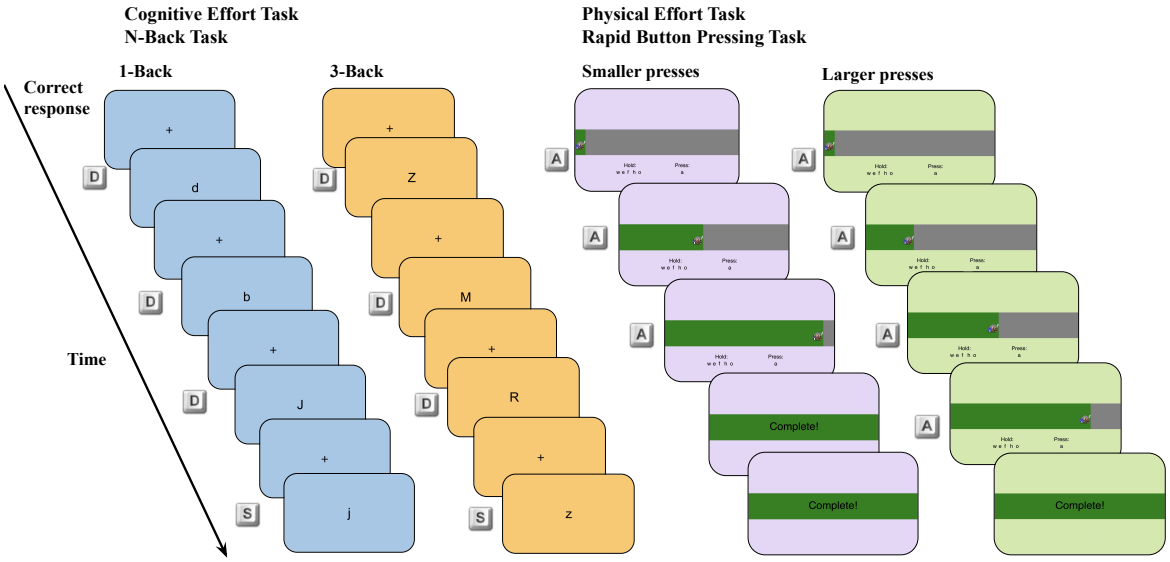

Figure S2: Travel task methods. Left panel: cognitive effort, N-Back working memory task. Participants responded whether the letter on the screen was the same ('s' key) or different ('d' key). The background color differed for the high effort (3-Back, orange) and low effort (1-Back, blue) conditions. Key icons next to each screen indicate the correct response. Right panel: physical effort, rapid key-pressing task. Participants rapidly pressed the 'a' key while holding down the 'w', 'e', 'f' (left hand), 'h' and 'o' keys (right hand). 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, 50% of maximum, purple) and low effort (larger presses, 100% of maximum, green) conditions. Adapted from Figure 2 of Bustamante et al., 2023 (2).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

A. Environment parameter	**Value**
Harvest time	2 seconds
Travel time	8.33 seconds
Cognitive Travel Tasks	1-Back, 3-Back
Physical Travel Tasks	Larger (100% max), Smaller (50%) number of presses
Block duration	7 minutes
Number blocks per condition	2
Initial reward	$\mathcal{N}(15, 1)$
Decay rate (κ)	β(14.13, 2.03)
Total Number of blocks	8
B. Best threshold simulation	**Value**
Best threshold (ρ, apples)	4.65
Reward rate (apples/second)	2.32
Number of harvests μ	9.95
Number of harvests SD	2.01

Table S4: Foraging environment parameters and results of best threshold simulation. A: column 1: environment parameter, column 2: value. B: Best exit threshold policy identified in Bustamante et al. (2023), rows indicate best threshold from simulation, reward rate achieved with best threshold (apples per second), mean and standard deviation of the number of harvests it took to reach the best threshold.

may reflect low engagement with the task or challenges meeting the response deadline. Ultimately, this affects the interpretability of MVT estimates (e.g., experienced harvest time longer than for other participants, fewer apples per second). The final sample included in behavioral analyses was 52 MDD participants (53 MDD participants in the physical effort condition) and 27 comparison participants.

8 MVT model additional methods

Because we are investigating individual differences in effort costs at the condition level, used a factorial model in which the MVT threshold is taken as fixed per-condition, determined by the overall rewards and delays in each condition and a per-condition effort-cost parameter. Thus, the model omits trial-by-trial learning of the threshold, and instead formally absorbs any such variation into the softmax choice stochasticity. We believe this simplification is warranted because the condition-wise effort costs of interest aggregate over per-trial threshold variability, and because we encouraged asymptotic behavior through pre-training and using a stable foraging environment throughout.

There were five parameters in the model, the inverse temperature (β, which controls the noise of the softmax choice function, with lower values indicating more noisy effects of rewards and thresholds on choices), the cognitive low ($c_{\text{cog low effort}}$) and high effort costs ($c_{\text{cog high effort}}$), and the physical low ($c_{\text{phys low effort}}$ and high effort costs ($c_{\text{phys high effort}}$). 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 estimate the correlation between individual differences in high cognitive and physical effort costs. Model priors were centered at zero and variances were selected to accommodate the magnitude of group-level posterior

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

distributions from the original Effort Foraging Task study (Experiment 2) (2). The prior distributions for group-level effects were $c_{\text{low effort}} \sim \mathcal{N}(0, 25)$, $c_{\text{high effort}} \sim \mathcal{N}(0, 15)$, $\beta \sim \mathcal{N}(0, 1)$. The prior on random effects variances were $c_{\text{low effort}} \sim \mathcal{N}(0, 25)$, $c_{\text{high effort}} \sim \mathcal{N}(0, 15)$, $\beta \sim \mathcal{N}(0, 1)$. The prior on the correlation matrix was unbiased as to the presence or absence of a correlation (LKJ Correlation Distribution prior=1, (14)). 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 four chains, Stan Development Team) (22). Convergence was assessed by visually inspecting model traces, and ensuring the $\hat{\text{R}}$ convergence diagnostic statistic was below 1.1. We also simulated the MVT model to estimate the best exit threshold with respect to reward and time given the foraging environment parameters. To test for diagnostic group (g) differences, we fitted a Hierarchical Bayesian MVT model in which each of the 5 group-level parameters (p) had a diagnostic group effect ($\beta_{g,p}$) added to it. For a participant (i) in diagnostic group (g_i) each participant-level parameter p_i was the sum of the group-level parameter p and the diagnostic group effect ($p_i = p + \beta_{g,p} * g_i$, where $g_{MDD}, g_{Comparison} \in 0.5, -0.5$). The diagnostic group effect parameter ($\beta_{g,p}$) had a prior distribution $\mathcal{N}(0, 5)$, values greater than zero indicate higher effort costs in the MDD relative to comparison group. To confirm the remitted and partial remitted participants did not change the results, we also used this model to test for group differences of the comparison group to the currently depressed subset of the MDD group.

9 MVT model evaluation

We used several methods to evaluate the MVT model fit. We inspected trace plots to ensure mixing between chains, and the $\hat{\text{R}}$ convergence diagnostic was below 1.1 for all parameters (using the rhat and mcmc_trace functions from the bayesplot package in R) (10). We conducted a posterior predictive check to confirm the fitted model captured foraging decisions. For each of 8,000 MCMC samples, for all trials across the entire dataset, the model sampled from the posterior predictive distribution from a Bernoulli distribution, generating a set of harvest (1) or exit (0) choices (using the bernoulli_logit_rng function in Stan) (23). We examined the correspondence of the posterior predictive samples to the empirical data. We tested whether the empirical probability of stay choices (across all participants and all trials in the dataset) fell within the posterior predictive distribution (i.e., the simulated probability of stay choices across all trials for each MCMC sample). To do so, we computed the distance from the median simulated probability of staying for every MCMC sample, as well as the distance from the empirical data. We tested the probability that the distance of the empirical data to the simulated median was larger than the distances of the simulated data. Additionally, we visually compared the overall exit threshold, as well as the change in exit threshold by diagnostic group, in simulated versus empirical data.

We assessed the log posterior likelihood per participant and used an unpaired t-test to compare goodness of fit for the model between the diagnostic groups. To do so, we computed the sum of log likelihoods within each participant, for each MCMC sample (using the bernoulli_logit_lpmf function in Stan, (23)). Then we aggregated across MCMC samples by exponentiating these values, summing them, and log-transforming them, resulting in one log posterior likelihood value per participant.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

10 Simulation to find best threshold

We considered the best threshold found by simulation in Bustamante et al. (2023) (2) for Experiment 2. To repeat the methods we used in that study, 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).

11 Canonical Correlation Analysis (CCA) methods

We conducted Canonical Correlation Analysis using the CCA package in the R programming language (11). The CCA (Figure S9) included 60 participants, 5 task variables (1-Back and 3-Back accuracy, cognitive and physical effort cost, overall exit threshold), and 11 survey variables (anhedonia, anxiety, behavioral apathy, cognitive function symptoms, depressed mood/suicidality, social apathy, physical anergia/slowing, Need for Cognition, Effortful Control, BMI, and age). We used the Wilks' Lambda asymptotic test for statistical significant of canonical correlation coefficients (using the p.asym function from the CCP package in R) (16).

12 Additional task measures methods

12.1 Task ability

Using a series of regression models, we tested whether diagnostic groups differed on effortful travel task performance. For the cognitive (N-Back) task we tested for differences in accuracy, reaction time, and missed trials. Using linear regression, we predicted N-Back accuracy (D') by N-Back level interacted with diagnostic group. We used logistic mixed-effects regression to predict N-Back reaction times (log transformed) across all trials by a 4-way interaction between N-Back level, correct or incorrect response, target or non-target trial, and diagnostic group, controlling for age. We used logistic mixed-effects regression to predict the percent of missed N-Back trials by diagnostic group.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

For the physical effort (rapid key-pressing) task we compared the groups on the required number of keypresses (determined during calibration) and the percent of keypresses completed during travel. Using linear regression, we predicted required keypresses by diagnostic group controlling for age. In a linear mixed effects regression, we predicted the percent of completed keypresses per travel interval by the effort level (smaller or larger number of presses) interacted with diagnostic group, controlling for age and BMI.

We tested whether cognitive and physical effort costs were dissociable from task ability (i.e., performance). Using data from all participants, in the first model we predicted cognitive effort cost by 1-Back and 3-Back D', controlling for age. In the second model we predicted cognitive effort cost by the change in D' from 1-Back to 3-Back (which in line with effort cost as a change score from low to high effort). In the third model we predicted physical effort cost by the percent of key presses completed in the larger number of presses, and the smaller number of presses condition, controlling for age and BMI.

12.2 Overall exit threshold

Overall exit threshold individual differences were estimated from a linear mixed effects regression model that predicted exit thresholds (log apples) in low effort orchards (which were least confounded by effects of effort) with effort type as a fixed effect, and random intercepts per participant. To test for diagnostic group differences, we added diagnostic group to the regression. To test for relationships with depression symptoms in the MDD group only, we ran a series of linear regressions predicting overall exit threshold by i) overall depression, and ii) each of the symptom domains separately (7 tests), controlling for age. We corrected for multiple comparisons across symptoms (FDR, 7 tests). We repeated these analyses zooming in to the currently depressed MDD group only and zooming out to all participants.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

	MDD		***Comparison***	
	Mean	**SD(range)**	**Mean**	**SD(range)**
Age (t=-0.083, df=57.2, p=0.934)				
	26.92	11.1(18-61)	27.11	9.64(19-59)
Age, behavioral analyses (t=-0.47, df=53.94, p=0.643)				
Childhood income (t=-2.35, df=61.73, p=0.022)				
	5.07	1.98(1-8)	6.00	1.52(3-9)
Childhood income, behavioral analyses (t=-2.12, df=62.94, p=0.038)				
Years education				
Mother (t=0.086, df=61.8, p=0.931)				
	15.13	3.41(5-20)	15.07	2.73(11-20)
Mother, behavioral analyses (t=0.46, df=62.41, p=0.645)				
Father (t=-1.17, df=61.93, p=0.245)				
	14.8	4.08(0-20)	15.8	3.26(12-20)
Father, behavioral analyses (t=-0.75, df=59.07, p=0.456)				
Self (t=-2.51, df=47.7, p<0.016*)				
	14.7	2.15(9-21)	16.0	2.28(13-20)
Self, behavioral analyses (t=-2.40, df=49.17, p<0.020)				
Education-self correlation with effort costs				
	Correlation	***t***	**df**	***p***
Cognitive effort cost, MDD group				
	-0.03	-0.24	50	0.809
Cognitive effort cost, all participants				
	0.09	0.80	77	0.427
Physical effort cost, MDD group				
	0.15	1.07	51	0.290
Physical effort cost, all participants)				
	0.06	0.53	78	0.599
	MDD		***Comparison***	
	N	**%**	**N**	**%**
Gender (χ^2=2.39, df=2, p=0.303)				
Gender, behavioral analyses (χ^2=3.75, df=2, p=0.154)				
Female	38	63.3%	15	55.6%
Male	19	31.7%	12	44.4%
Non-binary	3	5%	0	0%
Race (χ^2=1.66, df=4, p=0.800)				
Race, behavioral analyses (χ^2=1.63, df=4, p=0.803)				
American Indian/Alaska Native	0	0%	0	0%
Asian	13	21.7%	8	29.6%
Black or African American	8	13.3%	3	11.1%
Native Hawaiian/Pacific Islander	0	0%	0	0%
White	27	45%	13	48.1%
Other or prefer not to say	6	10%	2	7.41%
More than one race	6	10%	1	3.7%
Ethnicity (χ^2=1.41, df=2, p=0.495)				
Ethnicity, behavioral analyses (χ^2=1.71, df=2, p=0.424)				
Latino or Hispanic	8	13.3%	4	14.8%

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

	MDD		***Comparison***	
	N	**%**	**N**	**%**
Not Latino or Hispanic	49	81.7%	23	85.2%
Other or prefer not to say	3	5%	0	0%
Total household income (χ^2=9.44, df=7, p=0.222)				
Total household income, behavioral analyses (χ^2=9.02, df=7, p=0.251)				
Less than $25,000	12	20%	1	3.7%
$25,000 to $34,999	7	11.7%	3	11.1%
$35,000 to $49,999	8	13.3%	2	7.41%
$50,000 to $74,999	9	15%	8	29.6%
$75,000 to $99,999	5	8.33%	5	18.5%
$100,000 to $149,999	6	10%	3	11.1%
$150,000 to $199,999	0	0%	1	3.7%
$200,000 or more	4	6.67%	2	7.41%
Prefer not to answer	9	15%	2	7.41%
Occupational status (χ^2=9.64, df=7, p=0.209)				
Occupational status, behavioral analyses (χ^2=10.39, df=7, p=0.167)				
Working full-time	9	15%	11	40.7%
Working part-time	7	11.7%	2	7.41%
Student full-time	9	15%	3	11.1%
Student part-time	3	5%	0	0%
Working & student	21	35%	9	33.3%
Homemaker	2	3.33%	0	0%
Retired	0	0%	0	0%
Volunteer worker	0	0%	0	0%
Seeking employment	6	10%	2	7.41%
Leave of absence	0	0%	0	0%
Disabled (other free response)	3	5%	0	0%
Relationship status (χ^2=2.94, df=3, p=0.400)				
Relationship status, behavioral analyses (χ^2=4.12, df=3, p=0.248)				
Single	48	80%	21	77.8%
Married	2	3.33%	3	11.1%
Divorced or separated	2	3.33%	0	0%
Widowed	0	0%	0	0%
Other	8	13.3%	3	11.1%
Alcohol frequency (χ^2=2.20, df=4, p=0.700)				
Alcohol frequency, behavioral analyses (χ^2=2.27, df=4, p=0.686)				
Never	15	25%	8	29.6%
Monthly or less	14	23.3%	7	25.9%
2-4 times a month	20	33.3%	5	18.5%
2-3 times a week	9	15%	6	22.2%
4 or more times a week	2	3.33%	1	3.7%
Alcohol amount (χ^2=2.46, df=4, p=0.651)				
Alcohol amount, behavioral analyses (χ^2=2.19, df=4, p=0.702)				
0 (N/A)	15	25%	8	29.6%
1 or 2	20	33.3%	10	37.0%
3 or 4	16	26.7%	5	18.5%
5 or 6	6	10.0%	4	14.8%

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

	MDD		***Comparison***	
	N	**%**	**N**	**%**
7, 8, or 9	3	5.0%	0	0%
10 or more	0	0%	0	0%
Caffeine amount (χ^2=2.59, df=4, p=0.623)				
Caffeine amount, behavioral analyses (χ^2=2.04, df=3, p=0.565)				
None	15	25%	7	25.9%
1 cup	29	48.3%	11	40.7%
2-3 cups	13	21.7%	9	33.3%
4-5 cups	2	3.33%	0	0%
6 or more cups	1	1.67%	0	0%
Tobacco (χ^2=2.71, df=1, p=0.099)				
Tobacco, behavioral analyses (χ^2=2.31, df=1, p=0.129)				
Uses tobacco	12	20%	1	3.7%
Does not use tobacco	48	80%	26	96.3%

Table S5: Demographic factors by diagnostic group. Column 1: demographic factor and response options. Continuous measures shown first with means (columns 2-3) and unpaired t-test statistics (columns 4-6, * indicates p<0.05). Years of education (self) was significantly different between groups, but not correlated with effort costs. Diagnostic groups were matched on all other variables, and this was also true within the subset of participants that were included in the task behavioral analyses (statistical tests denoted by 'behavioral analyses').

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

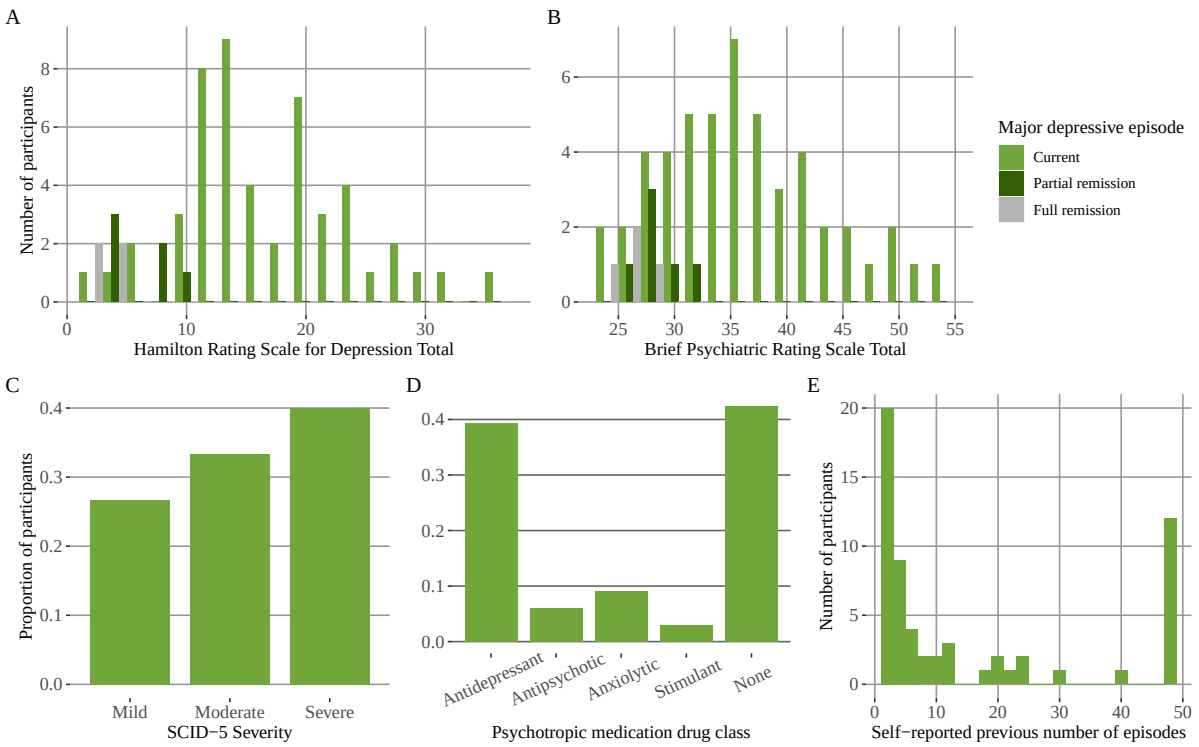

Figure S3: MDD sample characteristics. Histograms, A & B, y-axis: number of MDD participants, x-axis: A: Hamilton Depression Rating Scale total, B: Brief Psychiatric Rating Scale total, fill indicates major depressive episode status (light green indicates current, dark green indicates partial remission, and grey indicates full remission). C, D, & E, y-axis: proportion of MDD participants, C: Structured Clinical Interview for DSM-5 severity rating, D: psychotropic medication drug class, E: self-reported previous number of major depressive episodes.

13 Relationship between cognitive and physical effort costs

Previously, we found a significant positive correlation between cognitive and physical effort costs in a large online study (Experiment 1 (MSIT) of Bustamante et al. (2023), N=537, mean correlation=0.566, 95% HDI=0.355, 0.766). In a smaller undergraduate sample (Experiment 2 (N-Back) of Bustamante et al. (2023), N=81) there was no conclusive evidence for or against the correlation, as the highest density interval (HDI) was wide (mean correlation=0.048, 95% HDI=-0.369, 0.462). The present study uses the same N-Back version of the task as Experiment 2 of the original study with a comparable sample size (N=80 participants in total) and yielded a similarly wide posterior distribution (mean correlation=-0.034, 95% HDI=-0.694, 0.644, Table S6). In both cases the credible interval overlapped with the posterior distribution from Experiment 1, and it may simply be the sample size is underpowered to detect the presence or absence of a correlation. Beyond this, there are several other differences between these task versions, MSIT involves interference control, whereas N-Back involves working memory. Furthermore, the N-Back version is longer in duration (56 versus 32 minutes of main task time), the longer travel time (8.33 seconds versus 20 seconds) requiring more sustained effort. More research is needed to understand under which conditions cognitive and physical effort-based decision making are connected versus dissociated.

Ours and previous research on the relationship between individual differences in cog-

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

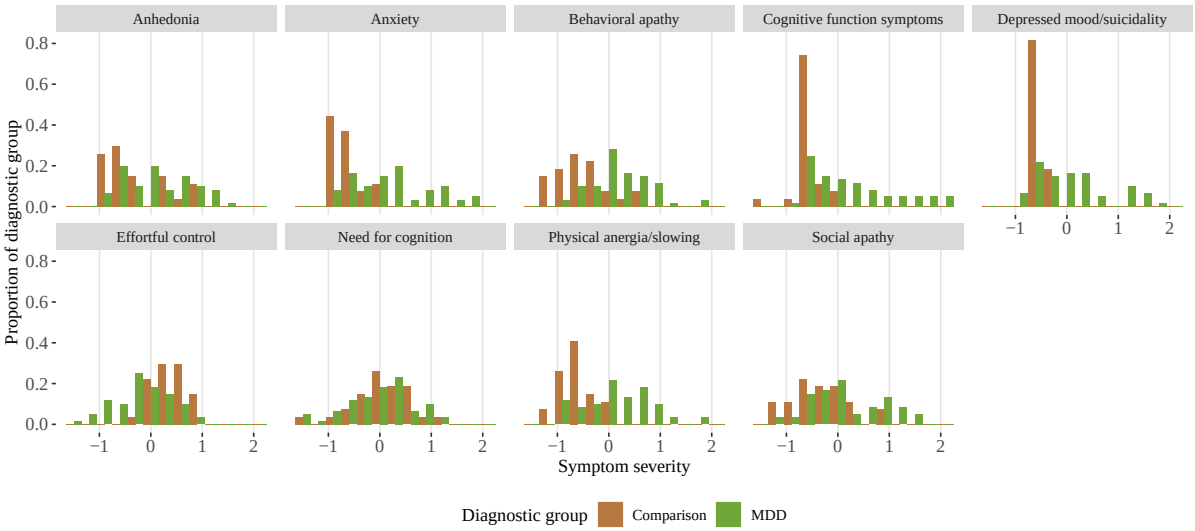

Figure S4: Symptom severity by domain and diagnostic group. Histogram, bar color indicates diagnostic group.

Parameter	Mean	Lower bound	Upper bound	*p*
Inverse Temperature (log)	-1.07	-1.30	-0.83	<0.001
Cognitive Low Effort Travel Cost	119	103	135	<0.001
Cognitive Effort Cost	14.8	8.9	21.3	<0.001
Physical Low Effort Travel Cost	126	110	142	<0.001
Physical Effort Cost	15.2	7.9	22.5	<0.001
Cognitive vs. Physical Effort Cost Correlation	-0.034	-0.694	0.644	>0.461

Table S6: Group-level parameter posterior distribution values. Column 1: parameter, column 2: mean of the group-level posterior distribution, column 3: lower bound of 95% credible interval, column 4: upper bound of credible interval, column 5: Bayesian p-values.

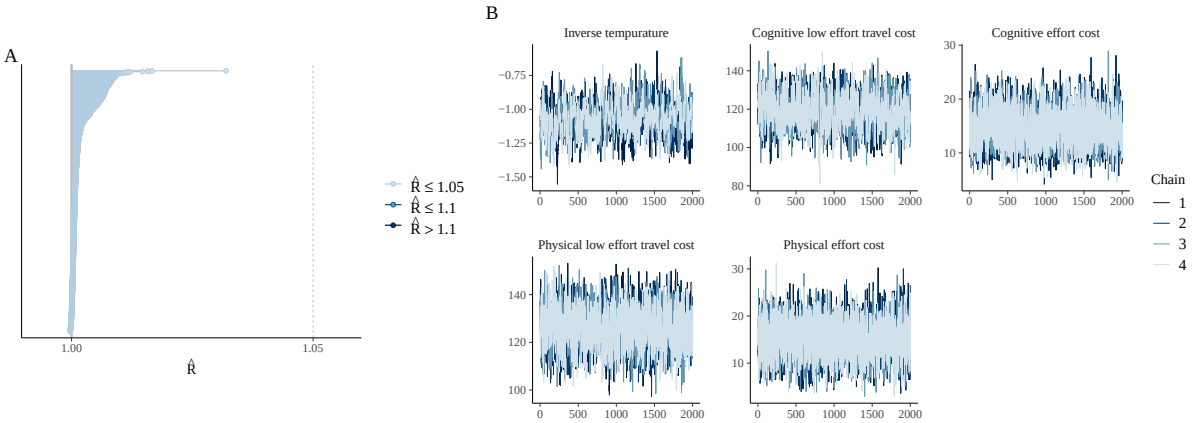

Figure S5: MVT model diagnostics. A: $\hat{R}$ convergence diagnostic plot for all parameters, all of which are below the 1.05 cutoff (vertical dotted line) indicating model convergence. B: MCMC trace plots, color indicates chain number, overlapping traces suggests model convergence.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

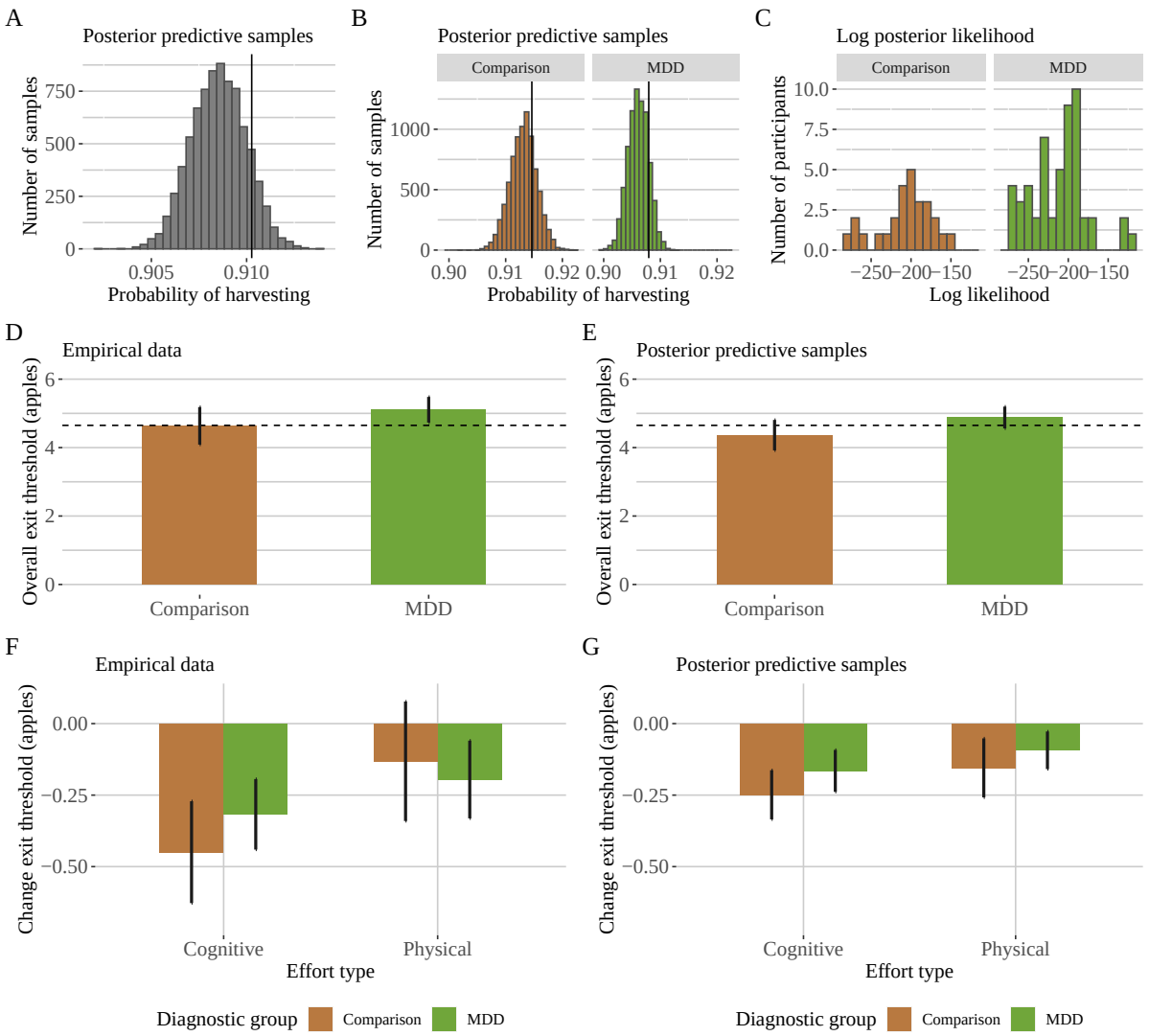

Figure S6: Posterior predictive check results by diagnostic group. A: x-axis, probability of choosing to harvest (1) or exit (0) across all trials, y-axis, number of MCMC samples, empirical observation indicated by vertical black line. B: x-axis, probability of choosing to harvest (1) or exit (0) across all trials for comparison participants (left) and MDD participants (right), y-axis, number of MCMC samples, empirical observations indicated by vertical black lines. C: x-axis, log posterior likelihood per participant for comparison (left) and MDD participants (right), y-axis number of participants. D: empirical data, x-axis, diagnostic group, y-axis, overall exit threshold (from low effort orchards), bars indicate group means, error bars indicate standard error of the mean, horizontal dotted line indicates best threshold from simulation. D: posterior predictive data, for each MCMC sample we computed the overall exit threshold per participant, and aggregated across samples to get the mean value per participant. Resulting plot shows group means and SEM matching C. E. empirical data, x-axis, effort type, y-axis, change exit threshold (high - low effort orchards), bars indicate group means, error bars indicate standard error of the mean, fill indicates diagnostic group. F. posterior predictive data, for each MCMC sample we computed the change in exit threshold per participant, and aggregated across samples to get the mean value per participant. Resulting plot shows group means and SEM matching E.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

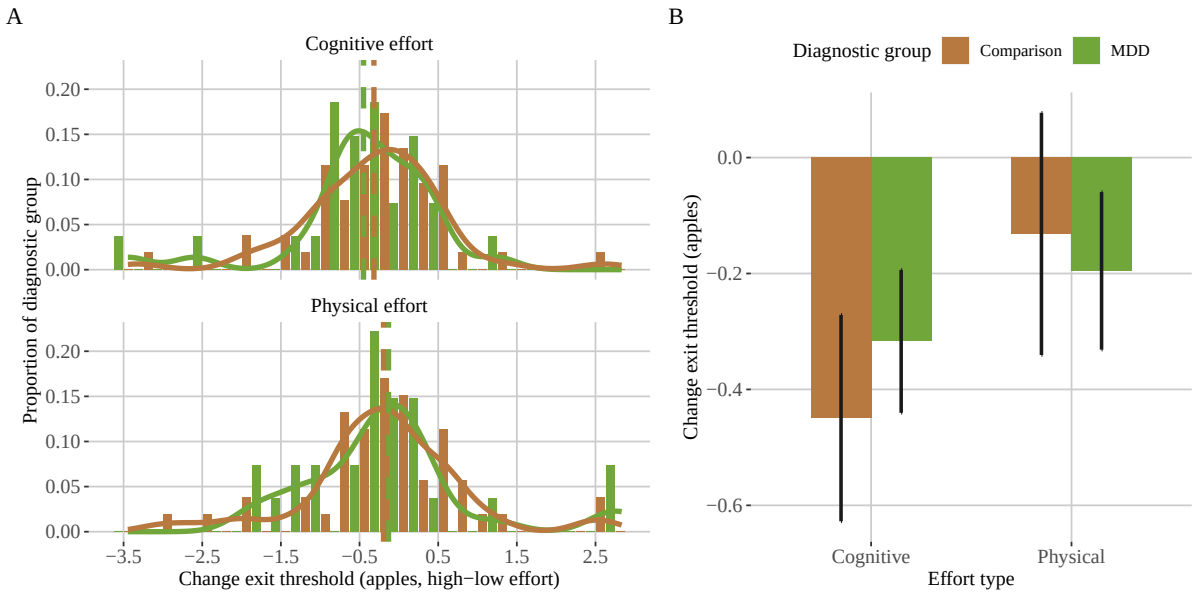

Figure S7: Model-agnostic change in exit threshold by diagnostic group. A: individual differences, error bars indicate 95% HDI, x-axis indicates cognitive effort condition, y-axis indicates physical effort condition. B: diagnostic group differences, x-axis indicates effort type and diagnostic group, y-axis indicates change in the exit threshold (high - low effort, apples).

nitive and physical effort decision making have found moderate correlations (e.g., correlation=0.43 in Lopez-Gamundi & Wardle, correlation=0.35 in Tran et al., 2020). This unshared variance between the effort domains leaves open the possibility of decoupling of their relationship to specific psychiatric symptoms. Indeed in Experiment 1 of Bustamante et al. 2023 (using the MSIT to elicit cognitive effort), we found cognitive, but not physical effort cost loaded strongly onto the dimension predictive of symptoms in a CCA (Dimension 1). Here, we demonstrated differential relationships of cognitive and physical effort cost to symptoms by conducting a comparison of correlations (Meng, Rosenthal, and Rubin's z statistic) (15).

14 Model-agnostic sensitivity to effort manipulation

The model-agnostic measure, change in exit threshold, showed that on average exit thresholds were lower in the cognitive, but not physical, high relative to low effort conditions (3-Back - 1-Back estimate=-0.361 apples, SE=0.103, df=80.59, t=-3.498, $p<0.001$; Larger - Smaller Number of Presses: -0.1809 apples, SE=0.1157, df = 81.541, t=-1.564, p=0.122). There was no reliable interaction between diagnostic group and change in exit threshold for cognitive (t=0.926, df=72.95, $p>0.357$) nor physical effort (t=0.097, df=76.84, $p>0.923$). The MVT model group-level posterior parameters indicated high effort cost is greater than zero for both effort types (Table S6). There was considerable individual variation in willingness to exert effort, signaling differences in perceived effort costs (see Figure 1).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Symptom	**Estimate**	**SE**	***t***	***p***	$p_{adjusted}$
A. Cognitive effort cost, MDD, controlling for medication use					
Diagnostic group difference	1.90	1.53	1.24	0.217	
Overall depression	-0.33	0.10	-3.38	0.001*	
Anhedonia	-3.16	1.13	-2.80	0.008	0.011*
Anxiety	-3.58	0.94	-3.82	0.000	0.003*
Behavioral apathy	-4.01	1.22	-3.29	0.002	0.003*
Social apathy	-1.08	1.10	-0.98	0.330	0.330
Cognitive function symptoms	-2.16	0.93	-2.32	0.025	0.029*
Depressed mood/suicidality	-3.62	1.09	-3.31	0.002	0.003*
Physical anergia/slowing	-3.74	1.10	-3.40	0.001	0.003*
B. Physical effort cost, MDD, controlling for medication use					
Diagnostic group difference	4.11	3.69	1.11	0.269	
Overall depression	0.14	0.27	0.52	0.609	
Anhedonia	8.24	2.94	2.80	0.008*	0.056
Anxiety	0.81	2.75	0.30	0.768	0.768
Behavioral apathy	7.64	3.26	2.34	0.024*	0.084
Social apathy	1.76	2.77	0.64	0.529	0.617
Cognitive function symptoms	4.24	2.66	1.59	0.118	0.275
Depressed mood/suicidality	3.92	2.92	1.34	0.186	0.326
Physical anergia/slowing	2.90	3.26	0.89	0.378	0.529
C. Overall exit threshold, MDD, controlling for medication use					
Diagnostic group difference	0.34	0.16	2.08	0.041*	
Overall depression	-0.03	0.01	-2.60	0.012*	
Anhedonia	-0.28	0.13	-2.16	0.036	0.050*
Anxiety	-0.26	0.11	-2.36	0.022	0.042*
Behavioral apathy	-0.33	0.14	-2.33	0.024	0.042*
Social apathy	-0.09	0.12	-0.79	0.434	0.434
Cognitive function symptoms	-0.17	0.11	-1.58	0.120	0.140
Depressed mood/suicidality	-0.33	0.12	-2.78	0.008	0.027*
Physical anergia/slowing	-0.35	0.12	-2.81	0.007	0.027*

Table S7: Symptoms effort cost regressions, controlling for psychotropic medication use, MDD participants. (A, MDD) Predict cognitive effort cost by overall depression severity and each symptom domain, controlling for cognitive task performance (3-Back D'), medication use, age, and years of education. (B, MDD) Predict physical effort cost by overall depression severity, and each symptom domain, controlling for physical task performance (% larger number of presses completed), BMI, medication use, age and years of education. (C, MDD) Predict overall exit threshold (log apples, from low effort conditions) by symptom severity, medication use, age, and years of education (* indicates $p<0.05$, FDR correction within symptom models).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

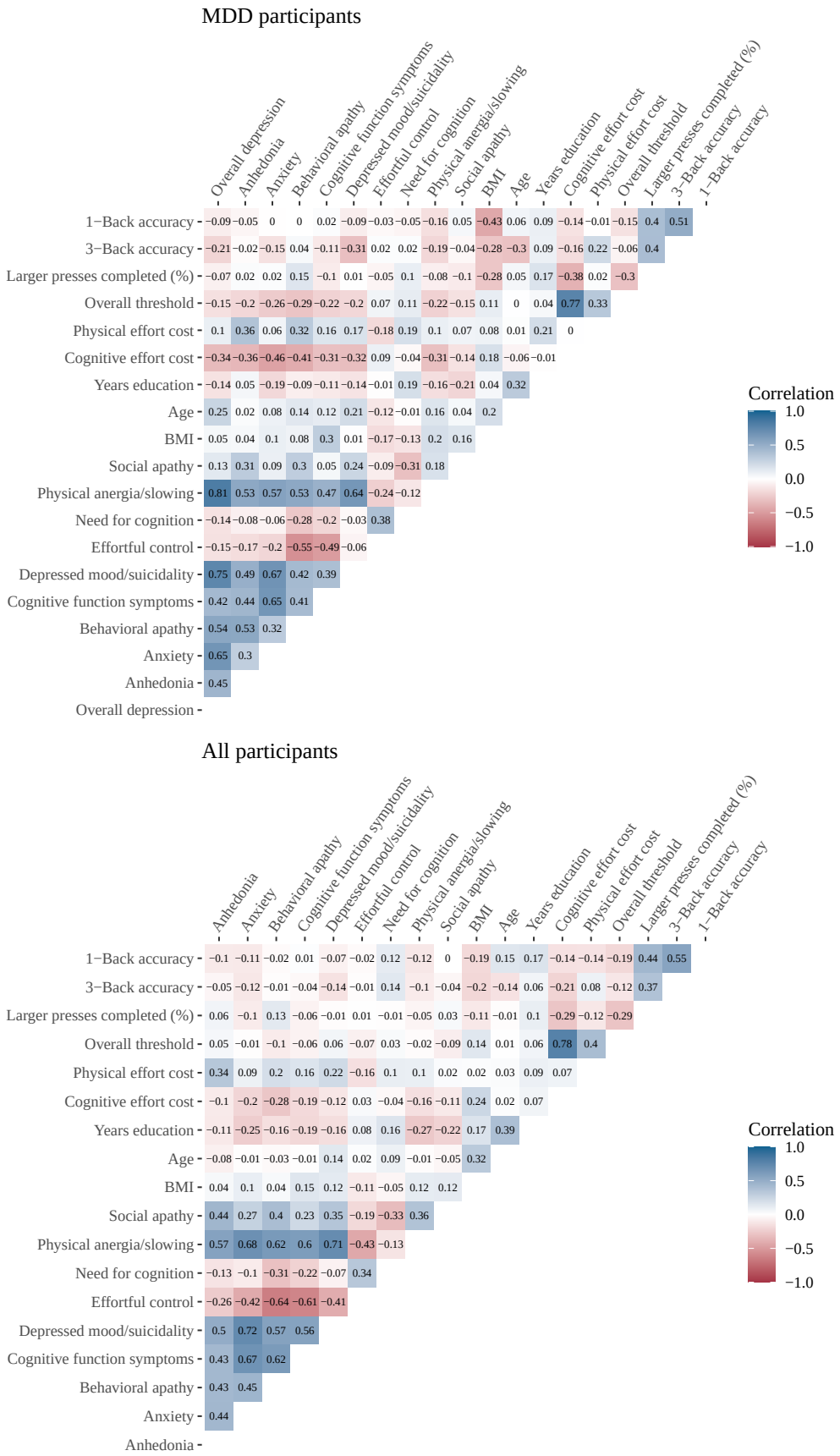

Figure S8: Task behavior symptom heatmap. Left MDD group only, right all participants (spearman correlation matrix).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Diagnostic group effect parameter	Mean	Lower	Upper	*p*
Cognitive Effort Cost	-0.96	-8.29	6.46	0.39
Physical Effort Cost	0.35	-7.74	8.42	0.46
Cognitive Low Effort Cost	1.96	-7.43	10.97	0.34
Physical Low Effort Cost	-1.93	-11.05	7.35	0.34
Inverse temperature (log)	-0.12	-0.50	0.26	0.27
Current depressed group effect parameter	Mean	Lower	Upper	*p*
Cognitive Effort Cost	-1.94	-10.71	6.64	0.32
Physical Effort Cost	-1.25	-10.09	7.37	0.60
Cognitive Low Effort Cost	1.96	-7.48	11.41	0.34
Physical Low Effort Cost	-1.80	-11.40	8.16	0.36
Inverse temperature (log)	-0.17	-0.53	0.18	0.17

Table S8: Diagnostic group difference MVT model. Group effect parameter for model that included all MDD participants. Current depressed group effect parameter for model that excluded participants in remission.

Predicted variable	**Estimate**	**SE**	***t***	***p***
Cognitive effort cost	-0.48	0.12	-3.93	0.000
Physical effort cost	0.14	0.33	0.43	0.668
1-Back D’	-1.05	1.14	-0.92	0.361
3-Back D’	-0.79	1.19	-0.66	0.512
Smaller number of presses completed (%)	0.04	0.60	0.06	0.951
Larger number of presses completed (%)	-0.12	0.13	-0.91	0.369
Overall exit threshold	-0.04	0.01	-3.28	0.002

Table S9: Self-reported overall depression regression results. Column 1: predicted variable, self-reported overall depression (PHQ-9) was used as a predictor variable in place of clinician-rated depression (HAMD). Results correspond to patterns identified using clinician rated depression.

15 CCA results

The CCA Figure S9 yielded one significant dimension using the Wilks’ Lambda asymptotic test for statistical significance of canonical correlation coefficients (Dimension 1 correlation=0.688, Wilks statistic=0.18, F-approx=1.66, df1=55, df2=207, p=0.006, Dimension 2 correlation=0.608, Wilks statistic=0.35, F-approx=1.37, df1=55, df2=172, p=0.085, Dimension 3 correlation=0.476, p=0.337, Dimension 4 correlation=0.463, p=0.412, Dimension 5 correlation=0.290, p=0.729, (16),using the p.asym function from the CCP package in R).

16 Additional task measures results

Across all the travel task measures tested we found few reliable diagnostic group differences (see Figure S10 and Table S11), including no difference in missed N-Back trials, cognitive task accuracy (D’), nor required keypresses determined in the calibration phase. We found the MDD group responded faster on average on the cognitive (N-Back) task (Figure S10). We found a significant effect of diagnostic group on percent of completed presses, in which the MDD group completed a larger percent of presses across conditions, but a significant diag-

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Symptom	**Estimate**	**SE**	***t***	***p***	$p_{adjusted}$
A. Cognitive effort cost, current MDD participants					
Overall depression	-0.37	0.13	-2.77	0.009*	
Anhedonia	-2.50	1.40	-1.78	0.083	0.097
Anxiety	-3.71	1.11	-3.35	0.002	0.013*
Behavioral apathy	-3.80	1.42	-2.68	0.011	0.026*
Social apathy	-1.58	1.30	-1.22	0.232	0.232
Cognitive function symptoms	-2.14	1.15	-1.86	0.071	0.097
Depressed mood/suicidality	-3.71	1.26	-2.96	0.005	0.019*
Physical anergia/slowing	-3.42	1.41	-2.42	0.021	0.036*
B. Physical effort cost, current MDD participants					
Overall depression	0.36	0.34	1.06	0.298	
Anhedonia	9.89	3.12	3.17	0.003	0.021*
Anxiety	1.41	2.97	0.47	0.638	0.744
Behavioral apathy	7.08	3.38	2.09	0.043	0.117
Social apathy	0.36	3.05	0.12	0.907	0.907
Cognitive function symptoms	6.02	2.97	2.03	0.050	0.117
Depressed mood/suicidality	4.30	3.10	1.39	0.174	0.291
Physical anergia/slowing	4.79	3.74	1.28	0.208	0.291
C. Cognitive effort cost, all participants					
Anhedonia	-0.79	0.95	-0.82	0.412	0.481
Anxiety	-2.07	0.84	-2.45	0.016	0.047*
Behavioral apathy	-2.29	0.96	-2.38	0.020	0.047*
Social apathy	-0.26	0.90	-0.29	0.774	0.774
Cognitive function symptoms	-1.33	0.76	-1.75	0.083	0.116
Depressed mood/suicidality	-2.40	0.98	-2.45	0.017	0.047*
Physical anergia/slowing	-1.93	0.95	-2.03	0.046	0.080
D. Physical effort cost, all participants					
Anhedonia	7.05	2.19	3.22	0.002	0.014*
Anxiety	1.08	2.15	0.50	0.617	0.703
Behavioral apathy	3.66	2.44	1.50	0.137	0.392
Social apathy	0.83	2.17	0.38	0.703	0.703
Cognitive function symptoms	2.37	1.93	1.23	0.224	0.392
Depressed mood/suicidality	3.31	2.44	1.35	0.180	0.392
Physical anergia/slowing	2.02	2.49	0.81	0.420	0.588

Table S10: Symptoms effort cost regressions, current MDD only, and all participants. (A, current MDD, C, all participants) Predict cognitive effort cost by overall depression severity and each symptom domain, controlling for cognitive task performance (3-Back D'), age, and years of education. (B, current MDD, D, all participants) Predict physical effort cost by overall depression severity, and each symptom domain, controlling for physical task performance (% larger number of presses completed), BMI, age and years of education (* indicates $p<0.05$, FDR correction within symptom models).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

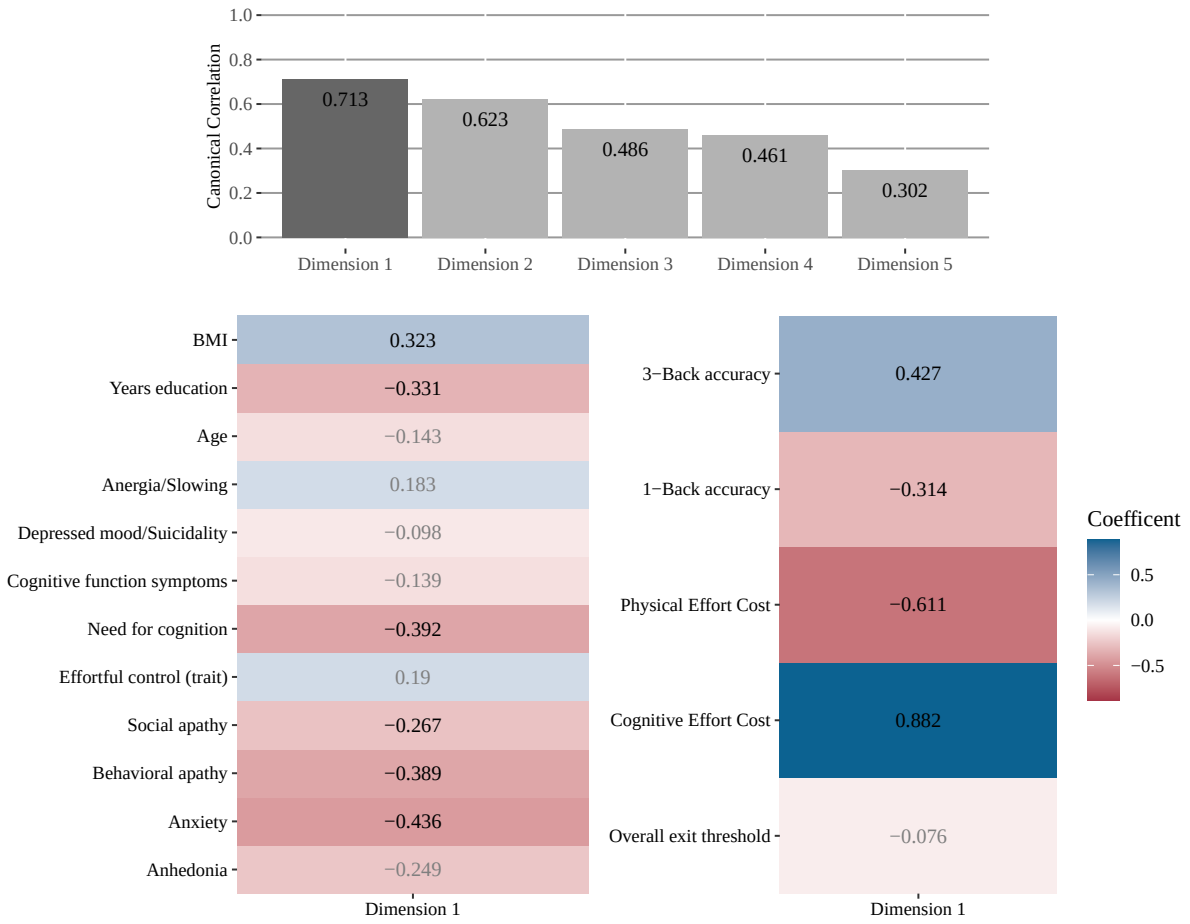

Figure S9: Canonical Correlation Analysis. MDD symptom intensity relationships to effort cost (MDD group only). The first dimension was significant, while all other dimensions were not. For the dimension 1 task coefficients, cognitive effort cost was positive, physical effort cost was negative, and 3-Back D' was positive. For the dimension 1 symptom coefficients behavioral apathy, need for cognition, anhedonia, and anxiety were all negative. This pattern suggests cognitive effort cost is negatively associated with symptom severity on this dimension, while physical effort cost is positively associated with symptom severity. The second dimension is displayed for completeness.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Predicted variable	**Predictor variable**	**Estimate**	**SE**	**df**	***t***	***p***
N-Back missed trials (%)	Group	-0.021	0.495	77	-0.042	0.966
N-Back RT (log)	Group	-0.163	0.123	78.63	-1.35	0.181
N-Back RT (log)	Group*correct*N-Back	-0.197	0.084	4955	-2.335	0.020
N-Back RT (log)	Group*correct* target	-0.273	0.009	2450	-3.206	0.001
Accuracy (D')	Group	-0.008	0.222	113.6	-0.034	0.973
Accuracy (D')	Group*N-Back level	0.117	0.202	77	0.577	0.566
Required keypresses	Group	-2.00	3.640	77	-0.550	0.584
Completed presses (%)	Group	0.272	0.970	101	0.281	0.780
Completed presses (%)	Group*effort level	-2.379	0.675	3237.4	-3.524	0.0004
Completed larger presses	Group	-1.490	1.510	79.33	-0.987	0.327
Completed smaller presses	Group	-0.010	0.358	1650	-0.027	0.978

Table S11: Travel task performance diagnostic group differences. Column 1: predicted variable in regression, column 2: predictor variable, column 3: regression estimate, column 4: standard error (SE), column 5: degrees of freedom, column 6: t-statistic, column 7: p-value.

nostic group by effort level interaction, in which the MDD group completed fewer presses. To decompose this effect, we ran the same regression separately for each effort level. There was no reliable group effect on the percent of smaller number of presses completed, but the MDD group completed a lower percent of required keypresses in the larger press condition.

We found that cognitive and physical effort costs were dissociable from task performance in the Effort Foraging Task, this may suggest a disconnect between effort selection and effort execution (as suggested in, (18)). There was no reliable association between cognitive effort cost and cognitive task performance (Figure S11, model 1: 3-Back D', $p>0.22$, 1-Back D', $p>0.79$, all participants, model 2: change D', $p>0.53$). Likewise, there was no reliable relationship between physical effort cost and the percent of key presses completed (model 3: larger number of presses, $p>0.91$, smaller number of presses, $p>0.20$, consistent with that was found in (6)).

Neither cognitive nor physical performance was reliably related to overall depression (predict Hamilton Depression Rating Scale Total controlling for age by 3-Back D', $p>0.16$, 1-Back D', $p>0.86$). While anxiety symptoms were associated with cognitive effort costs, they were not associated with cognitive task performance (MDD group predict by anxiety symptoms controlling for age by 3-Back D', $p>0.21$, 1-Back D', $p>0.79$). On the other hand, anhedonia symptoms were related to a lower percentage of completed keypresses in the low ($t=2.68$, $p<0.010$), but not high ($p>0.311$) physical effort condition.

Overall depression was not related to cognitive task performance (Figure S11, predict HAMD total controlling for age by 3-Back D', $p>0.16$, 1-Back D', $p>0.86$) nor for physical task performance (predict HAMD total controlling for age and BMI by Larger number of presses, $p>0.073$, Smaller number of presses, $p>0.370$). While anxiety symptoms were associated with cognitive effort costs, they were not associated with cognitive task performance (MDD group predict by anxiety symptoms by 3-Back D', controlling for age, $p>0.21$, and 1-Back D', $p>0.79$). Anhedonia symptoms were related to the percent of smaller number of presses completed ($t=2.70$, $p<0.010$) but not to the percent of larger number of presses ($p>0.318$).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

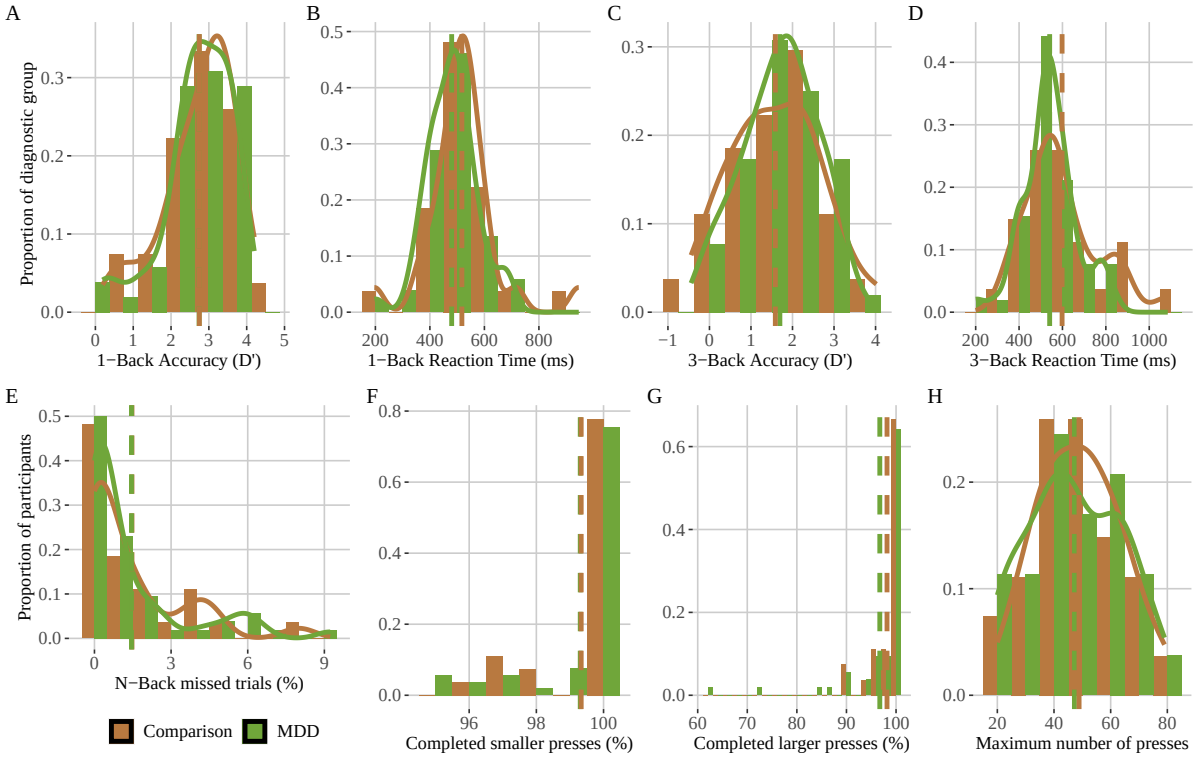

Figure S10: Diagnostic group difference in travel task performance. A: 1-Back accuracy (D'), B: 1-Back Reaction Time (ms), C: 3-Back accuracy (D'), D: 3-Back Reaction Time (ms), E: percent of N-Back trials missed responding before the RT deadline, F: Physical low effort performance (percent completed smaller number of presses), G: Physical high effort performance (percent completed larger number of presses), H: Maximum number of presses determined in a calibration phase. There were no diagnostic group differences except that the MDD group responded faster on average on the cognitive (N-Back) task (A & C), and completed fewer presses in the high effort physical condition (G).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

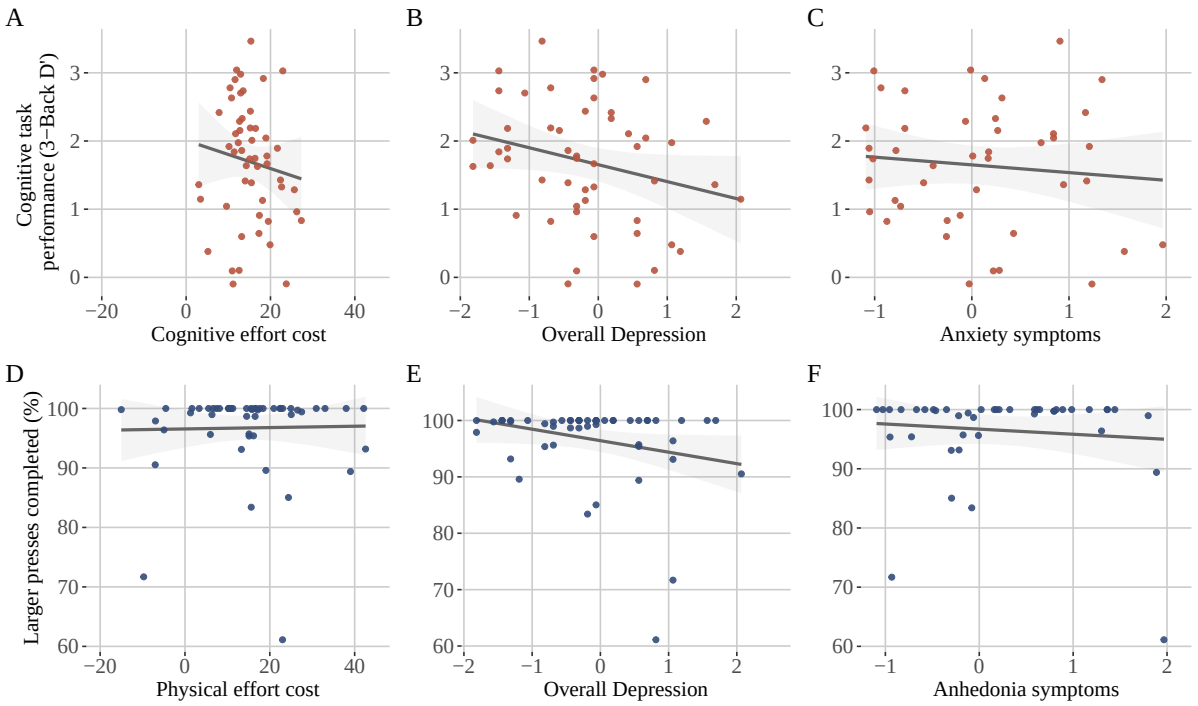

Figure S11: Travel task performance relationship to effort costs, overall depression, anhedonia, and anxiety. Only MDD group. Cognitive travel task performance (3-Back Accuracy (D'), y-axis) not reliably related to, A: cognitive effort cost, B: overall depression (HAMD, z-score), nor C: anxiety symptoms (x-axis, z-score). Physical travel task performance (percent of larger number of presses completed, y-axis) not reliably related to, D: physical effort cost, E: overall depression, nor F: anxiety symptoms (x-axis).

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Symptom	Estimate	SE	*t*	*p*	$p_{adjusted}$
A. Overall exit threshold, current MDD					
Overall depression	-0.03	0.01	-2.30	0.027*	
Anhedonia	-0.23	0.15	-1.51	0.138	0.193
Anxiety	-0.25	0.13	-1.94	0.060	0.122
Behavioral apathy	-0.30	0.16	-1.86	0.070	0.122
Social apathy	-0.16	0.14	-1.20	0.239	0.279
Cognitive function symptoms	-0.12	0.13	-0.97	0.336	0.336
Depressed mood/suicidality	-0.36	0.13	-2.72	0.010*	0.068
Physical anergia/slowing	-0.33	0.15	-2.15	0.038*	0.122
B. Overall exit threshold, all participants					
Anhedonia	-0.03	0.10	-0.26	0.796	0.796
Anxiety	-0.10	0.10	-1.07	0.288	0.503
Behavioral apathy	-0.14	0.11	-1.25	0.213	0.503
Social apathy	-0.03	0.10	-0.29	0.769	0.796
Cognitive function symptoms	-0.08	0.09	-0.89	0.374	0.524
Depressed mood/suicidality	-0.18	0.11	-1.67	0.099	0.503
Physical anergia/slowing	-0.13	0.11	-1.22	0.228	0.503

Table S12: Overall exit threshold relationship to symptoms (current MDD group, and all participants). Predicting individual differences in overall exit thresholds (log, from low effort conditions) by overall depression severity, and each symptom domain, controlling for age and years of education (* indicates $p<0.05$, FDR correction within symptom models).

References

[1] Yuen-Siang Ang, Patricia Lockwood, Matthew A. J. Apps, Kinan Muhammed, and Masud Husain. Distinct Subtypes of Apathy Revealed by the Apathy Motivation Index. *PloS One*, 12(1):e0169938, 2017.

[2] Laura A. Bustamante, Temitope Oshinowo, Jeremy R. Lee, Elizabeth Tong, Allison R. Burton, Amitai Shenhav, Jonathan D. Cohen, and Nathaniel D. Daw. Effort Foraging Task reveals positive correlation between individual differences in the cost of cognitive and physical effort in humans. *Proceedings of the National Academy of Sciences*, 120(50):e2221510120, December 2023. Publisher: Proceedings of the National Academy of Sciences.

[3] J. T. Cacioppo, R. E. Petty, and C. F. Kao. The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3):306–307, June 1984.

[4] Stephane Champely, Claus Ekstrom, Peter Dalgaard, Jeffrey Gill, Stephan Weibelzahl, Aditya Anandkumar, Clay Ford, Robert Volcic, and Helios De Rosario. pwr: Basic Functions for Power Analysis, March 2020.

[5] J. Cohen. A power primer. *Psychological Bulletin*, 112(1):155–159, July 1992.

[6] Adam J Culbreth, Erin K Moran, Wasita Mahaphanit, Molly A Erickson, Megan A Boudewyn, Michael J Frank, Deanna M Barch, Angus W MacDonald, III, J Daniel Ragland, Steven J Luck, Steven M Silverstein, Cameron S Carter, and James M

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Gold. A Transdiagnostic Study of Effort-Cost Decision-Making in Psychotic and Mood Disorders. *Schizophrenia Bulletin*, page sbad155, October 2023.

[7] David E. Evans and Mary K. Rothbart. Developing a model for adult temperament. *Journal of Research in Personality*, 41(4):868–888, August 2007.

[8] Maurizio Fava, Dan V. Iosifescu, Paola Pedrelli, and Lee Baer. Reliability and validity of the Massachusetts general hospital cognitive and physical functioning questionnaire. *Psychotherapy and Psychosomatics*, 78(2):91–97, 2009.

[9] Michael B. First. Structured Clinical Interview for the DSM (SCID). In *The Encyclopedia of Clinical Psychology*, pages 1–6. John Wiley & Sons, Ltd, 2015. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118625392.wbecp351.

[10] Jonah Gabry, Tristan Mahr, Paul-Christian Bürkner, Martin Modrák, Malcolm Barrett, Frank Weber, Eduardo Coronado Sroka, Teemu Sailynoja, and Aki Vehtari. bayesplot: Plotting for Bayesian Models, February 2024.

[11] Ignacio González and Sébastien Déjean. CCA: Canonical Correlation Analysis, 2021.

[12] M. Hamilton. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23:56–62, February 1960.

[13] K. Kroenke, R. L. Spitzer, and J. B. Williams. The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9):606–613, September 2001.

[14] Daniel Lewandowski, Dorota Kurowicka, and Harry Joe. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001, October 2009.

[15] Xiao-li Meng, Robert Rosenthal, and Donald B. Rubin. Comparing correlated correlation coefficients. *Psychological Bulletin*, 111:172–175, 1992. Place: US Publisher: American Psychological Association.

[16] Uwe Menzel. CCP: Significance Tests for Canonical Correlation Analysis (CCA), 2022.

[17] Paul A. Nakonezny, Thomas J. Carmody, David W. Morris, Benji T. Kurian, and Madhukar H. Trivedi. Psychometric evaluation of the Snaith-Hamilton pleasure scale in adult outpatients with major depressive disorder. *International Clinical Psychopharmacology*, 25(6):328–333, November 2010.

[18] Randall C. O'Reilly, Thomas E. Hazy, Jessica Mollick, Prescott Mackie, and Seth Herd. Goal-Driven Cognition in the Brain: A Computational Framework, April 2014. arXiv:1404.7591 [q-bio].

[19] John E. Overall and Donald R. Gorham. The Brief Psychiatric Rating Scale. *Psychological Reports*, 10(3):799–812, June 1962. Publisher: SAGE Publications Inc.

[20] RCoreTeam. R: A language and environment for statistical computing, 2015.

[21] Robert L. Spitzer, Kurt Kroenke, Janet B. W. Williams, and Bernd Löwe. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10):1092–1097, May 2006.

medRxiv preprint doi: https://doi.org/10.1101/2024.02.18.24302985; this version posted June 21, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[22] Stan. CmdStanR: the R interface to CmdStan — cmdstanr-package, 2021.

[23] Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, 2021.

[24] David Wechsler. Wechsler Test of Adult Reading: WTAR. *Psychological Corporation*, 2001.