Neural systems underlying the learning of cognitive effort costs

Ceyda Sayalı[1] and David Badre[1,2]

[1]Department of Cognitive, Linguistic, and Psychological Sciences
[2]Brown Institute for Brain Sciences
Brown University

Address correspondence to:
Ceyda Sayalı
Box 1821
Brown University
Providence, RI 02912-1978
tel 401-863-1544, fax 401-863-2255
email ceyda_sayali@alumni.brown.edu

Competing Interests Statement
The authors declare that there are no conflicts of interest.

Abbreviated title: Neural Systems of Effort Learning
Number of pages: 38
Number of figures: 9
Number of tables: 1
Word counts:
Abstract: 249 words
Introduction: 648 words
Discussion: 1432 words

Acknowledgements: We would like to thank Louise Stolz, Jordan Rubin-McGregor and Adriane Spiro for assisting in data collection, and Eliana Vassena for her helpful comments. This work was supported by a MURI from the Office of Naval Research (N00014-16-1-2832) and a grant from James S. McDonnell Foundation.

Abstract

People balance the benefits of cognitive work against the costs of cognitive effort. Models that incorporate prospective estimates of the costs of cognitive effort into decision making require a mechanism by which these costs are learned. However, it remains open what brain systems are important for this learning, particularly when learning is not tied explicitly to a decision about what task to perform. In this fMRI experiment, we parametrically manipulated the level of effort a task requires by increasing task switching frequency across six task contexts. In a scanned learning phase, participants implicitly learned about the task switching frequency in each context. In a subsequent test phase outside the scanner, participants made selections between pairs of these task contexts. Notably, during learning, participants were not aware of this later choice phase. Nonetheless, participants avoided task contexts requiring more task switching. We modeled learning within a reinforcement learning framework, and found that effort expectations that derived from task-switching probability and response time (RT) during learning were the best predictors of later choice behavior. Interestingly, prediction errors (PE) from these two models were differentially associated with separate brain networks during distinct learning epochs. Specifically, PE derived from expected RT was most correlated with the cingulo-opercular network early in learning, whereas PE derived from expected task switching frequency was correlated with the fronto-parietal network late in learning. These observations are discussed in relation to the contribution of cognitive control systems to new task learning and how this may bear on effort-based decisions.

Keywords: Cognitive effort, effort avoidance, cognitive control, effort cost, reward

Significance Statement

On a daily basis, we make decisions about cognitive effort expenditure. It has been argued that we avoid cognitively effortful tasks to the degree subjective costs outweigh the benefits of the task. Here, we investigate the brain systems that learn about task demands for use in later effort-based decisions. Using reinforcement learning models, we find that learning about both expected response time and task switching frequency affect later effort-based decisions and these are differentially tracked by distinct brain networks during different epochs of learning. The results indicate that more than one signal is used by the brain to associate effort costs with a given task.

Introduction

People avoid cognitively effortful tasks when given the option (Kool et al., 2010; Westbrook et al., 2013). Based on this observation, it has been proposed that effortful tasks incur a subjective cost that acts as a disutility to drive demand avoidance. Further, in order to balance the expected effort cost of a task against its expected rewards, people must learn to predict the level of cognitive effort required to successfully perform a task (Shenhav et al., 2013; Botvinick, 2007).

In the laboratory, demand avoidance has been tested using effort discounting paradigms. In these tasks, participants learn the association between a unique task identifier and a level of subjective task difficulty during a learning phase. In the following decision phase, participants engage in repeated decision-making between an easy task paired with smaller reward or a difficult task paired with higher reward. Several fMRI studies showed that people discount the value of an offer as a function of their subjective cost during the decision phase (Massar et al., 2015; Chong et al., 2017; Westbrook et al., 2019; Westbrook et al., 2020). In this context, a recent fMRI study found that a network of brain regions that have been shown to track subjective value (SV) in other decision making contexts, also tracked costs and benefits during effort-based decision-making (Westbrook et al., 2019). This study concluded that effort is indeed encoded as a SV and tracked by general purpose reward-related brain regions during effort-based decision-making. This SV network included core nodes in the ventromedial prefrontal cortex (vmPFC) and ventral striatum, as well as mid-dorsal Anterior Cingulate Cortex (dACC). In contrast to this value network, control-related brain regions of the fronto-parietal network (FPN) tracked the decision difficulty between offer options, but not the SV of the task.

These observations were partially consistent with those from a previous fMRI study of effort based decision making (Sayali & Badre, 2019). In that study, we separated learning and decision making phases of an effort selection task. In this task, participants implicitly learned about the cost associated with each of six parametrically increasing effort levels, without being informed about the subsequent decision phase. During the decision phase, they selected which task to perform based on this prior experience. As with Westbrook et al. (2019), we found that during the decision phase, though FPN activity tracked task difficulty, it was not related to effort based predictions. In our study, the only predictor of effort avoidance during task performance was default mode network activity (DMN). Notably, this network includes the vmPFC which overlaps the SV network highlighted by Westbrook et al. (2019). Though, we did not find evidence of a relationship between other nodes of the SV network and effort avoidance, including the ventral striatum and dACC. Overall, however, studies of effort-based decisions appear to agree that SV-related regions, like vmPFC, encode effort costs, while other networks involved in cognitive control, like the FPN, track difficulty but not the subjective cost driving avoidance decisions.

Importantly, however, these prior studies have focused on the decision stage, when people are familiar with the tasks involved and are weighing costs and benefits of cognitive effort. Few experiments have tested how the effort costs associated with tasks themselves are acquired. The one fMRI experiment to date to focus on effort learning observed that performance-based prediction errors drove the acquisition of effort costs (Nagase et al., 2018). Further, expected costs positively activated dACC, as well as vmPFC, a region in the SV network. However, in this task, effort decisions were made while effort costs were also being learned. This still leaves open the possibility that these regions might be tracking the effort costs because an explicit decision about task value is due. As such, it is not presently known if the SV network is involved when learning effort costs or tracking task value in the absence of a decision about what task to perform.

2. Methods

Overview of the Current Study

The current study asks what neural mechanisms underlie the implicit learning of effort costs from task experience, while in the absence of effort-based decisions. We hypothesize that 1) effort costs are learned on a trial-by-trial basis and predict later effort selections, 2) performance-based prediction errors drive the learning of effort costs, 3) SV brain regions will track expected effort costs. In order to test these hypotheses, we estimated expected costs and prediction errors during an implicit learning phase based on the cost prediction error model of effort. We then tested correlates of these parameters using model-based fMRI analysis of the brain during learning.

2.1. Participants

Three-hundred and seventeen adults were pre-screened in the lab using the Need for Cognition (NfC) Inventory. All participants were compensated for their time either monetarily or with course credit. We only included participants scoring 16/72 or lower on the NfC inventory. Based on our prior work (Sayali and Badre, 2019), this screening procedure increases the probability of including demand avoiding participants. We screened the remaining 119 participants based on neurological or psychiatric diagnosis, drug use, or contraindication for MRI. 78 participants met the fMRI eligibility criteria, agreed to be re-contacted, and had not participated in a previous study that tested effort-based decision-making. From these, 29 agreed to participate in our fMRI study. However, 7 of these cancelled their appointment on the day of their scan or withdrew. Given that our participants were prescreened as highly demand avoiding, this rate of withdrawal was not unexpected. However, we consider the implications of this sampling procedure for generalizability of this study in the discussion. Therefore, a total of 22 right-handed adults (aged 18-35; 12 female) who scored lower than 16/72 on NfC, with normal or corrected to-normal vision were scanned in the fMRI experiment. 3 additional participants were excluded prior to analysis because they either failed to follow instructions (2 participants) or moved more than a voxel in all sessions (1 participant). Thus, a total of 19 participants were included in the behavioral and fMRI analyses. Participants provided informed consent in accordance with the Research Protections Office at Brown University.

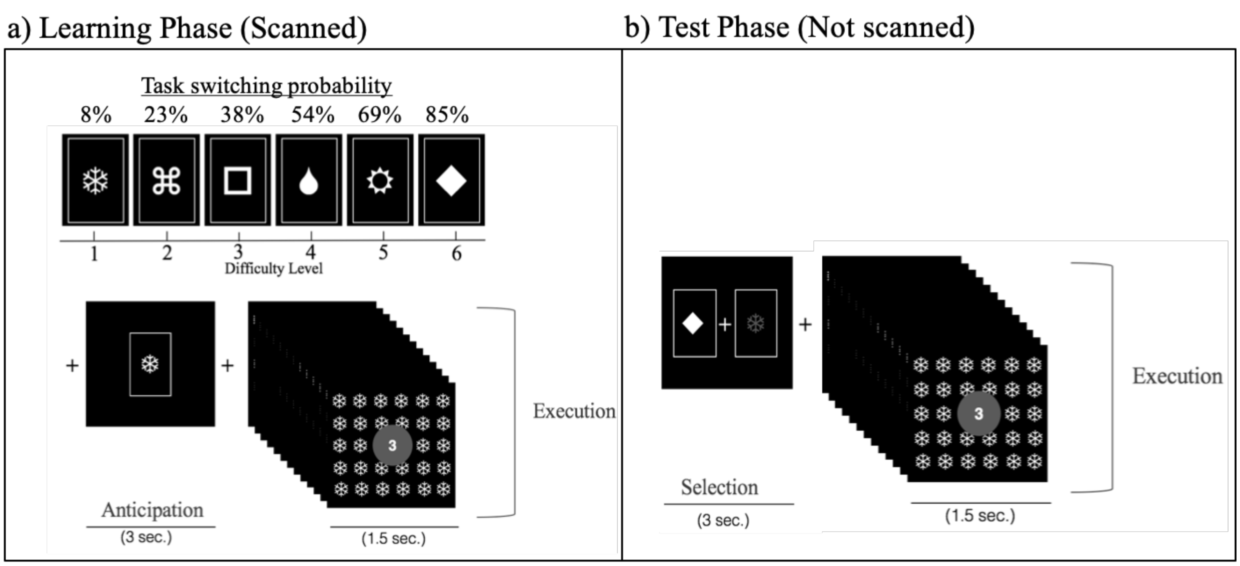

Figure 1. Illustration of the task design. (a) In the Learning Phase, participants learned the association between difficulty levels and corresponding unique task identifiers (top) while being scanned. Anticipation and Execution epochs (bottom) were separately optimized in a hybrid fMRI design. During the Anticipation epoch, participants viewed a symbol icon (a virtual card deck) for 3 seconds. Then, they performed the difficulty level associated with that symbol, while the symbol was tiled on the background. (b) Following the Learning Phase, participants completed the Selection Phase outside the scanner. During a Selection epoch, participants chose between two of the symbols they learned in the Learning Phase. During the Execution epoch, they performed the difficulty level that was associated with their choice.

2.2. Overview of the Behavioral Task

Participants performed a parametric variant of the DST, closely following the procedures from Sayali and Badre (2019; Fig 1). As an overview, the task was performed in two phases. In an initial, scanned Learning phase, the participants performed six tasks and associated each with a virtual "card deck", denoted by a particular symbol that tiled the screen when that task was performed. The six tasks differed from each other in the proportion of task switching required.

More frequent task switching requires greater cognitive control. In a second Test phase, participants chose which deck they wished to perform (Selection epoch) and then performed a series of trials drawn from that deck (Execution epoch). We now provide detail on each of these phases of the behavioral task.

2.2.1. The Task "Decks" and Basic Task Run Structure

Throughout all phases of the experiment, participants performed blocks consisting of 13 trials "drawn" from a virtual deck (Fig 1a). On each trial, the participant categorized a centrally presented digit as either odd/even (parity judgment) or greater/less than 5 (magnitude judgment). The color of a circle (green or blue) surrounding the digit cued whether to perform the parity or magnitude judgment on each trial. The participant indicated their categorization by pressing the left or the right arrow key on the keyboard. The response deadline was 1.5 sec. Trials were separated by .2 sec. The mappings of color to task and from each categorization to a left or right arrow press were provided in an instruction before the experiment and were consistent throughout both phases of the experiment. Response mappings and color-to-task mappings were counterbalanced across participants. Digits 1-4 and 6-9 were used with equal frequency across both tasks and were randomized for order of presentation.

In order to manipulate cognitive effort between the decks, we varied the frequency of task switching required by a particular deck (Fig 1a), as in how often a trial-to-trial transition required going from a parity to magnitude judgment or vice versa. We assumed that more task switches in a deck would also require more cognitive control for that deck and so more cognitive effort. As a short hand for this logic, we will refer to the proportion of task switching in a deck as its *effort level*.

The probability of switching tasks increased across decks over six effort levels: 8%, 23%, 38%, 54%, 69%, 85% (Fig 1a, top). Importantly, all effort levels included the same number of parity and magnitude trials. Thus, any differences in effort preference among the tasks could not be attributed to differences in the difficulty or effort of the categorizations themselves, but rather were attributable to the task switching manipulation. Further, the lowest effort level still included one switch. We did not include "pure blocks" of only parity judgments or only magnitude judgments with zero switches, as doing so would have made these decks qualitatively different from all other decks, in terms of the amount of each task being performed.

At the beginning of each run, a shape was presented for 3 sec to indicate which deck was about to be performed, and this shape was also tiled in the background throughout the performance of the run. Participants were told that this shape represented the symbol on the back of the virtual card deck from which trials for that sequence were drawn. Thus, each effort level could be associated with a particular deck symbol. Participants could learn this relationship through experience with the tasks. The mappings between deck symbols and effort levels were randomized across participants.

2.2.2. Practice Phase

During an initial "Practice Phase" participants gained familiarity with the trial structure, the categorization decisions, and the color and response mappings. After being given instructions regarding the categorization and response rules, they practiced two 13-trial blocks. During the presentation of these blocks, the researcher closely monitored the performance of the participant and assisted the participant if needed. This phase was repeated if the participant asked for a repeat of the task rules, and thus the data from this phase was not recorded.

Each task sequence included the presentation of a deck symbol (3 s), and the subsequent performance of the categorization trials. In the first run of the Practice Phase, feedback was presented after the button press as either 'Correct' in pink, 'Incorrect' in yellow, or 'Please respond faster' if the participant failed to answer before the deadline (1.5s). In the second run of the Practice phase, feedback was omitted as would be the case in the actual experiment. The deck symbols that participants encountered in the Practice Phase were for practice only and were not presented during the Learning or Test phases to avoid any association of difficulty or error feedback with a particular deck symbol during this phase.

2.2.3. Learning Phase

In the Learning Phase (Fig. 1a), participants learned the association between the six deck symbols and each effort level. Each deck was performed 15 times in random order with the other decks. This phase was performed during fMRI scanning, and participants used an MRI-compatible button box to indicate their decisions. The Execution trials were optimized as blocks. The Anticipation and Execution events were separated in time by a jittered time interval (mean 4 secs) so that signal related to each could be analyzed independently. The Learning phase was separated into two, approximately 20 minute-long scanning runs.

2.2.4. Test Phase

In the Test Phase (Fig 1b), two decks were presented and the participant chose which to perform (Selection epoch). The Test Phase took place outside the magnet, in a behavioral testing room immediately following the MRI session. The participants were told to freely choose between decks prior to a 3 sec deadline. We note that in contrast to other DST variants (Gold et al., 2014), participants in this task were not told about the difficulty manipulation to avoid biasing choices based on participants' explicit beliefs about effort. Once the participant made their selection, the selected deck turned blue and both options remained on the screen until the end of the 3 sec deadline. If the participant missed the deadline, the same choice pair was repeated at the end of the experiment until all selections had been made and executed.

Each choice pair was selected from the set of fifteen unique (un-ordered) pair combinations of all six decks, excluding self-pairing (e.g., deck #1 paired with deck #1). Each deck was presented either on the left or the right side of the screen, counterbalanced for location across trials. The Selection epoch was followed by performance of the selected effort level task deck (Execution epoch). The sequence of events, timing, and response mappings during Execution were the same as during the Learning phase. The Test phase was separated into four, approximately 15 minute-long blocks. In each run, each pair was presented 3 times in a pseudo-randomly intermixed order, making a total of 180 decision trials across 4 blocks.

2.3. *Behavioral Data Analysis*

Trials with response times (RT) below 200 ms were excluded from further analysis. Execution trials on which participants missed the response deadline were also excluded (approximately %1 of trials in both phases). Task performance was analyzed with effort level and experimental phases (Learning or Test phase) as within-subject variables. One participant never chose the 5^{th} effort level during the Test Phase, and thus was necessarily excluded from the effort*phase performance analysis (see below).

Choice behavior was assessed by calculating the probability of selecting each effort level across all selection trials on which that effort level was presented as an option. The decision difference analyses calculated the choice probability and the decision time to select the easier task across all decisions with the same difference in difficulty levels between effort options. For example, choice probability associated with a difficulty difference of 1 would be computed by averaging the probability of choosing the easier task across all choice pairs that differed by 1 effort level (i.e., 1 vs 2, 2 vs 3, 3 vs 4, 4 vs 5 and 5 vs 6).

Data were analyzed using a mixed-design analysis of variance (ANOVA) (within subject factor: Effort). If the sphericity assumption was violated, Greenhouse-Geisser correction was used. Significant interactions were followed by simple effects analysis, the results of which are presented with False Detection Rate (FDR) correction. Alpha level of .05 was used for all analyses. Error bars in all figs stand for within-subject error.

2.3.1. Computational model fitting

To test our hypothesis regarding effort cost learning, we modeled the acquisition of effort associations within a reinforcement learning framework (Sutton and Barto, 2018; Daw, 2011). We specifically used a prediction error (PE)-based model that has been applied during an effort selection task (Nagase et al., 2018). Following this prior work, we tested separate models that learned effort costs from prediction errors in 1) task-switching probability, 2) response time during effort execution, 3) error rates during effort execution. We tested two models that relied on prediction error in response time, as described below. Thus, overall, we tested four models of effort cost learning.

2.3.1.1. Task Switch Cost Model

The Task Switch Cost Model attempts to learn the likelihood of a switch based on the context.

To begin, the model assumes an even likelihood of a switch or repeat and so the expected cost on the first trial was initialized to 0.5. Then, on each effort run of the Learning Phase, expected cost of the effort task is updated by a PE of that effort level:

$$ExpectedCost(t) = ExpectedCost(t-1) + \alpha * \left(PredictionError(t)\right),$$

PE on each effort run was computed as the difference between the expected cost (i.e., the likelihood of a task switch) and the experienced cost on trial, t (i.e., the actual task-switching likelihood of the effort level):

$$PredictionError(t) = ActualCost(t) - ExpectedCost(t-1),$$

The rate at which the Expected cost is updated as a function of PE is controlled by a free parameter, α, which represents the learning rate of the participant. Given that our design did not include effort selections during the time of effort learning, we did not estimate learning rates based on a fit of effort selections. Previous simulation studies (Wilson & Niv, 2015) showed that model-based fMRI is insensitive to the change of learning rate parameters, as most experimental designs do not have the statistical power to detect differences in model parameters, specifically when the contrast-to-noise ratio and the number of trials is relatively low.

Consistent with this observation, we simulated learning rates between 0.2 and 0.7, in steps of 0.1 to test the effect of learning rates on the estimated BOLD responses of the SV Network ROI (Westbrook et al., 2019) for the expected cost regressor in Linear-Effort Level GLM model (see below for description of these models and ROIs). The results showed that there was no effect of learning rate parameters on the estimated BOLD response (Fig 2; $F(1.01,18.24)=0.75$, = .39, η_p^2 =.04). Consequently, we adopted a learning rate of 0.2 for the following analysis.

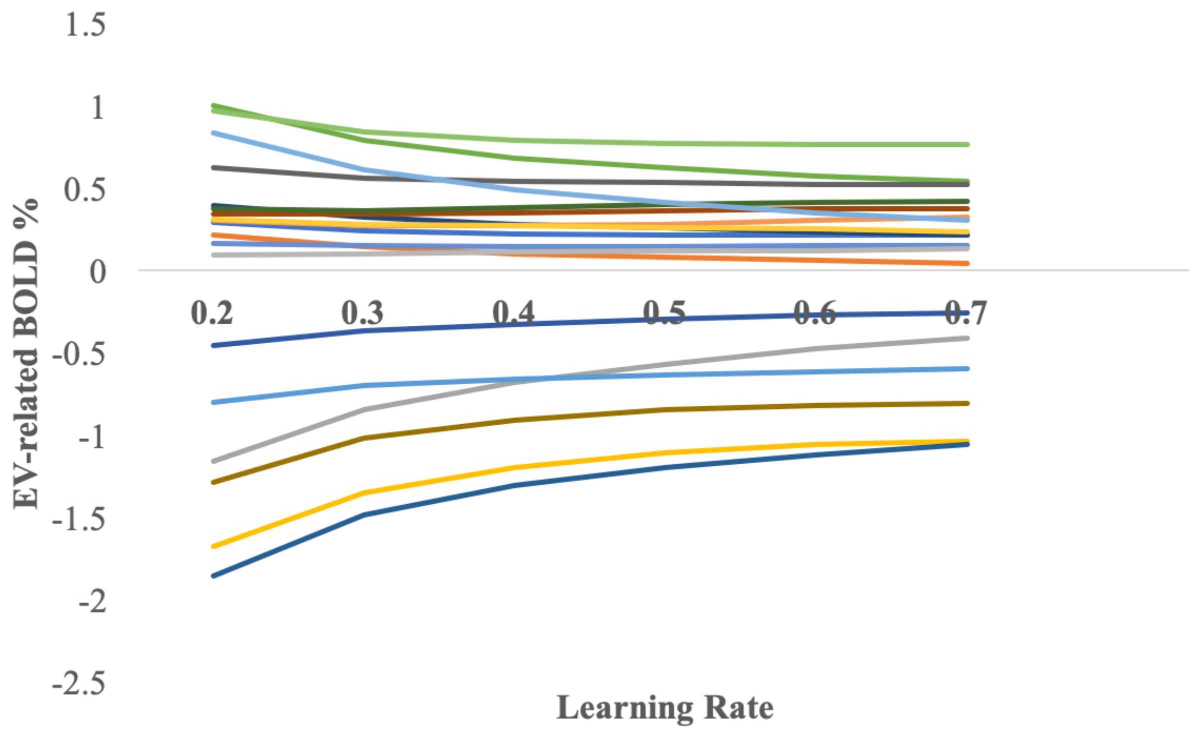

Figure 2. Regressor coefficients for expected cost regressor at the end of execution epoch estimated for SV Network ROI (Westbrook et al., 2019), as a function of learning rate. Each curve represents a single subject.

2.3.1.2. Response Time Cost Models

The Response Time Cost Model learns the average time it takes to perform an effortful task based on context. We used two separate RT models that initialized the expected cost parameter in different ways.

First, we tested a model that initialized the expected cost at the average RT during the repeat trials of the first effort run in the Learning Phase. Second, we tested a model that initialized expected cost as the first RT on the first trial of the first effort run in the Learning Phase. Then, for both initializations, average RT across all trials of each effort run is calculated and the expected cost of each effort run is updated by a PE of that effort level with a learning rate of 0.2:

$$ExpectedCost(t) = ExpectedCost(t-1) + \alpha * \left(PredictionError(t)\right),$$

PE on each trial was computed as the difference between the expected cost (i.e., expected RT of that effort level so far) and the experienced cost on trial, t (i.e., the actual average response time in that effort run):

$$PredictionError(t) = ActualCost(t) - ExpectedCost(t-1)$$

2.3.1.3. Error Rate Cost Model

We additionally tested a separate model that assumed effort costs to be average error rates during effort execution during the Learning Phase. In this model, expected cost was initialized at a 0.05 error rate and was updated based on the average error rate of a given effort run using the same Prediction Error algorithm explained above.

2.4. MRI procedure

Whole-brain imaging was performed with a Siemens 3T Prisma MRI system using a 64-channel head coil. A high-resolution T1-weighted 3D multi-echo MPRAGE image was collected from each participant for anatomical visualization. Each of the two runs of the experimental task involved between 450 and 660 functional volumes depending on the participant's response time, with a fat-saturated gradient-echo echo-planar sequence (TR = 2s, TE=28ms, flip angle = 90°, 38 interleaved axial slices, 192 mm FOV with voxel size of 3x3x3 mm). Head motion was restricted with padding, visual stimuli were rear projected and viewed with a mirror attached to the head coil.

2.4.1. fMRI Analysis

Functional images were preprocessed in SPM12 (http://www.fil.ion.ucl.ac.uk/spm). Before preprocessing, data were inspected for artifacts and variance in global signal (tsdiffana, art_global, art_movie). Functional data were corrected for differences in slice acquisition timing by resampling slices to match the first slice. Next, functional data were realigned (corrected for motion) using 4th degree B-spline interpolation and referenced to the mean functional image. Functional and structural images were normalized to Montreal Neurological Institute (MNI) stereotaxic space using affine transformation followed by a nonlinear warping based on a cosine basis set along with regularization, and then resampled into 2x2x2 mm voxels using 4th degree B-spline. Lastly, images were spatially smoothed with an 8 mm full-width at half-maximum isotropic Gaussian kernel.

A temporal high-pass filter of 128 (.0078 Hz) was applied to our functional data in order to remove noise. Changes in MR signal were modeled under assumptions of the general linear model (GLM). Two GLMs were devised: a linear effort-level GLM and an independent effort-level GLM (described below). Both GLMs included nuisance regressors for the six motion parameters (x, y, z, pitch, roll, yaw) and four run regressors for the 'Linear Effort-Level GLM' and one run regressor for the 'Independent Effort Level GLM'. The number of run regressors was different across GLMs because 'Independent Effort GLM' included the regressor for each effort level separately.

2.4.1.1. Linear Effort-Level GLM

The linear effort-level GLM tested which voxels in the brain parametrically increased or decreased linearly with effort level during learning. Two event regressors were used. First, Execution events were modeled as an impulse function at the end of the presentation of the last trial stimulus of the sequence. Second, the Anticipation event regressor modeled each anticipation event with a fixed boxcar of 3 secs. We used orthogonalized parametric modulators on these event regressors to test the linear effect of effort level. The Execution event regressor was modulated in order by (a) a Prediction Error parametric regressor, and (b) an Effort Level parametric regressor corresponding to the effort level of that task sequence (1 through 6). The Anticipation event regressor was modulated in order by (a) an expected cost regressor that scaled with the estimated expected cost of the upcoming effort execution (b) an Effort Level parametric regressor based on the presented effort level deck symbol (1 through 6). The Execution and Anticipation event regressors, along with their parametric modulators, were modeled separately for each scanning run within the GLM. Two run regressors and a linear drift over the whole experiment were included as regressors of no interest.

2.4.1.2. Independent Effort Level GLM

The independent effort level GLM sought to characterize the signal change related to each effort level independently of each other or of any particular function (e.g., linear). This GLM included twelve event regressors, one for each effort level (1 through 6) by epoch (Execution and Anticipation). Events in the Execution regressors were modeled as boxcars that onset with the presentation of the first trial stimulus of the sequence and ended with the participant's response to the final item. Events in the Anticipation regressors were modeled with a 3 sec boxcar at the onset of deck symbol. Two run regressors and a linear drift over the whole experiment were included as regressors of no interest.

For both GLMs, SPM-generated regressors were created by convolving onset boxcars and parametric functions with the canonical hemodynamic response (HRF) function. To account for error due to differences in the HRF shape from the canonical, nuisance regressors were also created that convolved the onset functions with the temporal derivative of the HRF. Beta weights for each regressor were estimated in a first-level, subject-specific fixed-effects model. For group analysis, the subject-specific beta estimates were analyzed with subject treated as a random effect. At each voxel, a one-sample t-test against a contrast value of zero gave us our estimate of statistical reliability. For whole brain analysis, we corrected for multiple comparison using cluster correction, with a cluster forming threshold of $p < .001$ and an extent threshold, k, calculated with SPM to set a family-wise error cluster level corrected threshold of $p<.05$ for each contrast and group. Note that the higher cluster forming threshold helps avoid violation of parametric assumptions such that the rate of false positive is appropriate (Eklund et al., 2016; Flandin & Friston, 2016).

2.4.1.3. ROI analysis

ROI definition is described below. For each ROI, a mean time course was extracted using the MarsBar toolbox (http://marsbar.sourceforge.net/). The GLM design was estimated against this mean time series, yielding parameter estimates (beta weights) for the entire ROI for each regressor in the design matrix.

We defined a fronto-parietal control (FPN) network ROI and Default Mode Network ROI derived from previously published cortical parcellations based on patterns of functional connectivity (Yeo et al., 2011; see Fig 3). The FPN (Network 12 in Yeo et al. 2011) included bilateral prefrontal cortex, bilateral parietal cortex, and SMA. The DMN network (Network 16 from Yeo et al. 2011) included ventromedial prefrontal cortex, orbitofrontal cortex, posterior cingulate cortex and parts of precuneus.

As a complement to these network definitions, we also included definitions of FPN and DMN based on our prior work on effort avoidance. Our previous study (Sayali and Badre, 2019, Fig 3 & Fig 4) parametrically manipulated implicitly learned effort cost/values in DST in fMRI. In that study, in order to have an unbiased analysis of the neural response function across effort levels, we conducted a Principal Component Analysis (PCA) of the whole brain to explore the shape of the neural functions of different brain areas that respond to effort execution during the Test Phase. Activity in these ROIs was associated in that study with effort avoidance decisions.

To draw connection with that prior work, we also defined fronto-parietal control and DMN ROIs based on previous study's Test Phase PCAs (Fig. 3).

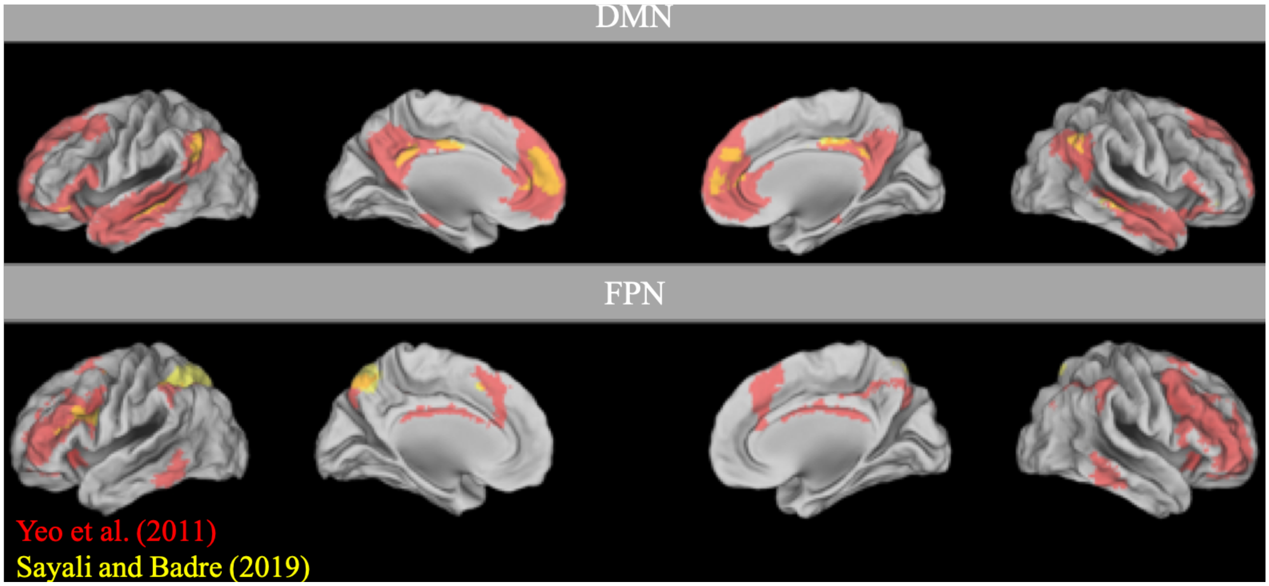

Figure 3. overlays the DMN (top) and FPN (bottom) ROIs on canonical brain as they are separately defined by Yeo et al (2011) and Sayali & Badre (2019) principle component ROIs.

We additionally defined Subjective Value (SV) Network ROIs based on the regions cited by Westbrook et al. (2019, Fig 3) in order to test the involvement of those brain regions that have been shown to significantly track subjective value during effort discounting. Accordingly, we drew a 6 mm sphere around each region that significantly correlated with the subjective value of effort.

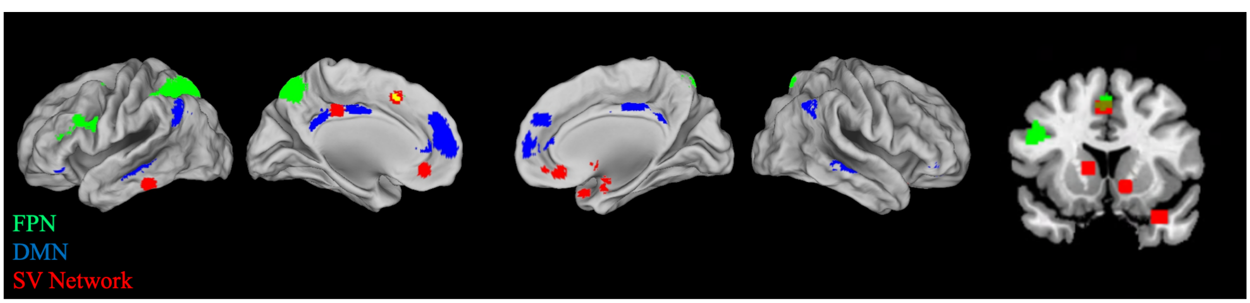

Figure 4. Sayali and Badre, 2019 ROIs of FPN and DMN overlayed on Subjective Value Network ROIs (Westbrook et al., 2019). FPN: green, DMN: blue, SV Network: red, Overlapping regions: yellow, purple.

3. Results

3.1. Demand avoidance behavior

Overall, participants performed accurately during task execution across both phases (mean error: 13% in the Learning Phase, 13% in the Test Phase), and participants missed the deadline on few trials (2.6% of Learning phase trials, *SE*=0.01, 1.2% of Test phase trials, *SE*=0.03). RTs but not error rates improved from the Learning to Test phase (RT: $F(1,17)=43.43$, $p< .001$, $\eta_p^2 =.72$; errors: $F(1,17)=0.22$, $p= .65$, $\eta_p^2 =.01$).

Task performance was impacted by the proportion of task switches across tasks. Across Learning and Test phases, there was a significant effect of effort level on both error rates and correct trial RTs (error rates: ($F(5,85)=20.07$, $p< .001$, $\eta_p^2 =.54$), RT: ($F(2.14,36.37)=143.19$, $p< .001$, $\eta_p^2 =.89$), and the effect of effort was linear for both (errors: $F(1,17)=40.82$, $p< .001$, $\eta_p^2 =.71$; RT: $F(1,17)=215.79$, $p< .001$, $\eta_p^2 =.93$). Additionally, both correct trial RTs and error rates of all effort levels reduced across the runs of the Learning phase (error rates: ($F(1,18)=6.32$, $p= .022$, $\eta_p^2 =.26$), RT: ($F(1,18)=24.17$, $p< .001$, $\eta_p^2 =.57$), indicating participants improved their performance during the Learning Phase.

Effort level affected task selections, as expected. On average, participants avoided the harder task 67% (*SE*=0.02) of the time (Fig. 5a), which differed from chance ($t(18)=6.87$, $p<0.001$). The probability of avoiding the harder task did not change across Test blocks ($F(3,54)=2.73$, $p=.53$, $\eta_p^2 =.13$), indicating that participants' effort selections were stable across time.

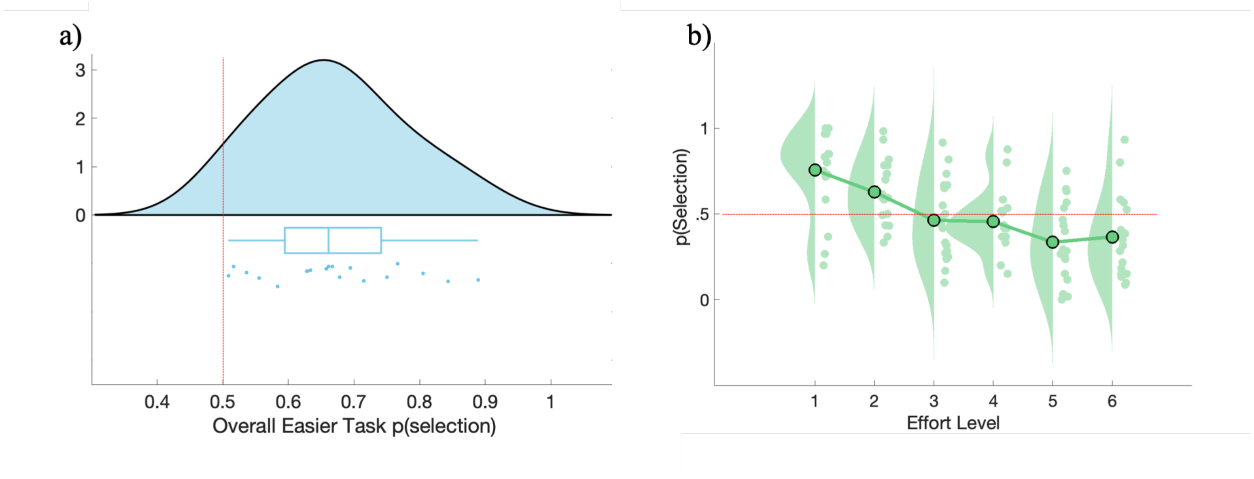

Figure 5. Demand avoidance across participants. 1A) Overall probability of selecting the easier task across all participants. The probability distribution of demand avoidance across subjects is plotted. Individual subjects are plotted as blue dots. All participants selected the easier task more than 50% of the time. 1B) Probability of selecting an effort level across participants. Probability distribution of task selections at each effort level is plotted. Individual subjects are plotted as dots to the right of each probability distribution. The probability of selecting a task significantly decreased across all participants with increasing effort level. Red dashed lines indicate chance (0.5) level.

Participants avoided higher effort levels at a greater rate than lower effort levels ($F(5,90)=8.86$, $p<.001$, $\eta_p^2=.33$), and the rate of task avoidance was linear across effort levels ($F(1,18)=47.88$, $p<.001$, $\eta_p^2=.73$; Fig. 5b), replicating Sayali and Badre (2019). The probability of selecting the easier task also significantly changed depending on the difference between effort levels for a given choice ($F(2.30,41.35)=5.14$, $p=.001$, $\eta_p^2=.22$), such that larger differences increased the probability of choosing the easier task. This effect was also linear across effort levels ($F(1,18)=13.48$, $p=.002$, $\eta_p^2=.43$). We observed a marginal increase in decision time across effort levels ($F(5,85)=1.74$, $p=.06$, $\eta_p^2=.09$).

3.2. Computational Modeling– Learning effects

We modeled effort learning using a reinforcement-learning model that acquires expectations about task performance (errors and RT) and task-switching probability for each task context (i.e., deck). In order to test our model estimates, we tested the effects of predicted response time, error rate and task-switching probability on subsequent effort selection behavior by conducting a mixed effects hierarchical regression analysis. Using a mixed-effect regression analysis, we tested which model explained selection rates better by testing the relationship between the finalized expected costs at each effort level during the Learning Phase and the selection rates in the Test Phase.

First, we tested the Error Rate Cost model that assumed effort costs to be the average error rates during effort execution in the Learning Phase. However, mixed-effects regression analysis showed that final expected error-rate costs did not explain selection rates during the subsequent Test phase, ($t(114)=-1.20$, $p=0.23$). We repeated the same analysis by calculating expected values using a learning rate that increments with 0.1 steps between 0.2 and 0.7. None of the models revealed a significant relationship between expected error rate costs and selection rates, so we did not include the error rate cost model in further analysis.

Second, we tested the two separate RT models that differentially initialized the expected cost parameter. The first RT model initialized the expected cost as the average RT from the repeat trials of the first effort run. The second initialized expected cost as the first trial RT on the first effort run of the Learning Phase. The model that assumed initial expected cost to be the first RT significantly correlated with subsequent task selection rates during the Test phase ($t(114)=-3.67$, $p< 0.001$). This model performed almost the same as the model that assumed initial expected cost to be the average Repeat RT of the first effort run ($t(114)=-3.66$, $p<0.001$). Thus, we picked one of these models (first trial RT) for model-based fMRI analysis.

Next, we tested whether task-switch probability explained selection rates. Note that the Task Switch Cost model updated expected costs by the difference between expected switch probabilities that are initiated at 0.5 and the actual switch probabilities of the effort task. Thus, given the trial numbers at learning, finalized expected costs always converged to the true switch probabilities of the effort tasks (see Table 1 for an example of average expected effort costs). Task-switch probability significantly predicted selection rates across participants, (t(114)=-6.83, $p<$ 0.001).

		Effort Level					
		1	**2**	**3**	**4**	**5**	**6**
Average Finalized Expected Cost	*RTcostModel1*	-0.61	-0.51	-0.45	-0.39	-0.35	-0.29
	RTcostModel2	-0.62	-0.52	-0.46	-0.40	-0.36	-0.30
	ERcostModel	0.10	0.13	0.15	0.14	0.15	0.16
	TScostModel	0.08	0.23	0.38	0.54	0.69	0.85

Table 1. Finalized average expected costs values for RT cost models, Task Switch Cost model (TScostModel) and Error Rate cost (ERcostModel) models. RTcostModel1 refers to the model expected cost of the first trial is initialized at first trial of the first effort run of the Learning Phase. RTcostModel2 refers to the model expected of the first trial is initialized at the average repeat trial RT of the first effort run.

Finally, we asked whether RT cost predicts selection rates over and above task-switch probability. In the presence of task-switching probability, neither RT model explained additional variance in selection rates (both *p*s>0.05). As the task-switch cost model alone explained 29% and RT cost models alone explained 10% of the variance in selection rates, model comparisons showed that including the RT cost expected costs in the same model did not significantly improve model fit (X2 (1, N = 2) = 0.4, $p > .05$). However, in order to understand the potential contribution of performance based effort learning and as RT cost does correlate with subsequent task selections, we included the RT cost model in our model-based fMRI analysis as well as the Task Switch Cost model.

3.3. The functional form of univariate brain activity over effort levels

We next sought to replicate and extend our prior observations from the Test phase (Sayali and Badre, 2019) to the Learning phase in a new sample of participants. We investigated the pattern of fMRI activity across effort levels during the Learning Phase using ROIs defined from the Test phase of Sayali & Badre (2019, see Methods), as well as ROIs in the *a priori* SV network (Westbrook et al., 2019). Note that we also separately tested independent ROIs in FPN and DMN defined from a large scale study of resting state functional connectivity (Yeo et al., 2011). However, these ROIs showed similar trends to those reported below, so we do not detail them further.

As expected, activation in FPN showed a linear increase across effort levels during task execution (Fig. 6 left panel). There was a significant effect of effort on FPN activation ($F(5,90)=11.12$, $p< .001$, $\eta_p^2 =.38$) such that activity linearly increased with greater effort requirements of the task ($F(1,18)=26.02$, $p< .001$, $\eta_p^2 =.59$). Additionally, FPN activation decreased over time (in terms of blocks; $F(1,18)=11.29$, $p= .003$, $\eta_p^2 =.39$). There was no effect of anticipated effort on the FPN activation during the Anticipation epoch (Fig. 6 right panel: $F(3.51,63.21)=1.22$, $p= .31$, $\eta_p^2 =.06$).

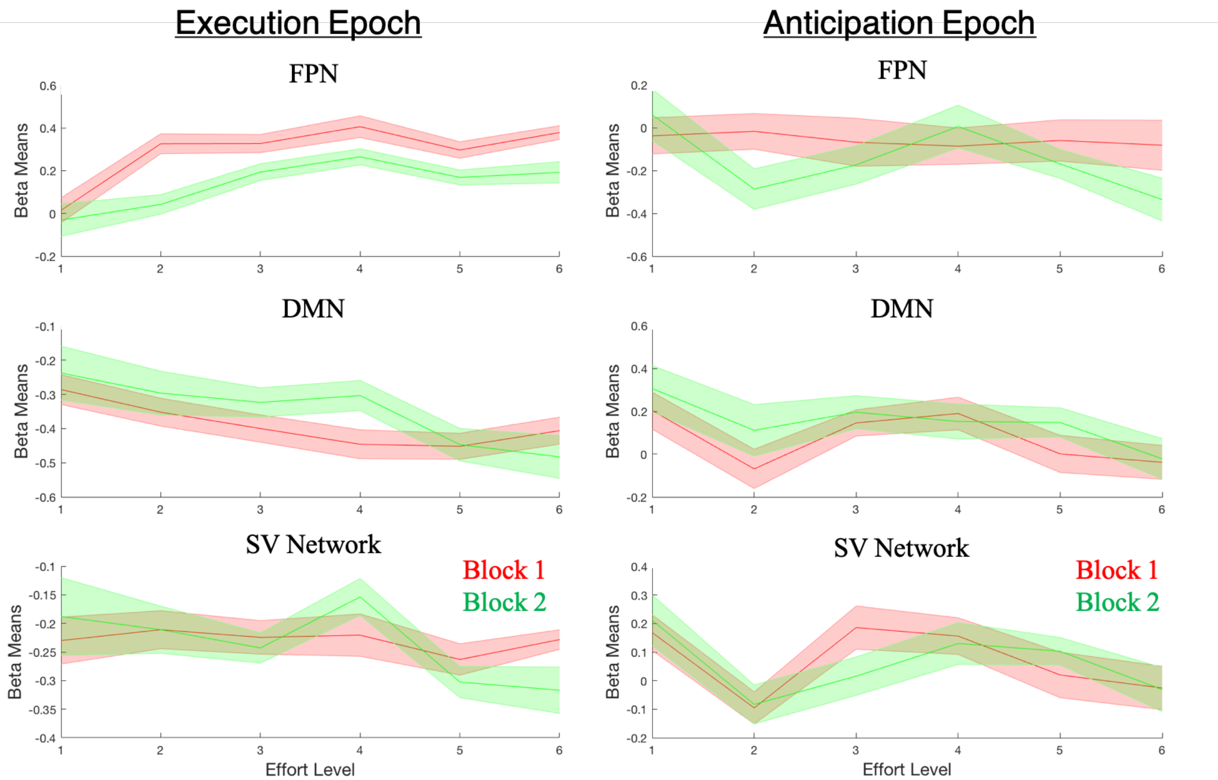

Figure 6. Regressor coefficients at the time of Execution and Anticipation Epochs for FPN component, DMN component and the Subjective Value Network.

Consistent with our prior observation from the Test Phase (Sayali and Badre, 2019), during the Learning Phase, the DMN showed a negative linear trend in activation across effort levels (Fig. 6 left panel). There was a significant effect of effort on DMN activity ($F(5,90)=3.31$, $p= .01$, $\eta_p^2 =.16$) and a linear decrease in activity with increasing effort requirements of the task ($F(1,18)=11.59$, $p = .003$, $\eta_p^2 =.39$). There was no effect of time (in terms of runs) on DMN activity ($F(1,18)=1.19$, $p= .29$, $\eta_p^2 =.06$). There was a marginal effect of effort on the DMN during the Anticipation epoch (Fig. 6 right panel, $F(5,90)=2.25$, $p= .06$, $\eta_p^2 =.11$; Yeo et al definition of DMN: $F(5,90)=2.34$, $p= .05$, ηp2 =.12), suggesting that DMN marginally reduced activity in the anticipation of increasing effortful task exertion.

We did not find a significant positive or negative trend in SV network activity across levels of effort execution ($F(5,90)=1.23$, $p= .30$, $\eta_p^2 =.06$) or across blocks during task execution

(F(1,18)=0.03, p= .26, η_p^2 =.002). During the anticipation interval prior to each imposed task, SV network activity was positively related to effort level (F(5,90)=2.44, p= .04, η_p^2 =.12, Fig. 6 right panel). Within-subject contrast results showed that the SV network showed a cubic trend across the anticipation of effort levels (F(1,18)=5.70, p = .028, η_p^2 =.24). There was no effect of run during the anticipation epoch on activity in SV Network (F(1,18)=0.02, p = .88, η_p^2 =.001).

In sum, the FPN and DMN exhibited similar modulation of brain activity during the Learning of effort tasks as previously observed during the Test phase. Further, FPN recruitment decreases with increased experience in the Learning Phase, which is expected based on the prior literature on task learning and experience (Ruge and Wolfensteller, 2010; Bhandari and Badre, 2020). The *a priori* SV Network showed a reliable omnibus effect of increase activity with greater anticipated effort prior to task performance.

3.4. Model-based fMRI

3.4.1. Task Switch Cost model

Our modeling analysis of learning behavior indicated that the expected likelihood of task switching acquired during the Learning phase was the best predictor of subsequent task selections during the Test phase. In the model, this expected likelihood is formed by incrementally adjusting an expected task switch probability based on task switch prediction errors. We next sought to test how the *a priori* hypothesized networks, FPN, DMN, and SV network, track prediction errors related to task switch probability. Thus, we tested the contribution to activity in these brain regions from the parametric functions of Prediction Error and expected cost parametric regressors in these ROIs, in addition to the Execution Level and the onset of Execution and Anticipation events. These effects were modeled within the Linear Effort Level GLM, described in the Methods.

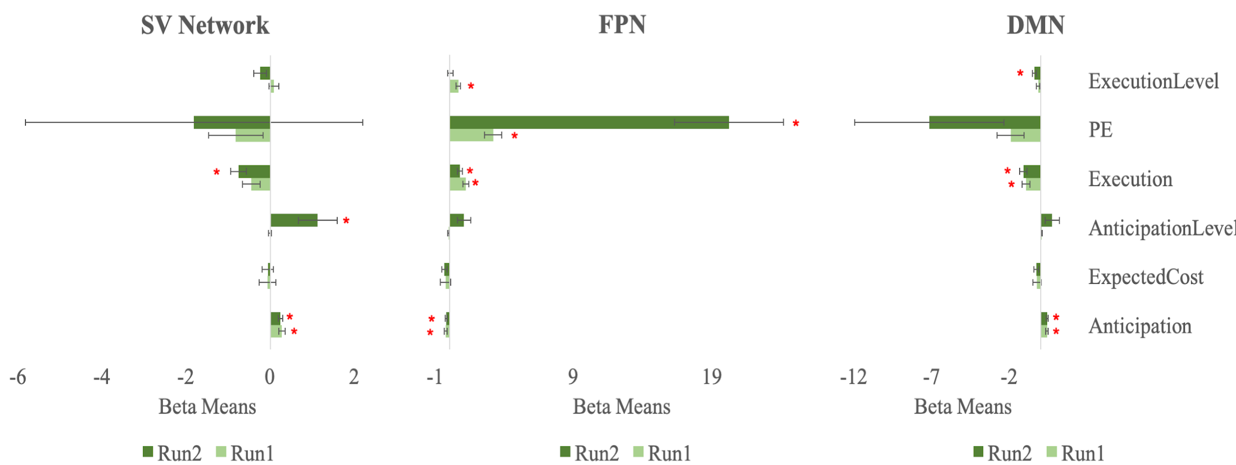

Figure 7. Regressor coefficients for Linear Effort Level GLM for DMN, FPN and Value Network ROIs. FPN showed positive, DMN and Value Network showed negative activation during Execution onset. FPN positively scaled with PE.

We focus here on the key model-based results related to prediction error and expected cost (see Fig. 7 for summary of complete results). The FPN correlated positively with prediction error ($p < 0.001$), such that FPN activation was greater when tasks switched more frequently during a run than expected based on prior experience (Fig. 7). This effect was strongest during the second relative to the first run of the task. Neither the DMN nor the SV network correlated with prediction error (both *p*s > 0.05), despite strong trends, particularly for the former. Neither FPN, DMN or SV networks correlated with expected costs.

In order to test specific nodes within the hypothesized SV network, we analyzed additional ROIs in left, right VS, two rVMPFC ROIs, lVMPFC and dACC that were included in Westbrook et al., (2019). While both left and right vMPFC showed significant deactivations during effort execution and anticipation, only dACC significantly tracked PE regressor during the first run and parametrically increased with effort anticipation in the second run.

Finally, to complement the ROI analyses, we performed a whole-brain voxel-wise analysis (Fig. 8) of the key model-based regressors. As shown in Figure 8, the expected cost regressor negatively correlated with one cluster at left middle frontal gyrus (x=-26, y=4, z=46). Two clusters positively correlated with prediction error (Figure 8). One cluster was centered at superior parietal lobule (x=-12, y=-68, z=56) and the other was centered at left middle frontal gyrus (x=-40, y=24, z=24).

3.4.2. RT cost model

As the RT cost model also predicted selection rates, we repeated our fMRI analysis using expected values and prediction errors derived from this model. Analysis of our a priori ROIs yielded results weaker but qualitatively similar results to that of the Task Switch Cost model. The FPN ($p < 0.001$) again correlated with prediction error in the 1st run, but the DMN and the SV network did not (both ps > 0.05). Neither FPN, DMN nor SV networks correlated with expected costs.

Also consistent with Task Switch Cost model, whole-brain voxel-wise analysis (Fig. 10) showed that expected cost regressor negatively correlated with one cluster at left precentral gyrus (x=-32, y=-10, z=36), additionally one at right lingual gyrus (x=10,y=-46,z=-4) and one at left insula (x=-46, y=12, z=-4).

Importantly, the whole brain analysis based on RT Cost PE did differ from the Task Switch Cost PE. Seven clusters survived the family-wise error-correction (Figure 8). Peaks positively correlated with PE were identified in precuneus (x=6, y=-92, z=22), dACC (x=-10, y=20, z=26), right supramarginal gyrus (x=60, y=-42, z=28), right superior frontal gyrus (x=24, y=46, z=10), right anterior insula (x=50, y=20, z=-6), left middle temporal gyrus (x=-54, y=-56, z=-6), and posterior cingulate gyrus (x=-4, y=-32, z=26). Notably, these clusters fall within the previously defined cingulo-operculum networks (CON) (Yeo et al., 2011), but not FPN.

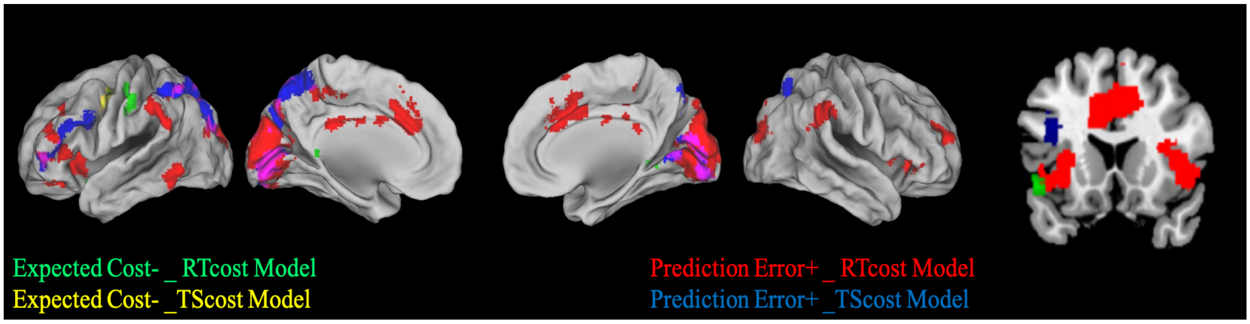

Figure 8. Whole Brain Results for Task Switch Probability Cost and Response Time Cost Models. All results thresholded at FWE p < .05.

Given this apparent difference in networks correlating with the two PEs, we ran a posthoc analysis testing whether Task Switch Cost PE and RT cost PE significantly differ in activating FPN (as defined in Yeo et al. 2011) vs. CON during PE. We used a priori ROI definitions of each network from Yeo et al. (2011, Network 8) and tested the effect of experimental run, model type (Task Switch Cost, RT cost) and ROI (FPN and CON) on BOLD response during PE. This analysis revealed a significant 3-way-interaction of model*ROI*run ($F(1,18)=12.28$, $p= .003$, $\eta_p^2 =.41$). Pairwise comparisons confirmed greater RT cost PE activation in CON the 1st run compared to Task Switch Cost PE. In contrast, there was greater Task Switch Cost PE activation in the FPN in the 2nd run compared to RT cost PE (Fig. 9). Thus, FPN and CON showed a differential relationship to PE derived from RT versus task switch frequency and in different phases of learning.

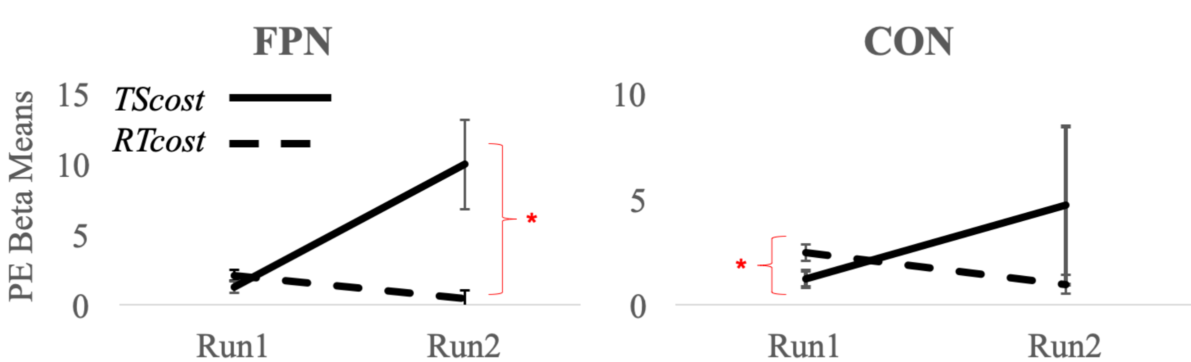

Figure 9. Three-way interaction between model, Yeo et al. 2011 defined ROIs and experimental run for PE regressor beta means obtained from Task Switch Cost and RT Cost Models. TScost: Task Switch Cost Model; RTcost: Response Time Cost Model.

4. **Discussion**

A central problem in cognitive neuroscience has concerned why people choose to perform certain tasks over others when prospective rewards are otherwise comparable. The widespread view is that cognitive tasks are subjectively costly and discount outcomes accordingly. Thus, an open question has been what aspects of a task make it costly and how does the brain track these demands. The present study adds a new perspective on this question by considering which demands are tracked during experience with a task, in the absence of any decision, but that later correlate with avoidance.

We find evidence that at least two factors are tracked: the frequency of task switching and the expected RT. These factors are correlated, and so it was hard to distinguish them behaviorally. Indeed, task switching frequency held a stronger relationship to choice and was able to fully explain the variance that was accounted for by expected RT.

Unexpectedly, however, using model-based fMRI, we found that trial-to-trial PE from each source held stronger relationships with distinct brain networks and at different stages of learning. Specifically, when RT was greater than expected, PE-related brain activity in control networks was greatest early in learning and this association declined by the second run. Further, this effect was differentially greater in the CON relative to FPN. In contrast, the correlation of brain activity with PE derived from the expected frequency of task switching was primarily in the second run of learning, and was differentially correlated with the FPN relative to CON.

We did not hypothesize these differences a priori and so future work must replicate them. Nonetheless, if verified, these results offer intriguing insights into the nature of task demands. At the surface, they suggest that multiple factors likely contribute separately to our assessment of task demands. Thus, rather than a single demand that drives all evaluations of cognitive effort, multiple factors are weighed, likely dynamically, depending on the situation. For example, there may be contexts where time-on-task is more or less salient than the frequency of task switching. And further, there may be more than one brain system that contributes to these decisions.

The early association of control network activity, and particularly the CON, with PE derived from RT is intriguing in light of the broader literature on task learning and automaticity. It is well established that when new tasks are performed, the first trial is associated with a higher RT and error rate that declines in log form toward an asymptote (Bhandari and Badre, 2018). Studies testing these early trial effects have repeatedly associated this period of rapid task adjustment with activation in control networks (Cole et al., 2010; Ruge and Wolfensteller, 2010; Mohr et al., 2016; Hampshire et al., 2019). A recent study that held the stimulus-response rules constant, but required adjustments only in control settings – analogous to the task switching demands that distinguish context in the present study – specifically found these rapid changes to be tied to adjustments in control settings (Bhandari and Badre, 2018). Further, these changes were specifically correlated with activation in the CON (Bhandari and Badre, 2020) along with striatum. This rapid change contrasted with slower changes in the FPN.

Following from this discussion, one hypothesis is that early learning is characterized by adjustments in control settings that adapt to the specific task context. Optimization in this stage is closely tied to changes in RT, and thus, PEs related to RT are informative regarding task demands. The CON may be important in tracking these errors of prediction.

This interpretation is broadly consistent with current views that ascribe a monitoring role to the CON network and dACC in particular (Shenhav et al., 2013; Botvinick, 2007). As one example, the Predicted Response-Outcome Model (PRO; Alexander and Brown, 2011) argues that the medial portion of PFC, including the dACC, implements two functions: 1) learning to predict response outcome requirements, 2) signaling prediction errors. Accordingly, dACC activity reflects changes in the predicted value of the effortful task.

The later sensitivity of the FPN to task switching frequency might similarly be related to the nature of task learning. The association of FPN with task switching has been widely observed (Kim et al., 2012). We also previously observed increasing activation in FPN with task switching frequency (Sayali and Badre, 2019), an effect we replicate here. However, the present results further show that the FPN is also more active when task switching is greater than expected. The later emergence of this association is notable given the overall activation decrease in the FPN with greater task experience.

When people learn control settings, in addition to the fast adjustments in RT noted above, there is also a slower process of learning that favors stable performance over variance in task demands (Bhandari and Badre, 2020). It is possible that the engagement of control, and the FPN, in this stable phase is attuned to unexpected and irregular events, particularly when they require control. Thus, deviations in the use of control itself can drive learning about task demands in this later phase. Future work should test these hypotheses directly and tie these brain dynamics directly to effort avoidance.

Based on previous studies of effort-based decisions and learning (Westbrook et al., 2019; Nagase et al., 2018), we hypothesized that the SV network would also track PE during the implicit learning phase. Surprisingly, we did not locate evidence of an association of either PE or anticipated cost with activation in the SV network. This seems at odds with prior observations, though at least one prior study also failed to locate an association of the SV network with the degree of physical effort required by a task (Prévost et al., 2010). However, given some trends, it is possible that our study simply lacked sensitivity in this network to locate an effect over noise.

Another possibility is that the SV network is affected by the requirement to make a decision. A clue in this regard comes from the anticipation phase, when we would predict less activation in SV prior to a difficult task. Instead, we observed greater activation in the SV network during anticipation of a more effortful task. This reversal might reflect that participants are anticipating an imposed task rather than evaluating which task to perform. Indeed, Schouppe et al. (2014) reported that the direction of activation in VS during the anticipation of effortful tasks changed as a function of whether the task was imposed or voluntarily selected.

Finally, we highlight two limitations of the study. First, as participants did not make task selections during learning, it was not possible to estimate individual learning rates through model fitting. This shortcoming indicates that current study cannot estimate the degree to which task switch cost or RT cost during effort learning is directly tied to effort selections. This will be an important gap to fill in future work.

Second, the present study prospectively sought to enroll and test demand avoiding participants. The rationale for this choice was that our previous study found significant individual differences in demand avoidance behavior on the basis of those who are demand seeking versus demand avoiding (Sayali and Badre, 2019), with demand seeking participants showing diminished or even reversed task choice behavior. We further found that participants in fMRI experiments tend to be disproportionately demand seeking. Thus, in order to reduce variability in our sample, we sought only demand avoiding participants.

Though this approach successfully decreased the individual variance in our study, this necessarily affects the generalizability of these results. Thus, our conclusions are most closely generalizable to demand avoiding individuals. Nonetheless, we argue that these effects are likely also applicable to less avoidant individuals when they choose to avoid effort. Indeed, in our prior work, we found that brain systems highlighted here, showed similar activity patterns across both demand avoiders and demand seekers as a function of task demands. However, it was the interpretation of these signals, in terms of their relationship to choice, that distinguished the groups.

What does experience with a task teach us about its demands? Our results indicate that at least two factors are critical. First, response time is tracked as an index of performance, particularly during our first experiences with a task, when we are rapidly adjusting our control settings and automatizing performance. We form expectations about those control settings over time, like task switching frequency, such that later deviations from these expectations become a prime source of learning about task demands. Thus, separate factors are tracked in the brain by control networks such as the FPN and CON, and offer differential sources of information about task demands and the prospects for cognitive effort.

References

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature neuroscience*, *14*(10), 1338.

Bhandari, A., and Badre, D. (2018). Learning and transfer of working memory gating policies. *Cognition, 172,* 89-100.

Bhandari, A., and Badre, D. (2020). Fronto-parietal, cingulo-opercular and striatal contributions to learning and implementing control policies. bioRxiv. **doi:** https://doi.org/10.1101/2020.05.10.086587

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 356-366.

Chong, T. T. J., Apps, M., Giehl, K., Sillence, A., Grima, L. L., & Husain, M. (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS biology*, *15*(2).

Cole MW, Bagic A, Kass R, Schneider WJJoN (2010) Prefrontal dynamics underlying rapid instructed task learning reverse with practice. 30:14245-14254.

Daw, N. D., & Tobler, P. N. (2014). Value learning through reinforcement: the basics of dopamine and reinforcement learning. In Neuroeconomics (pp. 283-298). Academic Press.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the national academy of sciences*, *113*(28), 7900-7905.

Flandin, G., & Friston, K. J. (2019). Analysis of family-wise error rates in statistical parametric mapping using random field theory. Human brain mapping, 40(7), 2052-2054.

Gold, J. M., Kool, W., Botvinick, M. M., Hubzin, L., August, S., & Waltz, J. A. (2015). Cognitive effort avoidance and detection in people with schizophrenia. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(1), 145-154.

Hampshire A, Daws RE, Neves ID, Soreq E, Sandrone S, Violante IR (2019) Probing cortical and sub-cortical contributions to instruction-based learning: Regional specialisation and global network dynamics. NeuroImage 192:88-100.

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. *Human brain mapping*, *33*(1), 130-142.

Massar, S. A., Libedinsky, C., Weiyan, C., Huettel, S. A., & Chee, M. W. (2015). Separate and overlapping brain areas encode subjective value during delay and effort discounting. *Neuroimage*, *120*, 104-113.

Mohr H, Wolfensteller U, Betzel RF, Mišić B, Sporns O, Richiardi J, Ruge HJNc (2016) Integration and segregation of large-scale brain networks during short-term task automatization. 7:13217.

Nagase, A. M., Onoda, K., Foo, J. C., Haji, T., Akaishi, R., Yamaguchi, S., ... & Morita, K. (2018). Neural mechanisms for adaptive learned avoidance of mental effort. *Journal of Neuroscience*, *38*(10), 2631-2651.

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M. L., & Dreher, J. C. (2010). Separate valuation subsystems for delay and effort decision costs. *Journal of Neuroscience*, *30*(42), 14080-14090.

Ruge H, Wolfensteller UJCC (2010) Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. 20:1656-1667.

Sayalı, C., & Badre, D. (2019). Neural systems of cognitive demand avoidance. Neuropsychologia, 123, 41-54

Schouppe, N., Demanet, J., Boehler, C. N., Ridderinkhof, K. R., & Notebaert, W. (2014). The role of the striatum in effort-based decision-making in the absence of reward. *Journal of Neuroscience*, *34*(6), 2148-2154.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-240.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Westbrook, A., Lamichhane, B., & Braver, T. (2019). The subjective value of cognitive effort is encoded by a domain-general valuation network. *Journal of Neuroscience*, *39*(20), 3934-3947.

Westbrook, A., van den Bosch, R., Määttä, J. I., Hofmans, L., Papadopetraki, D., Cools, R., & Frank, M. J. (2020). Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science*, *367*(6484), 1362-1366

Wilson, R. C., & Niv, Y. (2015). Is model fitting necessary for model-based fMRI?. *PLoS computational biology*, *11*(6).