

Neural Mechanisms for Adaptive Learned Avoidance of Mental Effort

Asako Mitsuto Nagase,^{1,3} Keiichi Onoda,² Jerome Clifford Foo,^{3,4,5} Tomoki Haji,^{6,7,8} Rei Akaishi,^{3,9} Shuhei Yamaguchi,² Katsuyuki Sakai,^{3,7} and Kenji Morita^{3,4}

¹Graduate School of Medical Research, ²Department of Neurology, Faculty of Medicine, Shimane University, Izumo 693-8501, Japan, ³Department of Cognitive Neuroscience, Graduate School of Medicine, ⁴Physical and Health Education, Graduate School of Education, The University of Tokyo, Tokyo 113-0033, Japan, ⁵Department of Genetic Epidemiology in Psychiatry, Central Institute for Mental Health, Medical Faculty Mannheim/University of Heidelberg, 68159 Mannheim, Germany, ⁶Brain Activity Imaging Center, Advanced Telecommunications Research Institute Promotions, Kyoto 619-0288, Japan, ⁷Brain Science Institute, Tamagawa University, Tokyo 194-8610, Japan, ⁸Center for Information and Neural Networks, National Institute of Information and Communications Technology, Osaka 565-0871, Japan, and ⁹Department of Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, New York 14627

Humans tend to avoid mental effort. Previous studies have demonstrated this tendency using various demand-selection tasks; participants generally avoid options associated with higher cognitive demand. However, it remains unclear whether humans avoid mental effort adaptively in uncertain and nonstationary environments. If so, it also remains unclear what neural mechanisms underlie such learned avoidance and whether they remain the same regardless of cognitive-demand types. We addressed these issues by developing novel demand-selection tasks where associations between choice options and cognitive-demand levels change over time, with two variations using mental arithmetic and spatial reasoning problems (males/females: 29:4 and 18:2). Most participants showed avoidance, and their choices depended on the demand experienced on multiple preceding trials. We assumed that participants updated the expected cost of mental effort through experience, and fitted their choices by reinforcement learning models, comparing several possibilities. Model-based fMRI analyses revealed that activity in the dorsomedial and lateral frontal cortices was positively correlated with the trial-by-trial expected cost for the chosen option commonly across the different types of cognitive demand. Analyses also revealed a trend of negative correlation in the ventromedial prefrontal cortex. We further identified correlates of cost-prediction error at time of problem presentation or answering the problem, the latter of which partially overlapped with or were proximal to the correlates of expected cost at time of choice cue in the dorsomedial frontal cortex. These results suggest that humans adaptively learn to avoid mental effort, having neural mechanisms to represent expected cost and cost-prediction error, and the same mechanisms operate for various types of cognitive demand.

Key words: avoidance learning; cognitive demand; decision making; mental effort; model-based fMRI; reinforcement learning

Significance Statement

In daily life, humans encounter various cognitive demands and tend to avoid high-demand options. However, it remains unclear whether humans avoid mental effort adaptively under dynamically changing environments. If so, it also remains unclear what the underlying neural mechanisms are and whether they operate regardless of cognitive-demand types. To address these issues, we developed novel tasks where participants could learn to avoid high-demand options under uncertain and nonstationary environments. Through model-based fMRI analyses, we found regions whose activity was correlated with the expected mental effort cost, or cost-prediction error, regardless of demand type. These regions overlap, or are adjacent with each other, in the dorsomedial frontal cortex. This finding helps clarify the mechanisms for cognitive-demand avoidance, and provides empirical building blocks for the emerging computational theory of mental effort.

Introduction

Humans tend to avoid mental effort in various situations associated with many types of cognitive demand. When making com-

plex decisions, humans tend to rely on heuristics instead of effortful reasoning (Tversky and Kahneman, 1974). Humans also discount reward values when mental effort is required (Botvinick

Received July 13, 2017; revised Jan. 26, 2018; accepted Jan. 31, 2018.

Author contributions: A.M.N. wrote the first draft of the paper; A.M.N. and K.M. edited the paper; A.M.N., K.S., and K.M. designed research; A.M.N., J.C.F., and T.H. performed research; A.M.N. and K.M. contributed unpublished

reagents/analytic tools; A.M.N., K.O., and K.M. analyzed data; A.M.N., K.O., J.C.F., R.A., S.Y., and K.M. wrote the paper.

This work was supported by the Impulsing Paradigm Change through Disruptive Technologies Program (Cabinet Office, Government of Japan) to S.Y., Funding Program for Next Generation World-Leading Researchers (LS030,

et al., 2009; Massar et al., 2015; Chong et al., 2017), and expend physical effort to reduce mental effort (Risko et al., 2014). Moreover, exertion of mental effort causes fatigue effects on subsequent choice behavior (Blain et al., 2016). To clarify the precise nature of mental-effort avoidance in the absence of other factors affecting decisions, such as reward or physical effort, previous researchers developed the demand-selection task paradigm (Botvinick, 2007). In this paradigm, participants freely choose one of two cues associated with high and low cognitive demands. By using several variations of the task, researchers have demonstrated the generality of cognitive-demand avoidance, to the extent that potential confounders, such as the rate of errors or the time on task, could not fully explain (Kool et al., 2010).

In daily life, the cognitive demands of choice options encountered are likely to change over time. Work using this demand-selection task (Kool et al., 2010) has examined the condition where participants needed to learn the association between novel cues and stable demand levels in every task block, finding that most participants consistently avoided higher-demand options. However, it has yet to be experimentally demonstrated whether humans adaptively learn to avoid higher cognitive demand through experience in situations where demand levels are not stationary, i.e., when the association between cues and demand levels fluctuates and changes over time.

Moreover, if humans exhibit this kind of experience-based adaptive learned avoidance of mental effort, exploring its neural basis is of particular interest. A number of studies have identified neural correlates of the level of imposed cognitive demand (Botvinick et al., 2001; Duncan, 2010; Mansouri et al., 2017; Shenhav et al., 2017) or anticipated cognitive demand (Sohn et al., 2007; Krebs et al., 2012; Vassena et al., 2014), the avoidance rating of experienced cognitive demand (McGuire and Botvinick, 2010), or the mental effort-discounting of reward values (Botvinick et al., 2009; Massar et al., 2015; Chong et al., 2017). However, the results of these studies are not yet sufficient to understand the neural mechanisms for adaptive learned avoidance of mental effort. Significantly, the previous imaging studies did not examine brain activity during learned avoidance based on trial-by-trial experience.

Furthermore, to clarify general neural mechanisms for mental-effort avoidance (i.e., those which operate regardless of demand type), it is necessary to test more than one type of cognitive demand. The previous imaging studies on anticipation or avoidance of cognitive demand tested only a single type of cognitive demand in each study (Sohn et al., 2007; Botvinick et al., 2009; Krebs et al., 2012; Vassena et al., 2014; Massar et al., 2015; Chong et al., 2017). Therefore, it remains unclear whether the same neural mechanisms underlie avoidance of various types of cognitive demand.

To address these questions, we formed two hypotheses. First, we hypothesized that humans adaptively learn through experience to avoid an option that presently requires higher cognitive demand in the situation where the demand level of options changes over time. This learning process was assumed to be approximated by reinforcement-learning models in which the ex-

pected cost of mental effort is updated according to prediction error (PE). Second, we hypothesized that the expected cost estimated from the model is represented in the same brain regions regardless of the types of cognitive demands. To test these hypotheses, we developed two tasks requiring different cognitive-demand types where associations between choice options and cognitive-demand levels change over time. We fitted participants' choices using various models, conducted model comparisons, and explored brain regions representing the expected mental effort cost and cost-prediction error (CPE) through model-based fMRI analyses.

Materials and Methods

Participants

There were 33 participants (four females; mean age, 25.5 ± 5.4 years) in Experiment 1 and 20 participants (two females; mean age, 24.7 ± 6.2 years) in Experiment 2. Six participants took part in both experiments. We paid all participants equally with book store gift cards (worth ¥6000) for their participation. No participants were taking any medicine or had prior history of neuropsychiatric disorders. All participants were right-handed and native Japanese speakers. Informed written consent was obtained from all participants before the experiment. The present study was approved by the ethics committee of the Graduate School of Medicine, the University of Tokyo.

Behavioral tasks

We include a method summary for behavioral tasks and analyses in the Results, and here we describe all the details. We conducted two experiments. These had the same structure but used different types of problems requiring different kinds of cognitive activity (Fig. 1). Specifically, we used mental division (arithmetic) problems in Experiment 1 and mental cube-folding (spatial reasoning) problems in Experiment 2. For both experiments, we prepared problems with two levels of cognitive demand, i.e., high-demand and low-demand problems.

In Experiment 1, we required participants to divide a five-digit number by 7 and report whether the remainder was small (≤ 3) or large (≥ 4) via a button press with no time limitations. In low-demand problems, the dividend (e.g., 35426) consisted of two consecutive two-digit numbers that were multiples of 7 followed by a single one-digit number from 1 to 6. In contrast, the dividend of high-demand problems did not contain any numbers that were multiples of 7 in mental calculation processes (e.g., 48106). More specifically, the set of low-demand problems consisted of seven patterns for the digits representing ten-thousands and thousands (14, 21, 28, 35, 42, 49, 56, 63, 70), seven patterns for the digits representing hundreds and tens (7, 14, 21, 28, 35, 42, 49, 56, 63), and six patterns for the digits representing ones, and a combination of them was pseudorandomly selected in each trial. Meanwhile, dividends in higher-demand problems satisfying the above-mentioned rule were selected on each trial by using a pseudorandom function. The five-digit numbers and the answer choices were presented as gray characters on a black background.

In Experiment 2, we required participants to judge whether a concurrently presented 3D cube with three visible colored faces matched an unfolded cube with all six faces and colors (purple, red, yellow, green, sky blue, and deep blue), and report the answer via a button press with no time limitations. The difference between low-demand and high-demand problems was whether the three faces shown on the 3D cube were adjacent on the unfolded cube. There were in total 78 patterns (match, 39; nonmatch, 39) for low-demand problems and 120 patterns (match, 60; nonmatch, 60) for high-demand problems, and in each trial, one problem was pseudorandomly selected from these patterns with duplication permitted. The set of high-demand problems used was selected from all the possibilities so that the correct ratio was expected to be ≥ 0.8 based on the results of pilot experiments. The 3D and unfolded cubes and the answer choices (match, circle; nonmatch, cross) were presented on a white background. The locations of the answer choices (circle and cross), in either the left-bottom or right-bottom, were fixed for each participant and counter-balanced across participants.

In both Experiments 1 and 2, at the start of each trial, two arrow cues appeared, one on the left and the other on the right of the screen (via the

Cabinet Office, Government of Japan) to K.S., and Grant-in-Aid for Scientific Research (No. 15H05876) of The Ministry of Education, Culture, Sports, Science and Technology in Japan to K.M. We thank Kazumasa Umeda, Motomu Katsurakawa, and Takuya Ibaraki for data collection assistance.

The authors declare no competing financial interests.

Correspondence should be addressed to either of the following: Asako Mitsuto Nagase, Graduate School of Medical Research, Shimane University, Enya-cho 89-1, Izumo 693-8501, Japan, E-mail: asako3210+academic@gmail.com; or Kenji Morita, Physical and Health Education, Graduate School of Education, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan, E-mail: morita@p.u-tokyo.ac.jp.

DOI:10.1523/JNEUROSCI.1995-17.2018

Copyright © 2018 the authors 0270-6474/18/382632-21\$15.00/0

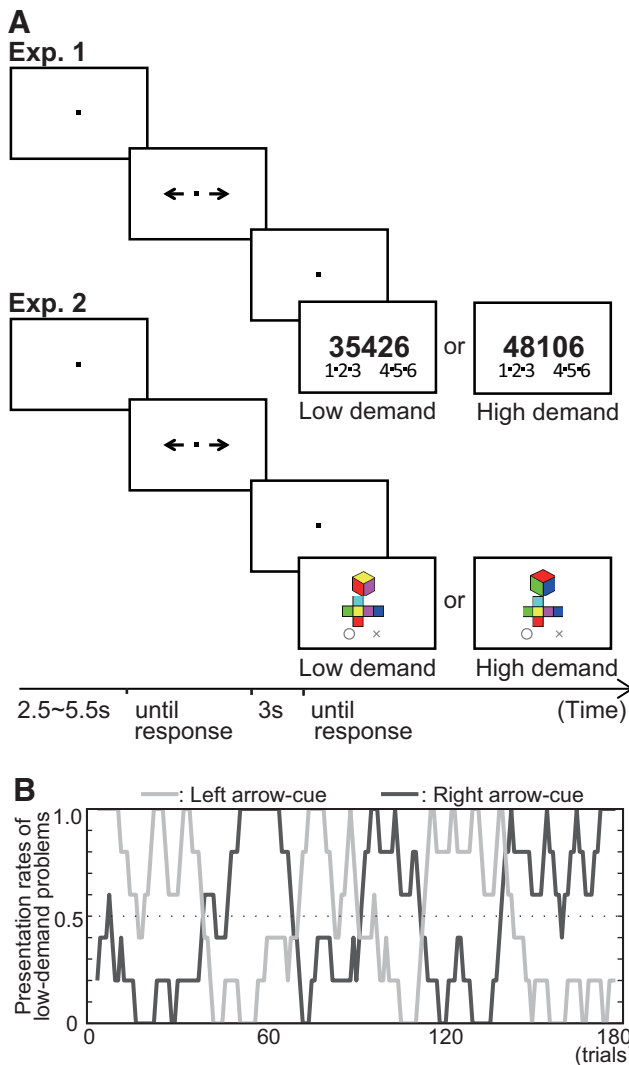


Figure 1. Behavioral paradigm. **A**, Participants chose an arrow cue at the start of each trial. After the choice, a problem was presented. In Experiment 1 (Exp. 1), the problem was mental arithmetic: to divide a five-digit number by 7 and report whether the remainder was small or large. In Experiment 2 (Exp. 2), the problem was spatial reasoning: to judge whether a 3D cube matched an unfolded cube. In both experiments, there were high-demand problems and low-demand problems, whose presentation rates were associated with the arrow cues and varied over time. **B**, An example of the presentation rates of low-demand problems (moving average of latest 5 trials) associated with the left arrow-cue (light gray) and the right arrow-cue (dark gray).

overhead mirror). Participants were asked to choose one of them, without time restriction, to specify whether to solve the “left-arrow” problem or the “right-arrow” problem on that trial (without seeing the problem itself). Three seconds after participants chose one of the arrow cues, the chosen problem was presented in the center of the screen. Participants were asked to answer the problem (no time limitation), and when a choice was made by a button press, an intertrial interval started, followed by the next trial. The intertrial intervals were jittered: 2.5, 4, or 5.5 s. We asked participants to answer problems as fast and accurately as possible, giving weight to the accuracy. We further instructed the participants that if they felt unsure about their answer, they were asked to redo the mental calculation or cube-folding until they were sure. We did not provide correct/incorrect feedback to participants so as to minimize the possible effects of mistakes on the way participants made decisions.

The probability that a high-demand or low-demand problem appeared depended on whether participants chose the left-arrow cue or the right-arrow cue at the start of each trial, and the probabilistic associations

between each of the cues and high-demand and low-demand problems changed over trials (Fig. 1B). Specifically, when programming the task, we divided the entire 180 trials into six blocks. In the first, third, and fifth blocks, the left and right arrow-cues were associated with low-demand problems in 80 and 20% of trials within each block, respectively. In the remaining (i.e., second, fourth, and sixth) blocks, the cue–demand associations were reversed, i.e., the left and right arrow-cues were associated with low-demand problems in 20 and 80% of trials within each block, respectively. The length (number of trials) of the first, third, and fifth block was 20, 30, or 40 trials; their order was pseudorandomized for each participant. The same was applied to the length of the second, fourth, and sixth block. Notably, participants were not informed of when the task entered a new block, or even of the existence of these different blocks.

In both Experiments 1 and 2, the low-demand and high-demand problems had very similar visual properties and the same task rule, so the differences between them were limited to the level of cognitive demand. We instructed the participants that there were two types of problems and the probabilities that the two types appeared varied across trials individually in the left-arrow problems and the right-arrow problems. As described above, the ratios of low-demand and high-demand problems in the left-arrow and right-arrow problems in each block were symmetric, i.e., either 80 and 20% or 20 and 80%. However, the moving average rates of the problem types were not generally symmetric, as shown in Figure 1B, and so we expected that the participants felt that the left and right probabilities of the two types varied individually as instructed. In the initial instruction about the two problem types, we did not explain that the difference between the two types was in the level of cognitive demands (but see below for our communications with participants during instruction and practice sessions).

Before scanning, participants received instructions and practiced the tasks for 5–20 trials in Experiment 1 and 30–60 trials in Experiment 2 to get acclimated to the tasks and recognize that there were two types of problems. During the practice session(s), we continued to give oral instructions to and accept questions from participants. After the practice session(s), we asked the participants to explain what they thought was the difference between the two types. When participants’ reports reflected the actual difference in the problem content or when they said that they thought the difference was in the level of difficulty or largely similar things, we told them that we could not say the answer was correct or not. When the difference that participants reported was wildly incorrect, we told them that the answer was wrong and let them practice more, and when the participants subsequently reported things closer to reality, we (sometimes) said that the answer was not largely incorrect. Notably, to ensure that participants made their choices freely and without bias from us, we never told participants to avoid high-demand problems. Also, we instructed participants that their choices did not affect how fast they could finish the experiments so as to minimize the possible effects of this factor on participants’ decision making.

In the scanner, in both Experiments 1 and 2, there were in total 180 trials, which were divided into four sessions, each consisting of 45 trials (note that these sessions were different from the abovementioned six blocks). In each session, when a participant completed (chose an arrow cue and answered the problem) up to the 10th, 20th, 30th, and 45th (i.e., the last) trial in the session in <160, 320, 480, and 750 (in Exp. 1) or 900 (in Exp. 2) s, respectively, a rest period was imposed until 160, 320, 480, and 750 (in Exp. 1) or 900 (in Exp. 2) s had passed so as to make the progression of trials as independent as possible from participants’ choices. We instructed participants that the left and right probabilities of each type of problems would change continuously across sessions.

Image acquisition

We used a Siemens 3T Trio scanner (Brain Science Institute, Tamagawa University). When scanned, participants wore a head strap with memory foam to reduce their head movements. Participants viewed the experimental stimuli via a mirror reflecting the projector screen. Functional images were acquired in an ascending order and by using T2-weighted echo planar imaging sequence [repetition time (TR) = 2500 ms; echo time (TE) = 25 ms; field of view (FOV): 192 × 192 mm; in-plane resolution: 3 × 3 mm; acquisition matrix, 64 × 64; 42 slices with a slice

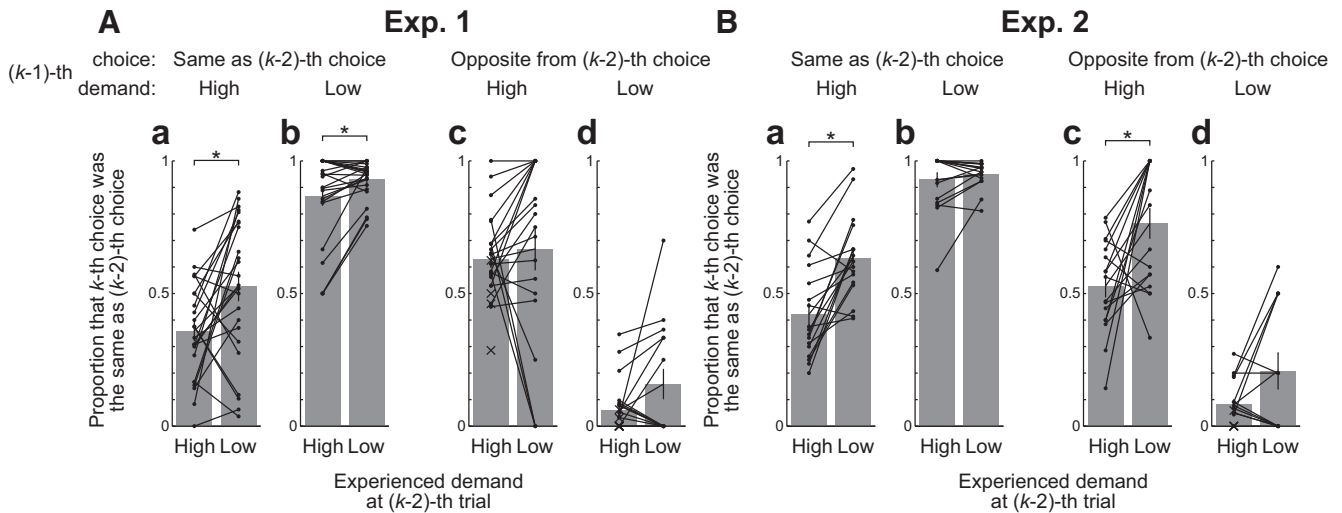


Figure 2. Effects of the experienced demand at two trials before on the current choice. The paired bars indicate the across-participants average proportions that the current (k -th) choice was the same as the choice at two trials before [$(k - 2)$ -th] when the experienced demand at the $(k - 2)$ -th trial was high (left bar) or low (right bar), for each case sorted by the choice at the $(k - 1)$ -th trial [same as (**a, b**) or opposite (**c, d**) from the $(k - 2)$ -th choice] and the experienced demand at the $(k - 1)$ -th trial [high (**a, c**) or low (**b, d**)] in Exp. 1 ($n = 24$; **A**) or Exp. 2 ($n = 17$; **B**). The average was taken across the participants who were judged as demand-avoiding based on the effects of the experienced demand at the previous trial (see Results). The error bars indicate the mean \pm SEM. The dots connected by lines indicate the data of individual participants who had paired data, while the crosses indicate the data of individual participants who lacked one of the paired data; both types of participants were included in the calculation of the average indicated by the bar heights. The paired data, represented by the dots, were compared by paired t test, and the cases indicated by asterisk were significant ($p < 0.05$): **Aa**, $p = 0.01$, $d = 0.73$, $t_{(23)} = 2.89$; **Ab**, $p = 0.02$, $d = 0.46$, $t_{(23)} = 2.59$; **Ba**, $p < 0.01$, $d = 1.27$, $t_{(16)} = 4.89$; **Bc**, $p = 0.01$, $d = 1.13$, $t_{(16)} = 2.98$.

thickness of 3.0 mm with no interslice gap]. In Experiment 1, ≥ 305 volumes were acquired, but just 305 were used for fMRI analysis in each of the four sessions for each participant. In Experiment 2, ≥ 365 volumes were acquired, but just 365 volumes were used for fMRI analysis in each of the four sessions for each participant. The first five scans in each session were discarded to allow for steady-state magnetization. We acquired high-resolution T1-weighted images using a MPRAGE sequence for all participants (TR = 2000 ms; TE = 1.98 ms; in-plane resolution, 1×1 mm; slice thickness, 1 mm).

Experimental design and statistical analysis

Behavioral analysis. In both experiments, one participant was excluded from analysis due to low correct rates ($< 80\%$ for both types). We first tested whether each participant chose the opposite option (left or right) after solving a high-demand problem more frequently than after solving a low-demand problem. Specifically, we conducted a χ^2 test on the contingency table consisting of the problem types (high or low demand) of the previous trials and the choices (same or opposite) of the current trials for each participant, and judged that s/he avoided high-demand problems if the frequency was significantly different ($p < 0.01$) and the adjusted standardized residuals were negative in the high \times same and low \times opposite cells of the contingency table. We next analyzed the effect of the demand experienced two trials prior (i.e., the trial before the previous trial) on the choice at the current trial for participants who showed avoidance of high-demand problems (judged through the χ^2 test as described above). Specifically, we compared the proportion that the k -th choice was the same as the $(k - 2)$ -th choice ($k = 3, 4, \dots, 180$) between the cases where the experienced demand at the $(k - 2)$ -th trial was high or low by using a paired t test as a planned comparison, for each case sorted by the $(k - 1)$ -th choice (same as or opposite from the $(k - 2)$ -th) and the $(k - 1)$ -th experienced demand (high or low). We sorted the data by the $(k - 1)$ -th choice and demand because otherwise spurious dependence could appear, i.e., the rate that the k -th choice was the same as the $(k - 2)$ -th choice could differ depending on the $(k - 2)$ -th demand even if choice was made depending solely on the choice and demand on the previous trial. It was generally rare that participants chose the opposite side after experiencing low demand, and there were participants who did not have data for the cases that included such a pattern (Fig. 2, crosses). Those unpaired data were omitted from the t test. As a measure of the effect size, we reported Cohen’s d calculated as follows:

$$d = \frac{\mu_2 - \mu_1}{\sqrt{(s_1^2 + s_2^2) / 2}}$$

where μ_1 , μ_2 and s_1 , s_2 are the means and sample SDs of each set of the paired data, respectively. The SEM was calculated by dividing the SD [defined with $1/\sqrt{n}$ (Figs. 2, 8) or $1/\sqrt{n - 1}$ (otherwise)] by \sqrt{n} , where n was the number of data included.

Model fitting, comparisons, simulations, and analyses of simulated behavioral data. We fitted the choices of the participants judged to have avoided high-demand problems in the χ^2 test using PE-based models (O’Doherty et al., 2007; Daw, 2011). We assumed that participants retained the expected cost (*ExpectedCost*) of mental effort for the left-arrow and right-arrow problems [*ExpectedCost_{left}*(k) and *ExpectedCost_{right}*(k) ($k = 1, 2, \dots, 180$: trials)]. At each trial k , either the left or right arrow was assumed to be chosen with the probabilities $P_{\text{left}}(k)$ and $P_{\text{right}}(k)$, respectively, depending on the expected cost, expressed as follows:

$$P_{\text{left}}(k) = \frac{\exp(-\beta \cdot \text{ExpectedCost}_{\text{left}}(k))}{\exp(-\beta \cdot \text{ExpectedCost}_{\text{left}}(k)) + \exp(-\beta \cdot \text{ExpectedCost}_{\text{right}}(k))}$$

$$P_{\text{right}}(k) = \frac{\exp(-\beta \cdot \text{ExpectedCost}_{\text{right}}(k))}{\exp(-\beta \cdot \text{ExpectedCost}_{\text{left}}(k)) + \exp(-\beta \cdot \text{ExpectedCost}_{\text{right}}(k))}$$

$$= 1 - P_{\text{left}}(k),$$

where β was a free parameter called the inverse temperature. If β was positive, the option with lower *ExpectedCost* was more frequently chosen, and the size of β represented the degree of exploitation over exploration. After solving a problem, the *CPE* was assumed to be calculated as follows: $CPE(k) = \text{ActualCost}(k) - \text{ExpectedCost}_{\text{Choice}}(k)$, where *ActualCost*(k) was the actual cost of the solved problem, and *Choice*(k) was Left or Right depending on which was chosen. As *ActualCost*, we considered five cases: (1) time spent solving the problem (solve time; in seconds) in individual trials [PE-Solve-Time (ST) model]; (2) demand level of the problem; more specifically, 1 and 0 for high-demand and low-demand problems, respectively [PE-High-Low (HL) model]; (3) incorrect solving; more

specifically, 1 and 0 for incorrect and correct solving, respectively (PE-Incorrect-Correct (IC) model); (4) sum of (1) and (3), with a weighting parameter for (3) ($w_{incorrect}$), i.e., solve-time + $w_{incorrect} \cdot incorrect$ -solving (PE-ST-IC model); and (5) sum of (2) and (3), with a weighting parameter for (3) ($w_{incorrect}$), i.e., demand-level + $w_{incorrect} \cdot incorrect$ -solving (PE-HL-IC model).

ExpectedCost for the chosen option was then assumed to be updated as follows: $ExpectedCost_{Choice(k)}(k+1) = ExpectedCost_{Choice(k)}(k) + \alpha \cdot CPE(k)$, where α was a free parameter representing the learning rate. *ExpectedCost* for the unchosen option was assumed to be unchanged. *ExpectedCost* for the option chosen at the first trial [$ExpectedCost_{Chosen}(1)$] was assumed as follows, depending on the five cases for *ActualCost*: (i) the mean solve time across all the trials in case (1) above (i.e., for the PE-ST model); (ii) 0.5 in case (2) (for the PE-HL model); (iii) the mean incorrect rate in case (3) (for the PE-IC model); (iv) sum of (i) and (iii) with the weighting parameter for (iii) in case (4) (for the PE-ST-IC model); and (v) sum of (ii) and (iii) with the weighting parameter for (iii) in case (5) (for the PE-HL-IC model).

ExpectedCost for the unchosen option at the first trial [$ExpectedCost_{unchosen}(1)$] was assumed to be either a free parameter or equal to $ExpectedCost_{Chosen}(1)$ (i.e., we examined both cases for each of the five cases for *ActualCost*): in total $5 \times 2 = 10$ models). Because the practice session(s) was rather short and we continued to give oral instructions to and accept questions from participants during the practice session(s), we considered it inappropriate to use performance measures during the practice session(s) for the initial values, and therefore we instead set the initial values as above.

In addition to these cost-based PE models, we also considered a PE model assuming that rest time in the scanner was a reward (*ActualReward*) for participants and they made choices based on the expectation of this reward (*ExpectedReward*) and updated *ExpectedReward* by reward PE (*RPE*; i.e., participants may have tried to choose low-demand problems so as to maximize “inactive time”). In this model, referred to as the PE-Rest model, *ActualCost*, *ExpectedCost*, and *CPE* were replaced with *ActualReward*, *ExpectedReward*, and *RPE*, respectively, and also β was replaced with $-\beta$. *ActualReward* was assumed to be the rest time (in seconds) from problem answer in the k -th trial to arrow presentation in the $k+1$ -th trial or to the end of session when the k -th trial was the last trial in the session. $ExpectedReward_{chosen}(1)$ and $ExpectedReward_{unchosen}(1)$ were assumed to be 4 s, given that the intertrial interval was 2.5, 4, or 5.5 s. We further considered a model combining this PE-Rest model and the best of the 10 cost-based PE models (which was the PE-HL model with $ExpectedCost_{unchosen}(1) = ExpectedCost_{Chosen}(1)$; see Results). In this model, referred to as the PE-HL-Rest model, *ExpectedCost* and *ExpectedReward* were assumed to be updated with individual learning rates (α_{cost} and α_{reward}), and choice was made based on the expected value that combined the expected cost and reward in a softmax manner with an inverse temperature β . The expected value was expressed as follows: $ExpectedValue_X(k) = w_{reward} \cdot ExpectedReward_X(k) - ExpectedCost_X(k)$, where X was Left or Right and w_{reward} was a free weighting parameter. The values of *ExpectedReward* and *ExpectedCost* at the first trial were assumed to be the same as those in the PE-Rest model and the PE-HL model with $ExpectedCost_{unchosen}(1) = ExpectedCost_{Chosen}(1)$.

In addition to the PE models, we also examined probabilistic Win-Stay-Lose-Shift (pWSLS) models, in which Win or Lose was followed by a selection of the same or different option, respectively, with exceptions with probability p that was a free parameter. We further examined full probabilistic-selection (PS) models, in which Win or Lose was followed by a selection of the same option with probabilities a and b , respectively, which were free parameters. For either type of model, Win and Lose were defined in two ways: (1) experiences of low-demand and high-demand problems, respectively (pWSLS-HL model and PS-HL model), or (2) solving correctly and incorrectly, respectively (pWSLS-IC model and PS-IC model).

For each of these models, we explored a set of free parameters that maximized the log-likelihood, expressed as follows:

$$\log L = \log \left\{ \prod_{k=1:180} P_{choice(k)}(k) \right\}$$

for each participant by using an optimization function (`fminsearch`) of the Matlab Optimization Toolbox and also a grid approach [more spe-

cifically, we conducted both exploration, using `fminsearch` assuming a set of initial parameters, and direct calculation of $\log L$ at various parameter values (grid points), and combined the results: see the codes for the analysis that is planned to be uploaded to a public database after publication]. For the pWSLS and PS models, $P_{choice}(1)$ was assumed to be 0.5. To compare the goodness of fit of each model while taking into account the penalty for a larger number of parameters, we calculated the Bayesian information criterion (BIC) as follows: $BIC = -2 \log L + q \log(180)$, where q was the number of free parameters.

Using the best-BIC PE model [PE-HL model with $ExpectedCost_{unchosen}(1) = ExpectedCost_{Chosen}(1)$; see Results], pWSLS-HL model, and PS-HL model, we performed simulations of task execution (180 trials) with the best-fit parameters for each individual demand-avoiding participant in each experiment. Specifically, for each demand-avoiding participant (judged based on the effects of the experienced demand at the previous trial through the above-mentioned χ^2 test) in each experiment, we extracted the best-fit parameters for each of the three models. Then, using these parameters and the actual sequences of low-demand and high-demand problems for the left-arrow and right-arrow cues used in the experiments, we generated 180 (number of trials) choices 100 times (i.e., performed 100 simulation runs) by using different sets of pseudo-random numbers in Matlab. We then analyzed the effect of the demand experienced two trials prior on the choice at the current trial in the pooled simulated choices for each participant. Specifically, we calculated the proportion that the k -th choice was the same as the $(k-2)$ -th choice when the $(k-2)$ -th demand was high or low, for each case sorted by the choice and demand on the $(k-1)$ -th trial, in the pooled simulated choices [178 (the initial two trials were omitted from the total 180 trials) \times 100 = 17,800 simulated choices] corresponding to each participant.

Functional imaging analysis. We used SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) for fMRI data processing and analysis. We realigned the volumes to the first images using a six-parameter rigid-body transformation. We corrected timing differences for each slice and normalized individual images. We applied a Gaussian kernel with a full-width at half-maximum of 8 mm for spatial smoothing. After excluding six participants from Experiment 1 and three participants from Experiment 2 with >3 mm head movements from those who met the performance criterion and showed avoidance of high-demand problems (see Results), we conducted general linear model (GLM) analysis of BOLD data (Exp. 1, $n = 21$; Exp. 2, $n = 15$). As pointed out by Mumford et al. (2015), when multiple parametric modulations exist for the regressor at the same time, SPM8 performs orthogonalization by default. We turned off this default operation by commenting out line 228 of `spm_get_ons.m` and lines 277–279 of `spm_fmri_design.m`, which call `spm_orth.m`, in reference to <http://imaging.mrc-cbu.cam.ac.uk/imaging/ParametricModulations> (but the line numbers that were commented out differed from those described on this website). All individual and group analyses in each experiment were done at the whole-brain level.

At the individual level, we examined the following three GLMs (Fig. 3) designed to explore the correlates of *ExpectedCost* for the chosen option (referred to as $ExpectedCost_{Chosen}$) and the *CPE*, adjusted for the response time for choosing an arrow (referred to as RT_{choice}), actual demand level of the problem (referred to as problem-demand), and solve time. These GLMs included the regressors at arrow-cue presentation with parametric modulations by $ExpectedCost_{Chosen}$ (derived from the best-BIC PE model; see Results) and RT_{choice} regressors with the duration from problem presentation to answer with parametric modulations by problem-demand (0 and 1 for low-demand and high-demand problems, respectively) and solve time, regressor with parametric modulation by *CPE* at the time of problem presentation (GLM1), midpoint between problem presentation and answer (GLM2), or time of answer (GLM3), and regressors for motor response (at both arrow choice and answer in GLM1 and GLM2 and only at arrow choice in GLM3) and head movements. We also considered variants of GLM1, which are described in Results. For each of these GLMs, we convolved each regressor with the SPM8's canonical hemodynamic response function and performed one-sample t tests for individual maps for the regressor(s) of interest across 21 and 15 participants in Experiments 1 and 2, respectively. We calculated the variance inflation factor (VIF) using the Canlab Matlab toolboxes (<https://>

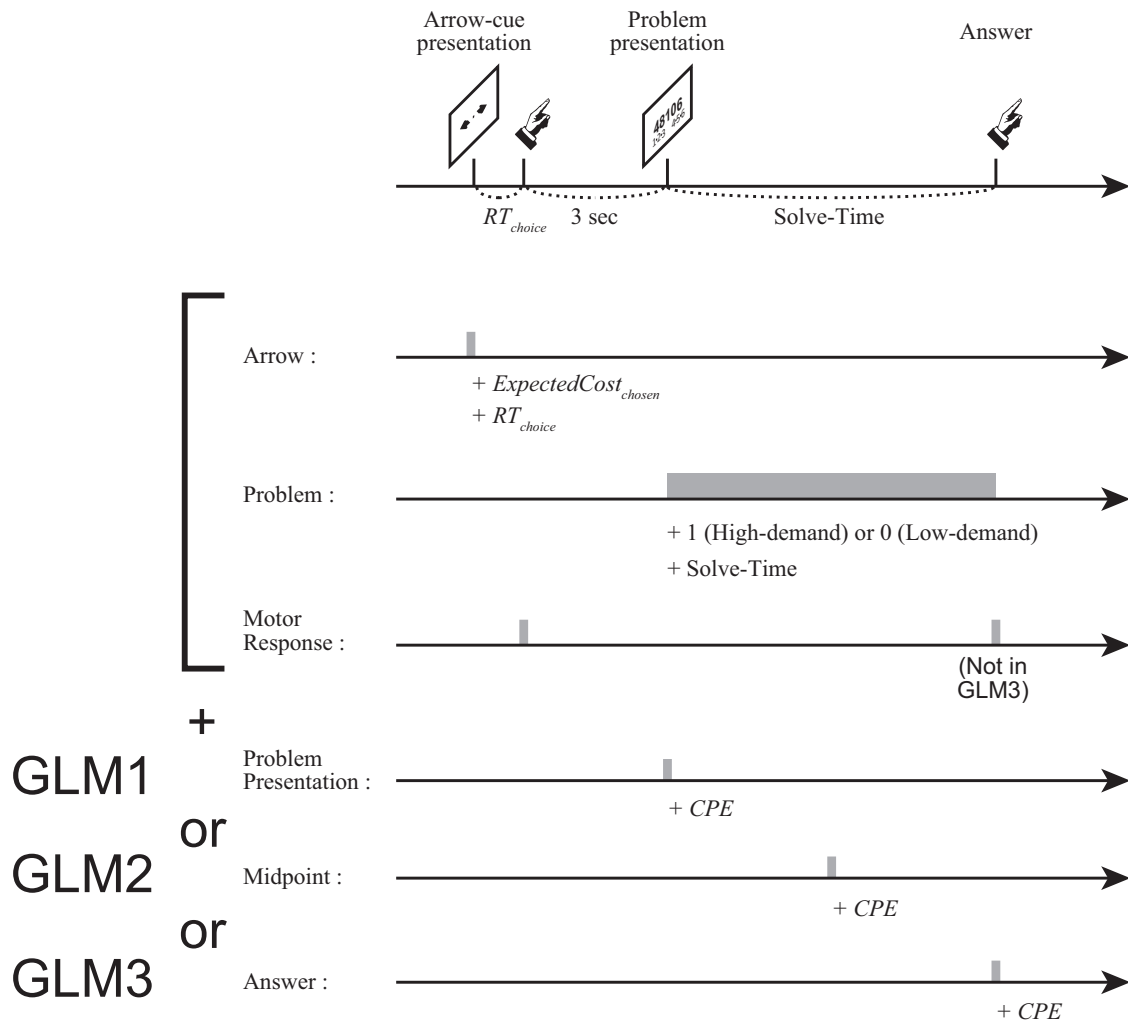


Figure 3. GLMs and regressors used in the fMRI analyses. We explored the correlates of the expected cost of mental effort for the chosen option ($ExpectedCost_{chosen}$) and the CPE estimated in the model by using three GLMs (GLM1–GLM3), which assumed three different possibilities regarding the time of CPE generation/representation. Each of these GLMs included the regressors at arrow-cue presentation with nonorthogonalized parametric modulations by $ExpectedCost_{chosen}$ and RT_{choice} regressors starting at problem presentation and having the duration of solve-time with nonorthogonalized parametric modulations by demand level (1 and 0 for high-demand and low-demand problems, respectively) and solve time, regressors for motor response at both arrow choice and answer (GLM1 and GLM2) or at arrow choice (GLM3), regressors for head movements (not illustrated here), and regressor at problem presentation (GLM1), midpoint between problem presentation and answer (GLM2), or answer (GLM3) with parametric modulation by CPE.

github.com/canlab/CanlabCore) and judged whether collinearity of the regressor of interest was at a tolerable level considering that 5 or 10 is typically used as a cutoff value of VIF for the collinearity issue (Mumford et al., 2015).

At the group level, we reported correlates detected by GLM1–GLM3 in each experiment with a threshold of cluster-level familywise error (FWE) corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$ for the cases where at least one cluster was found with this threshold, or more specifically, for the positive correlates of $ExpectedCost_{chosen}$ in GLM1–GLM3 in both experiments, negative correlates of $ExpectedCost_{chosen}$ in GLM1–GLM3 in Experiment 1, positive correlates of CPE in GLM1–GLM3 in both experiments, and negative correlates of CPE in GLM1 and GLM3 in Experiment 2. For the other cases where results for individual experiments were reported, or more specifically, for the negative correlates of $ExpectedCost_{chosen}$ in GLM1–GLM3 in Experiment 2 and negative correlates of CPE in GLM1–GLM3 in Experiment 1 and in GLM2 in Experiment 2, we reported correlates with a threshold of voxel-level uncorrected $p < 0.001$ with voxel-size of ≥ 5 if we found any.

To detect common regions in the correlates found in Experiments 1 and 2, we conducted conjunction analyses, to which we applied a binary mask. We used two masks with different thresholds: the strict mask and the relaxed mask. The strict mask consisted of common voxels between

the results of Experiments 1 and 2 with the threshold of cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$. The relaxed mask consisted of common voxels with the threshold of voxel-level uncorrected $p < 0.01$ (which mask was used for which analyses is described in Results and tables; the relaxed mask was used when no cluster was detected with the threshold of the strict mask in either experiment or no cluster was detected as a result of conjunction analysis with the strict mask). We then reported correlates detected in the masked conjunction analyses with a threshold of cluster-level uncorrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$.

Results

Behavioral tasks and analyses

We conducted two experiments. These had the same task structure but used different types of problems that imposed different kinds of cognitive demand (Fig. 1A). In Experiment 1, we used mental arithmetic problems. We asked participants to divide a five-digit number by 7 and report whether the remainder was small (≤ 3) or large (≥ 4). In Experiment 2, we used spatial reasoning (mental cube-folding) problems. We asked participants to judge whether a 3D cube with three visible colored faces matched a con-

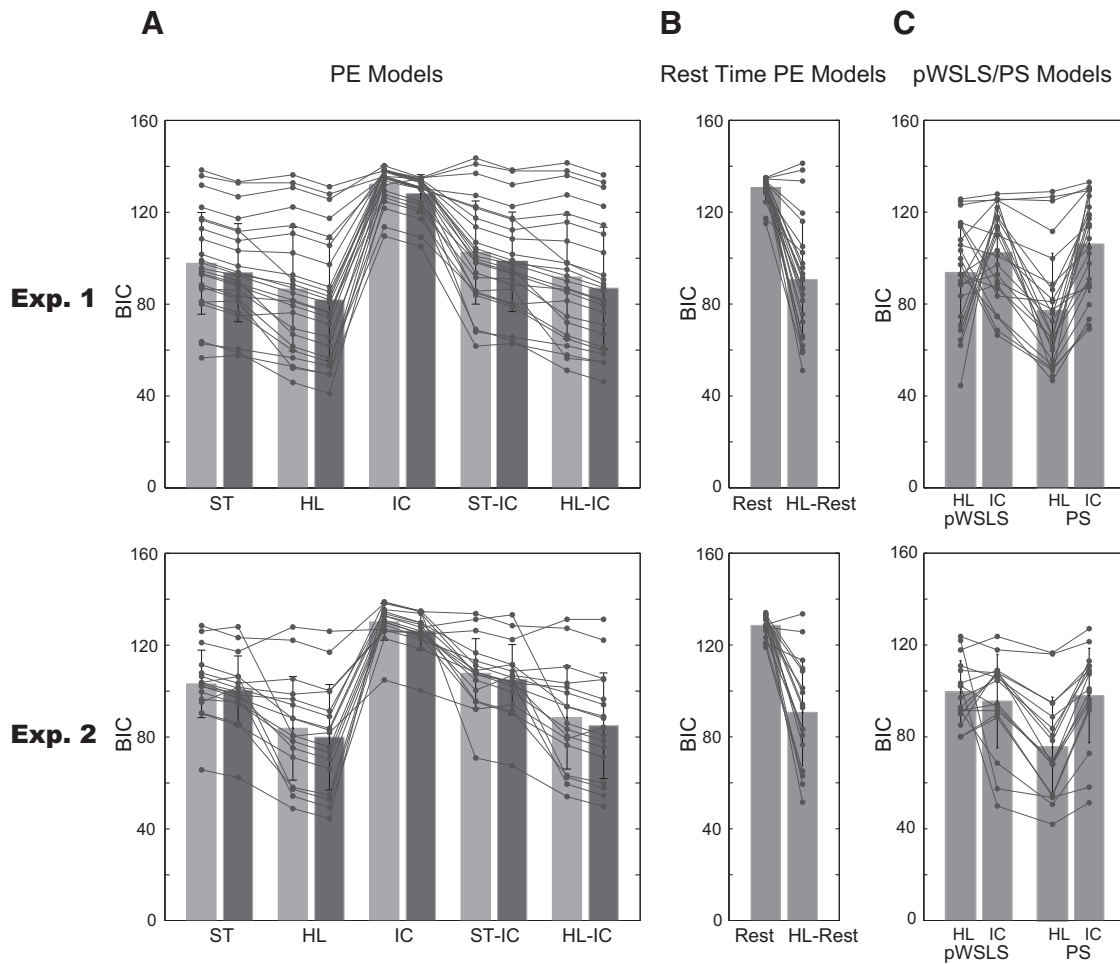


Figure 4. BIC scores of the models fitted to the choices of demand-avoiding participants. **A**, The bars indicate the mean \pm SEM of BIC scores for 10 variants of PE models in Experiments 1 (Exp. 1) and 2 (Exp. 2). The horizontal axis indicates the 10 PE models: five assumptions on *ActualCost* [Solve-Time (ST), High-Low (HL), Incorrect-Correct (IC), ST-IC, and HL-IC: see the Results for details] \times 2 assumptions on *ExpectedCost_{unchosen}*(1) [free parameter (light-gray bars) or equal to *ExpectedCost_{chosen}*(1) (dark-gray bars)]. The black dots connected with the lines indicate individual demand-avoiding participants. **B**, Results for the additionally considered PE models that regarded the rest time in the scanner as reward for participants. **C**, Results for pWSLS models and PS models.

currently presented unfolded cube or not. For both experiments, we prepared two sets of problems that required different levels of cognitive demand, i.e., low-demand problems and high-demand problems. In Experiment 1, the dividend in low-demand problems (e.g., 35426) consisted of two consecutive two-digit numbers that were multiples of 7 followed by a single one-digit number from 1 to 6. Meanwhile, the dividend in high-demand problems (e.g., 48106) did not contain any numbers that were multiples of 7. In Experiment 2, the difference between low-demand and high-demand problems was whether the three faces shown on the 3D cube were neighboring on the unfolded cube. In both experiments, the probability that a high-demand or low-demand problem appeared at each trial depended on a cue that participants chose at the start of the trial: there were two cues, the left and right arrows, and the probabilistic associations between each of the cues and high-demand and low-demand problems changed across trials, such as shown in Figure 1B. After participants chose a cue, a problem, either high or low demand, was presented, and they were asked to answer it. There was no time limit for response.

There were 33 participants (four females; mean age, 25.5 ± 5.4 years) in Experiment 1 and 20 participants (two females; mean age, 24.7 ± 6.2 years) in Experiment 2. Six participants took part in both experiments. Most of the participants were males, so the

results cannot with certainty be generalized to females. The response time for choosing an arrow cue (RT_{choice}) was 0.95 ± 0.08 s (mean \pm SEM) in Experiment 1 and 1.25 ± 0.21 s in Experiment 2. The mean correct answer rates for high-demand and low-demand problems were 0.94 ± 0.008 (mean \pm SEM) and 0.98 ± 0.010 , respectively, in Experiment 1, and 0.92 ± 0.024 and 0.99 ± 0.004 respectively, in Experiment 2. In both experiments, correct answer rates for low-demand problems were higher than those for high-demand problems on average (paired t test, $t_{(32)} = -3.0, p = 4.9 \times 10^{-3}$, in Exp. 1; $t_{(19)} = -2.7, p = 1.3 \times 10^{-2}$, in Exp. 2). The mean solve times (i.e., times for problem solving) for high-demand and low-demand problems were 10.00 ± 0.70 s (mean \pm SEM) and 1.86 ± 0.09 s, respectively, in Experiment 1, and 12.44 ± 1.13 s and 4.60 ± 0.32 s, respectively, in Experiment 2. In both experiments, participants took longer for high-demand problems on average (paired t test, $t_{(32)} = 12.5, p = 8.1 \times 10^{-14}$, in Exp. 1; $t_{(19)} = 8.3, p = 1.1 \times 10^{-7}$, in Exp. 2). To ensure the quality of data used, we set a performance criterion for inclusion. Specifically, we assumed a participant faithfully executed the problems if the correct answer rate was ≥ 0.8 for either low-demand or high-demand problems. As a consequence, one participant in each experiment was excluded from the following analyses.

As an initial analysis of the participants' learning and choice behavior, we inferred whether each participant learned to avoid

Table 1. BIC scores and best-fit parameters of the models fitted to the choices of demand-avoiding participants

Models	BIC/parameters	Experiment 1 (<i>n</i> = 24)			Experiment 2 (<i>n</i> = 17)		
		Median	25th percentile	75th percentile	Median	25th percentile	75th percentile
A, PE models							
ST	BIC	95.1	84.0	114.8	103.1	96.1	108.8
EC _{unchosen} (1): free	EC _{unchosen} (1)	7.51	4.69	10.61	11.05	6.44	13.16
	α	0.65	0.53	0.85	0.41	0.32	0.60
	β	0.29	0.17	0.61	0.28	0.18	0.42
ST	BIC	91.3	81.8	109.7	99.7	95.0	106.3
EC _{unchosen} (1) = EC _{chosen} (1)	α	0.66	0.57	0.87	0.53	0.38	0.65
	β	0.25	0.16	0.53	0.25	0.17	0.41
	BIC	83.3	64.2	106.5	80.7	67.9	97.0
EC _{unchosen} (1): free	EC _{unchosen} (1)	0.75	0.67	0.87	0.75	0.71	0.96
	α	0.84	0.72	0.96	0.85	0.63	0.92
	β	3.13	2.23	4.49	3.32	2.47	4.48
HL	BIC	78.5	59.2	101.4	81.9	63.3	93.4
EC _{unchosen} (1) = EC _{chosen} (1)	α	0.80	0.71	0.96	0.87	0.64	0.97
	β	3.13	2.23	4.46	3.15	2.44	4.32
	BIC	135.2	128.6	138.0	130.9	126.8	134.9
EC _{unchosen} (1): free	EC _{unchosen} (1)	0.70	0.49	1.11	0.90	0.41	1.29
	α	0.23	0.15	0.38	0.20	0.10	0.40
	β	2.18	−4.15	3.14	2.43	−0.44	3.86
IC	BIC	130.7	123.9	134.0	126.2	124.5	130.8
EC _{unchosen} (1) = EC _{chosen} (1)	α	0.29	0.18	0.46	0.20	0.12	0.32
	β	2.90	0.72	3.39	3.60	2.24	4.31
	BIC	100.3	85.6	119.7	108.0	99.2	114.0
EC _{unchosen} (1): free	EC _{unchosen} (1)	7.59	4.70	10.70	11.30	6.44	14.44
	α	0.65	0.53	0.84	0.40	0.32	0.60
	β	0.29	0.17	0.60	0.28	0.18	0.40
	$w_{incorrect}$	0.67	0.00	6.96	0.00	0.00	0.00
ST-IC	BIC	96.4	83.4	114.6	104.1	98.3	111.5
EC _{unchosen} (1) = EC _{chosen} (1)	α	0.67	0.57	0.86	0.53	0.38	0.65
	β	0.26	0.16	0.53	0.26	0.17	0.36
	$w_{incorrect}$	0.00	0.00	3.14	0.00	0.00	1.28
	BIC	88.5	69.2	111.6	85.9	73.1	102.1
EC _{unchosen} (1): free	EC _{unchosen} (1)	0.75	0.67	0.89	0.76	0.70	0.96
	α	0.83	0.72	0.93	0.85	0.64	0.92
	β	3.38	2.18	4.40	3.16	2.36	4.58
	$w_{incorrect}$	0.00	0.00	0.14	0.00	0.00	0.08
	BIC	83.7	64.2	106.5	84.3	68.5	98.6
EC _{unchosen} (1) = EC _{chosen} (1)	α	0.80	0.71	0.96	0.87	0.64	0.97
	β	3.13	2.23	4.36	3.15	2.44	4.32
	$w_{incorrect}$	0.00	0.00	0.01	0.00	0.00	0.00
	BIC	132.3	129.0	134.0	131.0	125.5	132.1
B, PE-Rest models							
Rest	BIC	132.3	129.0	134.0	131.0	125.5	132.1
	α	0.40	0.06	0.97	0.12	0.06	0.45
	β	0.01	−0.05	0.11	0.05	−0.05	0.12
HL-Rest	BIC	88.7	69.2	103.8	92.8	73.6	109.0
	α_{cost}	0.80	0.68	0.96	0.87	0.63	0.97
	α_{reward}	0.00	0.00	0.03	0.00	0.00	0.01
	β	−3.21	−4.54	−2.45	−3.14	−4.32	−2.24
	w_{reward}	−0.03	−0.43	0.60	0.76	−0.35	2.02
C, pWSLS/PS models							
pWSLS-HL	BIC	92.9	79.0	110.8	97.1	91.1	109.4
	<i>p</i>	0.19	0.14	0.27	0.21	0.18	0.27
pWSLS-IC	BIC	111.1	88.6	124.7	100.9	92.7	110.8
	<i>p</i>	0.76	0.67	0.85	0.80	0.76	0.83
PS-HL	BIC	69.4	57.8	87.4	70.1	54.6	90.2
	<i>a</i>	0.95	0.91	0.98	0.97	0.91	0.98
	<i>b</i>	0.47	0.30	0.58	0.48	0.46	0.58
PS-IC	BIC	111.1	88.6	124.7	100.9	92.7	110.8
	<i>a</i>	0.76	0.67	0.85	0.80	0.76	0.83
	<i>b</i>	0.50	0.41	0.55	0.66	0.50	0.77

This table shows the median, 25th, and 75th percentiles of BIC scores and best-fit parameter estimates for 10 variants of PE models (5 assumptions on *ActualCost* [Solve-Time (SV), High-Low (HL), Incorrect-Correct (IC), ST-IC, and HL-IC] × 2 assumptions on EC_{unchosen}(1) [free parameter or equal to EC_{chosen}(1)]), two types of additionally considered PE models (Rest and HL-Rest models), and pWSLS and PS models. EC_{unchosen} the expected cost for the unchosen option; EC_{chosen} the expected cost for the chosen option; α , learning rate; β , inverse temperature; $w_{incorrect}$, weighting parameter; α_{cost} the learning rate for cost; α_{reward} the learning rate for reward; w_{reward} , weighting parameter; *p*, probability to select the different and same option after Win and Lose, respectively; *a*, probability to select the same option after Win; *b*, probability to select the same option after Lose.

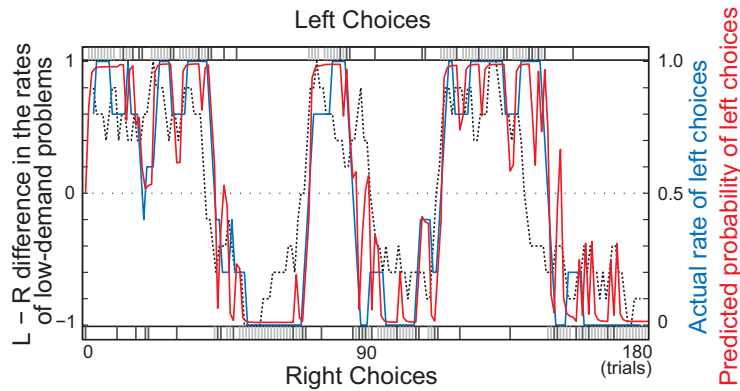


Figure 5. Example of participant's choices and choice probability predicted by the PE-HL model. Short vertical bars at the top and bottom indicate participant's left and right choices, respectively, with dark or light gray indicating that high-demand or low-demand problems were experienced, respectively. The black dashed line in the middle indicates the left–right difference in the presentation rates (moving average of latest 5 trials) of low-demand problems plotted against the left scale. The blue and red solid lines indicate the participant's actual left-choice rate (moving average of latest 5 trials) and the left-choice probability predicted by the PE-HL model plotted against the blue and red scales on the right, respectively. The best-fit parameters for this participant (in Exp. 1) were as follows: [learning rate, inverse temperature] = [0.68, 3.62].

high-demand problems from the dependence of choices on the previous trials. We reasoned that participants wanting to avoid high-demand problems (whether consciously or not) would stay at the same side (left or right) if a low-demand problem appeared in the previous trial but would rather switch to the opposite side if a high-demand problem appeared. We thus examined whether such a bias existed by conducting a χ^2 test [on 2×2 factors: problem types (high or low demand) in the previous trial \times choice (same side or opposite side) in the current trial]. In the results, significant bias ($p < 0.01$) existed in 26 of 32 (81.3%) and 17 of 19 (89.5%) participants in Experiments 1 and 2, respectively. Among these cases, 24 of 32 (75.0%) and 17 of 19 (89.5%) participants in Experiments 1 and 2, respectively, showed avoidance of high-demand problems. This indicates that these participants (i.e., the majority) learned to avoid high-demand problems in the situation where the probabilistic associations between cues and demand levels changed over time. Overall, these demand-avoiding participants in Experiments 1 and 2 experienced low-demand problems in $63.6 \pm 1.0\%$ (mean \pm SEM) and $64.6 \pm 1.2\%$, respectively, and chose the same option as in the previous trial in $74.9 \pm 2.1\%$ and $79.5 \pm 1.8\%$ (in Trials 2–180), respectively. On the other hand, the remaining two participants in Experiment 1 had the opposite bias, indicating that this minority of participants learned (chose) to seek high-demand problems.

Next, we analyzed whether the choices of the demand-avoiding participants as judged above depended also on the demand experienced two trials before the present trial, i.e., in the trial before the previous trial. The paired bars in Figure 2 show the proportions that the choice at the k -th trial was the same as the choice at the $(k - 2)$ -th trial when the demand experienced at the $(k - 2)$ -th trial was high (left bar) or low (right bar), for each case sorted by the choice and demand at the $(k - 1)$ -th trial in Experiment 1 (Fig. 2Aa–d) and Experiment 2 (Fig. 2Ba–d). As shown in the figure, the proportion that the k -th choice was the same as the $(k - 2)$ -th choice was significantly higher when the demand experienced at the $(k - 2)$ -th trial was low than when it was high, with medium-to-large effect sizes, in two cases in Experiment 1 (Fig. 2Aa; $p = 0.01$, $d = 0.73$, $t_{(23)} = 2.89$; Fig. 2Ab; $p = 0.02$, $d = 0.46$, $t_{(23)} = 2.59$) and in two cases in Experiment 2 (Fig. 2Ba; $p < 0.01$, $d = 1.27$, $t_{(16)} = 4.9$; Fig. 2Bc; $p = 0.01$, $d = 1.13$, $t_{(16)} = 2.98$). In this way, in both

experiments, the choices of the demand-avoiding participants did depend on the demand experienced two trials prior.

Detailed analyses of learning and choice behavior

To analyze learning and choice behavior in detail, we fitted the choices using PE-based models (O'Doherty et al., 2007; Daw, 2011), considering that PE-based models have been suggested to be able to approximate reinforcement learning of reward values (McClure et al., 2003; O'Doherty et al., 2003; Daw et al., 2006) as well as avoidance learning of pain (Seymour et al., 2004; Roy et al., 2014; Zhang et al., 2016), physical-effort cost (Skvortsova et al., 2014), or sustained effort (selecting circles on the screen) concurrently with reward learning (Scholl et al., 2015). In particular, we assumed that (1) participants had (whether consciously or not) expectations of the cost of mental effort, referred to as

the *ExpectedCost* below, needed to solve a left or right problem (denoted by *ExpectedCost_{left}* and *ExpectedCost_{right}*), (2) participants chose either the left or the right problem according to *ExpectedCost_{left}* and *ExpectedCost_{right}* in a “softmax” manner, i.e., avoided an option with a higher *ExpectedCost* with a higher probability, and (3) the *ExpectedCost* for the chosen option (denoted by *ExpectedCost_{Chosen}*) was updated according to the CPE: $ActualCost - ExpectedCost_{Chosen}$, where the *ActualCost* was the cost actually experienced.

Given that participants took much longer times for high-demand problems than for low-demand problems on average as shown above, it is possible that the time spent to solve the problem constituted *ActualCost*, while it is also conceivable that *ActualCost* directly reflected the demand level of the problem itself. Moreover, because the correct answer rates also differed between the low-demand and high-demand problems, incorrect solving could also constitute or contribute to *ActualCost*, even though the incorrect answer rates were rather low and we did not provide correct/incorrect feedback to participants. With these considerations, we considered the following five cases for constituent(s) of *ActualCost*: (1) time spent solving the problem (solve time) in individual trials [referred to as the PE-Solve-Time(ST) model]; (2) demand level of the problem; more specifically, 1 and 0 for high-demand and low-demand problems, respectively [PE-High-Low(HL) model]; (3) incorrect solving; more specifically, 1 and 0 for incorrect and correct solving, respectively [PE-Incorrect-Correct(IC) model]; (4) sum of (1) and (3), with a weighting parameter for (3) (PE-ST-IC model); and (5) sum of (2) and (3), with a weighting parameter for (3) (PE-HL-IC model).

For each of these five cases, we considered two models assuming that *ExpectedCost* for the unchosen option (*ExpectedCost_{unchosen}*) at the first trial was either a free parameter or a value equal to *ExpectedCost_{Chosen}*, resulting in $5 \times 2 = 10$ models.

We fitted these models to the participants' choices, individually for each participant who showed avoidance of high-demand problems (as judged by the χ^2 test above), by exploring parameters that maximized the log-likelihood. We then compared the fitted models according to the BIC. As a result, the model assuming case (2) *ActualCost* [i.e., the PE-High-Low(HL) model] and $ExpectedCost_{unchosen} = ExpectedCost_{Chosen}$ at the first trial had the

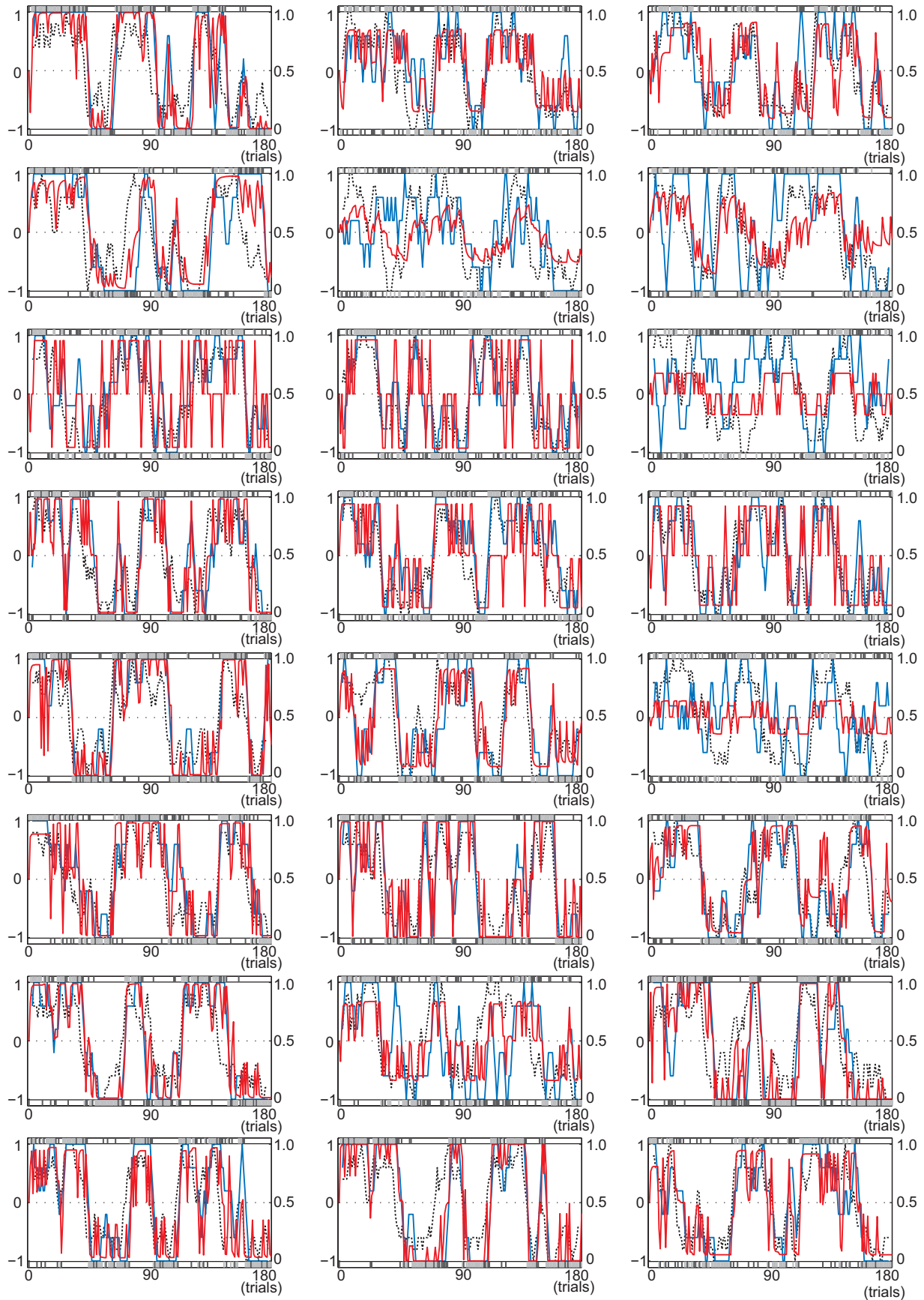


Figure 6. Results of model fitting for all the participants who showed avoidance in Experiment 1. The configurations are the same as those of Figure 5.

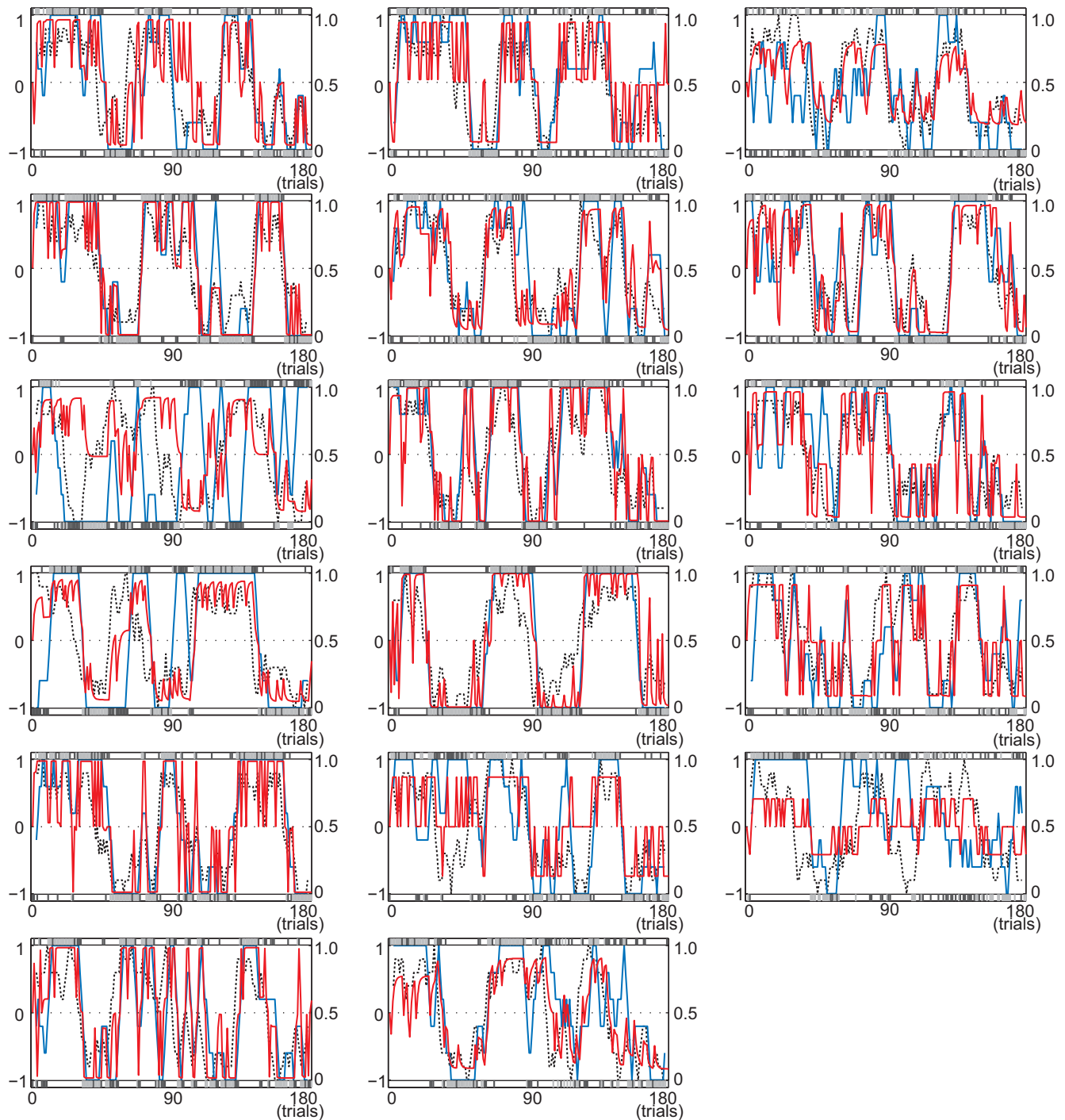


Figure 7. Results of model fitting for all the participants who showed avoidance in Experiment 2. The configurations are the same as those of Figure 5.

best (i.e., least) BIC score for most of the participants in both experiments (23 of 24 in Exp. 1; 14 of 17 in Exp. 2; Fig. 4A; Table 1; hereafter we refer to this model as the PE-HL model). This result indicates that mental-effort cost in our experiments was experienced and/or registered as (nearly) binary variables corresponding to the binary demand levels of the problems, rather than variables reflecting the solve time or mistakes (we will return to this later). An example of the fit by this model is shown in the red solid line in Figure 5, and the results of all the analyzed participants in Experiments 1 and 2 are shown in Figures 6 and 7, respectively.

We additionally examined two more PE models that assumed that the rest time in the scanner was reward for participants and

they made choices based on the expectation of this reward (referred to as the PE-Rest model) or on the expectations of both this reward and the mental-effort cost, which was assumed to be 0 and 1 for low-demand and high-demand problems, respectively, inheriting the assumption of the PE-HL model that gave the best BIC score (referred to as the PE-HL-Rest model). However, these models gave larger (i.e., worse) BIC scores than the PE-HL model in almost all cases (except for one participant in Exp. 1 for both PE-Rest and PE-HL-Rest models; Fig. 4B).

We also examined models having different structures from the PE models. In particular, we considered a pWLS model, in which Win or Lose was followed by a selection of the same or

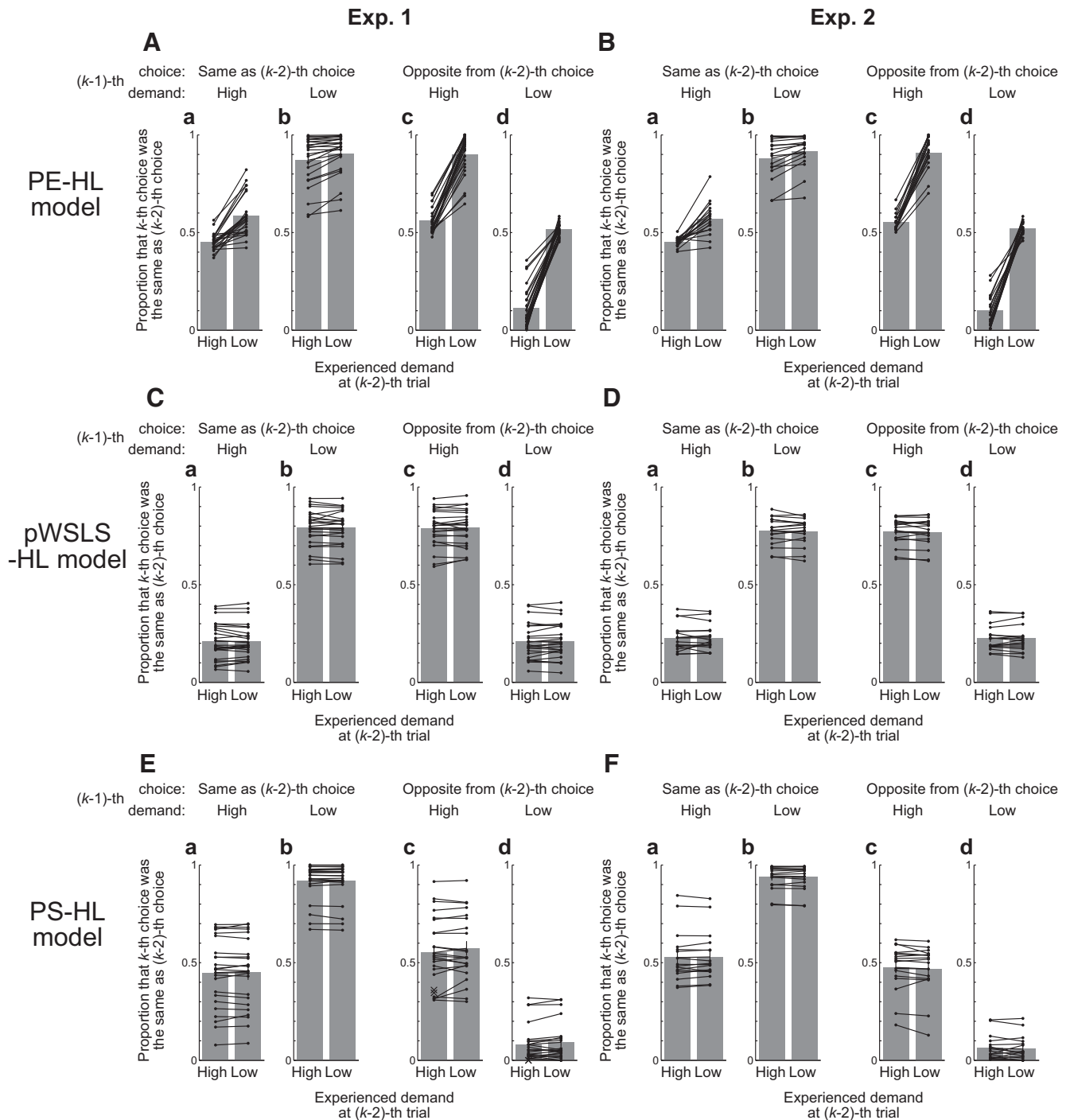


Figure 8. Effects of the experienced demand at two trials before on the current choice in pooled simulated choices. **A, B**, The dots connected by lines indicate the proportions that the k -th choice was the same as the $(k - 2)$ -th choice when the $(k - 2)$ -th demand was high (left) or low (right), sorted by the $(k - 1)$ -th choice and demand (**a–d**), in pooled simulated choices corresponding to each demand-avoiding participant, which were generated by performing 100 simulation runs of task execution (180 trials) in Experiment 1 (Exp. 1; **A**) or Experiment 2 (Exp. 2; **B**) with actual demand sequences (high or low in the left or right) used for the participant by the PE-HL model with best-fit parameters for the participant. The bars indicate the average of the proportions corresponding to individual participants, and the error bars indicate the mean \pm SEM. **C–F**, Same as **A** and **B** except that the pWLSL-HL model (**C, D**) or the PS-HL model (**E, F**) was used instead of the PE-HL model.

different option, respectively, with exceptions with a certain probability that was a free parameter. Win and Lose were defined in two ways: (1) experiences of low-demand and high-demand problems, respectively (pWLSL-HL model), or (2) solving correctly and incorrectly, respectively (pWLSL-IC model). As a result, for either type of pWLSL models, participants for whom the given type of pWLSL model gave smaller (i.e., better) BIC scores than the PE-HL model (i.e., the best-BIC PE model) were out-

numbered by those who had the opposite pattern (pWLSL-HL model: 14 of 24 in Exp. 1; 16 of 17 in Exp. 2; pWLSL-IC model: 20 of 24 in Exp. 1; 15 of 17 in Exp. 2; Fig. 4C). We further considered a full probabilistic selection (PS) model, in which Win or Lose was followed by a selection of the same or a different option with arbitrary probabilities that were free parameters, with the same two definitions of Win and Lose as above (PS-HL model and PS-IC model). As a result, the number of participants for whom

Table 2. Neural correlates of the expected cost of mental effort for the chosen option revealed by using GLM1

Area	Left, right, bilateral	Cluster		Peak (T)	MNI coordinates (mm)		
		Size	p(FWE)		x	y	z
A, Experiment 1							
Positive correlation							
MFG/inferior frontal gyrus/anterior insula	Right	691	<0.001	6.0	33	50	4
aMFG/anterior insula	Left	213	0.001	6.2	−30	47	7
dmFC/dACC	Bilateral	354	<0.001	6.3	0	14	49
Inferior parietal lobule	Right	230	<0.001	5.6	45	−52	49
Negative correlation							
Rostromedial PFC	Bilateral	137	0.006	5.4	0	59	13
vmPFC/striatum	Bilateral	291	<0.001	7.6	3	41	−17
Primary motor cortex/primary somatosensory cortex	Right	92	0.031	5.3	42	−19	64
Temporal lobe	Left	749	<0.001	6.7	−42	−37	25
Precuneus	Left	180	0.002	5.2	−6	−67	22
B, Experiment 2							
Positive correlation							
aMFG/orbitofrontal cortex	Left	129	0.010	7.6	−33	44	4
Anterior insula	Left	86	0.044	6.1	−27	26	4
dmFC/dACC/superior frontal gyrus	Bilateral	520	<0.001	7.4	6	20	49
Anterior insula/aMFG/orbitofrontal cortex	Right	582	<0.001	13.7	33	20	4
Inferior parietal lobule	Right	142	0.006	5.2	48	−49	55
Cerebellum	Left	143	0.006	8.5	−18	−61	−41
Negative correlation							
Striatum ^a	Bilateral	30	0.425	4.5	0	5	−8
C, Conjunction analysis							
Positive correlation (strict mask ^b)							
dmFC/dACC	Bilateral	215	0.002	6.3	6	20	46
Anterior insula ^c	Right	57	0.164	4.9	30	20	4
aMFG ^c	Left	50	0.211	5.5	−30	47	7
aMFG	Right	212	0.002	5.6	33	44	1
Inferior parietal lobule ^c	Right	75	0.088	4.9	45	−49	49
Negative correlation (relaxed mask ^d)							
vmPFC ^e	Bilateral	68	0.112	4.3	−6	53	−5

A, B, The brain area, laterality of hemisphere, cluster size, cluster *p* value, peak *T* value, and peak coordinate of the regions that were positively and negatively correlated with the expected cost of mental effort for the chosen option (*ExpectedCost_{Chosen}*) predicted by the PE-HL model in Experiment 1 (*n* = 21; A) or Experiment 2 (*n* = 15; B) revealed by using GLM1. Statistical thresholds were set to be cluster-level FWE corrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001 except for the case of the negative correlation in Experiment 2, for which the thresholds were voxel-level uncorrected *p* < 0.001 and voxel size ≥ 5. C, The information for the regions obtained in the conjunction analyses shown in Figure 9A. Methods for analyses and statistical thresholds were the same as those used for Figure 9.

^aVoxel-level uncorrected *p* < 0.001 and voxel size ≥ 5.

^bA binary mask consisting of common voxels at cluster-level FWE corrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001.

^cCluster-level uncorrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001; with a binary mask consisting of common voxels at cluster-level FWE corrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001.

^dA binary mask consisting of common voxels at voxel-level uncorrected *p* < 0.01.

^eCluster-level uncorrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001; with a binary mask consisting of common voxels at voxel-level uncorrected *p* < 0.01.

the PS-HL model gave a smaller (better) BIC score than the PE-HL model was comparable to the number of those who had the opposite pattern in Experiment 1 [12 participants for each; though the average BIC score across participants was smaller (better) in the PS-HL model], and the PS-HL model gave a smaller (better) BIC score than the PE-HL model in a large number of participants in Experiment 2 (13 of 17; Fig. 4C). The PS-IC model gave larger (worse) BIC scores than the PE-HL model in most participants (23 of 24 in Exp. 1; 15 of 17 in Exp. 2; Fig. 4C).

As seen above, in terms of BICs, while the PE-HL model gave better fit than the pWSLS models and the PS-IC model, the PS-HL model outperformed the PE-HL model. Nevertheless, because the PS models, as well as the pWSLS models, assume that choice depends solely on the outcome of the previous trial, these models were expected to be unable to explain the observed considerable dependence of the actual choices on the experienced demand at two trials before (Fig. 2). To confirm this, we performed simulations of task execution by the PS-HL and pWSLS-HL models, as well as the PE-HL model, with the best-fit parameters for each individual demand-avoiding participant and actual demand sequences (high or low for the left and right arrow-cues) used for the participant in the experiments. We performed 100 simulation runs of task execution (180 trials) for each demand-avoiding

participant in each experiment, and examined the proportions that the *k*-th choice was the same as the (*k* − 2)-th choice when the (*k* − 2)-th demand was high or low, sorted by the (*k* − 1)-th choice and demand, in the pooled simulated choices corresponding to each participant (178 × 100 = 17,800 simulated choices; see Materials and Methods). As expected, dependence of the current choice on the demand at two trials before hardly appeared in the cases of the PS-HL and pWSLS-HL models, in contrast to the case of the PE-HL model (Fig. 8). In this way, although the PS-HL model gave good fit in terms of BIC scores, this model could not adequately explain the considerable dependence on multiple trials back observed in the actual choices, which could potentially be captured by the PE-HL model. Therefore, in the following we present model-based fMRI analyses by using the results of fitting by the PE-HL model.

fMRI analyses

We searched the whole brain for regions where changes in hemodynamic response for the presentation of the left or right arrow-cue were positively or negatively correlated with the *ExpectedCost* for the chosen option (*ExpectedCost_{Chosen}*). We also searched for regions correlated with the *CPE*. Regarding the time of *CPE* generation, there were multiple possibilities. To see this, we need to return to

Table 3. Neural correlates of the expected cost of mental effort for the chosen option revealed by using GLM2

Area	Left, right, bilateral	Cluster		Peak (T)	MNI coordinates (mm)		
		Size	<i>p</i> (FWE)		<i>x</i>	<i>y</i>	<i>z</i>
A, Experiment 1							
Positive correlation							
MFG/inferior frontal gyrus/anterior insula	Right	514	<0.001	5.6	30	50	4
aMFG	Left	103	0.033	5.5	−30	47	7
dmFC/dACC	Bilateral	196	0.002	5.7	0	14	49
Inferior parietal lobule	Right	144	0.009	5.0	45	−52	49
Negative correlation							
Rostromedial PFC	Bilateral	233	0.001	6.0	−9	53	16
vmPFC/striatum	Bilateral	392	<0.001	7.5	3	41	−17
Primary motor cortex/primary somatosensory cortex	Right	125	0.017	5.4	42	−22	64
Temporal lobe	Right	190	0.003	5.3	69	−34	13
Parietal operculum/temporal lobe	Left	1217	<0.001	6.8	−42	−37	25
Precuneus	Bilateral	218	0.001	5.4	−3	−58	16
B, Experiment 2							
Positive correlation							
Anterior insula/aMFG/orbitofrontal cortex	Left	266	<0.001	7.6	−27	23	4
dmFC/dACC/superior frontal gyrus	Bilateral	489	<0.001	7.3	6	20	49
Anterior insula/aMFG/orbitofrontal cortex	Right	663	<0.001	12.9	33	20	4
Inferior parietal lobule/precuneus	Right	262	<0.001	5.4	45	−49	55
Cerebellum	Left	141	0.007	8.0	−18	−61	−41
Negative correlation							
vmPFC ^a	Left	12	0.818	4.5	−3	35	−20
Striatum ^a	Bilateral	48	0.201	4.9	0	5	−8
C, Conjunction analysis							
Positive correlation (strict mask ^b)							
aMFG	Right	223	0.002	6.0	33	47	4
dmFC/dACC	Bilateral	136	0.018	5.8	3	20	46
Inferior parietal lobule ^c	Right	91	0.064	4.9	45	−49	49
Negative correlation (relaxed mask ^d)							
vmPFC	Bilateral	131	0.020	4.8	−6	35	−17

Configurations are the same as those in Table 2, except that GLM2 was used and C corresponds to Figure 9B.

^aVoxel-level uncorrected $p < 0.001$ and voxel size ≥ 5 .

^bA binary mask consisting of common voxels at cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$.

^cCluster-level uncorrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$; with a binary mask consisting of common voxels at cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$.

^dA binary mask consisting of common voxels at voxel-level uncorrected $p < 0.01$.

the result of model fitting that the PE-HL model assuming binary costs gave the best BIC among the PE models. As described before, this result indicates that mental-effort cost in our experiments was experienced and/or registered as (nearly) binary variables corresponding to the binary demand levels of the problems, rather than variables reflecting the solve time or mistakes. Looking more closely, this would involve (at least) two possibilities. The first possibility is that the demand level itself was registered as “actual cost” and used to update the expected cost, potentially even before the cost was experienced (i.e., before the problem was solved). The second possibility is that the experienced cost at the time of answer was in fact more closely approximated by binary variables than by the solve time in our experiments. Although response time is generally thought to relate to subjective difficulty, it is possible that the within-participant variances of the solve time for the problems of the same demand types in our experiments were due in a large part to factors that did not linearly relate to the experienced cost. For example, participants might sometimes solve a problem sluggishly so that solve time was large but experienced cost was not large: even though they were asked to solve as fast and accurately as possible, how well participants complied with this instruction was somewhat elusive given that there was no time limitation and no feedback/penalty, and also participants’ mood could fluctuate during the experiment so as to modulate the solve time and experienced cost in potentially different or nonlinearly related ways.

If the first possibility mentioned above holds, *CPE* may be generated at the time of problem presentation (or soon after it) because the binary demand level of the problem could be recognized almost instantaneously, although *CPE* could also be represented at the time of answer if the neural process for updating the expected cost could operate only after the process for problem-solving was ended. On the other hand, if the second possibility mentioned above holds, *CPE* would be generated at the time of answer. There was yet another possibility that *CPE* was generated at a time between problem presentation and answer. Therefore, by using separate GLMs (Fig. 3, GLM1–GLM3), we examined three possibilities for when *CPE* was generated/represented: at the time of problem presentation, at the midpoint of problem presentation and answer, and at the time of answer. Each of these GLMs was adjusted for the actual demand level of, and solve time for, the problem and also the response time for choosing an arrow (RT_{choice}), which could reflect decision difficulty, a potential confounder (cf., Heekeren et al., 2004; Shenhav et al., 2014; Shenhav et al., 2016b) although RT_{choice} was hardly correlated with $ExpectedCost_{Chosen}$ [Exp. 1 (mean \pm SEM): $r = 0.07 \pm 0.03$; Exp. 2: $r = 0.02 \pm 0.02$], the relative-cost ($ExpectedCost_{Chosen} - ExpectedCost_{unChosen}$; Exp. 1: $r = 0.08 \pm 0.03$; Exp. 2, $r = 0.04 \pm 0.02$), or the absolute-difference ($|ExpectedCost_{Chosen} - ExpectedCost_{unChosen}|$; Exp. 1: $r = -0.03 \pm 0.03$; Exp. 2: $r = 0.01 \pm 0.02$) in our experiments. In any of these GLMs, VIF of the regressors for $ExpectedCost_{Chosen}$

Table 4. Neural correlates of the expected cost of mental effort for the chosen option revealed by using GLM3

Area	Left, right, bilateral	Cluster		Peak (T)	MNI coordinates (mm)		
		Size	p (FWE)		x	y	z
A, Experiment 1							
Positive correlation							
MFG	Right	307	<0.001	5.4	33	44	1
dmFC/ACC	Bilateral	128	0.016	5.2	0	14	49
Negative correlation							
Rostromedial PFC	Bilateral	200	0.002	6.0	−9	53	16
vmPFC/striatum	Bilateral	322	<0.001	6.6	−9	17	−11
Primary motor cortex/primary somatosensory cortex	Right	112	0.026	5.7	42	−22	64
Temporal lobe/inferior occipital gyrus	Right	341	<0.001	5.2	69	−37	13
Precuneus	Left	178	0.004	4.9	−3	−61	13
middle occipital gyrus/temporal lobe/parietal operculum	Left	1335	0.000	6.6	−48	−79	28
B, Experiment 2							
Positive correlation							
aMFG/orbitofrontal cortex	Right	342	<0.001	7.4	24	47	−11
aMFG	Left	101	0.024	7.3	−33	44	4
Anterior insula	Left	85	0.044	5.5	−27	23	4
Anterior insula	Right	136	0.007	10.8	33	20	4
Superior frontal gyrus/dmFC/dACC	Bilateral	408	<0.001	8.1	24	8	67
Inferior parietal lobule	Right	177	0.002	5.4	48	−49	55
Cerebellum	Left	129	0.009	7.0	−21	−61	−41
Negative correlation							
vmPFC ^a	Left	5	0.955	4.2	−3	56	−5
vmPFC ^a	Left	13	0.795	4.5	−3	35	−20
Striatum ^a	Bilateral	35	0.340	5.0	0	5	−8
Middle occipital gyrus ^d	Left	6	0.940	5.0	−30	−88	37
C, Conjunction analysis							
Positive correlation (strict mask ^b)							
aMFG	Right	163	0.009	5.3	42	41	22
dmFC/dACC ^c	Bilateral	74	0.109	5.5	3	20	46
Negative correlation (relaxed mask ^d)							
vmPFC	Bilateral	116	0.031	4.7	−6	35	−17

Configurations are the same as those in Table 2, except that GLM3 was used and C corresponds to Figure 9C.

^aVoxel-level uncorrected $p < 0.001$ and voxel size ≥ 5 .

^bA binary mask consisting of common voxels at cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$.

^cCluster-level uncorrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$; with a binary mask consisting of common voxels at cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$.

^dA binary mask consisting of common voxels at voxel-level uncorrected $p < 0.01$.

and CPE was on average, across sessions and participants, < 5 , and thus we judged that collinearity was at a tolerable level.

Tables 2A, B, 3A, B, and 4A, B show the correlates of $ExpectedCost_{Chosen}$ from the individual experiments in the analyses using GLM1, GLM2, and GLM3, respectively. For positive correlates in both experiments and negative correlates in Experiment 1, clusters detected at a threshold of cluster-level FWE corrected $p < 0.05$, voxel-level uncorrected $p < 0.001$ were reported. For negative correlates in Experiment 2, no clusters were detected with the same threshold, and clusters detected at a relaxed threshold of voxel-level uncorrected $p < 0.001$ and voxel-size ≥ 5 were reported. To identify regions commonly correlated with $ExpectedCost_{Chosen}$ in both experiments, we performed conjunction analyses. For positive correlations, we applied a binary mask consisting of common voxels between the results of the individual analyses for Experiments 1 and 2 with the threshold of cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$ (hereafter we refer to the mask with this threshold as the strict mask). For negative correlations, we applied a binary mask consisting of common voxels between the results in Experiments 1 and 2 at voxel-level uncorrected $p < 0.01$ (hereafter we refer to the mask with this threshold as the relaxed mask). Figure 9 and Tables 2C, 3C, and 4C show the results of the masked conjunction analyses using the three GLMs, which are reported with the threshold of cluster-level uncorrected $p < 0.05$, voxel-level uncorrected $p < 0.001$. Regarding positive correlations,

conjunction analyses with the strict mask using any of the three GLMs revealed clusters in dorsomedial frontal cortex (dmFC)/dorsal anterior cingulate cortex (dACC) and the right anterior middle frontal gyrus (aMFG), while other cluster(s) were also detected in GLM1 and GLM2. As for negative correlations, conjunction analyses with the relaxed mask using any of the three GLMs revealed the vmPFC, although it should be noted that the relaxed mask was literally relaxed (note that it differed from the threshold for negative correlates in Exp. 2), and therefore we regarded this result for negative correlation as a trend.

It was conceivable that the PE model's variables other than $ExpectedCost_{Chosen}$ in particular, *relative-cost* ($ExpectedCost_{Chosen} - ExpectedCost_{unchosen}$) and/or *absolute-difference* ($|ExpectedCost_{Chosen} - ExpectedCost_{unchosen}|$) were also represented in the brain at the time of arrow-cue presentation commonly for both tasks. In fact, however, there were strong positive correlations between $ExpectedCost_{Chosen}$ and *relative-cost* [$r = 0.89 \pm 0.01$ (Exp. 1) and 0.90 ± 0.02 (Exp. 2)] and negative correlations between $ExpectedCost_{Chosen}$ and *absolute-difference* [$r = -0.74 \pm 0.05$ (Exp. 1) and -0.77 ± 0.05 (Exp. 2)]. Presumably because of these, when regressor for *relative-cost* or *absolute-difference* was added to GLM1, VIF of the regressor for *relative-cost* or *absolute-difference* was > 10 , precluding valid analysis. We also considered GLM in which regressor for $ExpectedCost_{unchosen}$ was added to GLM1, but VIF of the regressor for $ExpectedCost_{unchosen}$ was > 10 ,

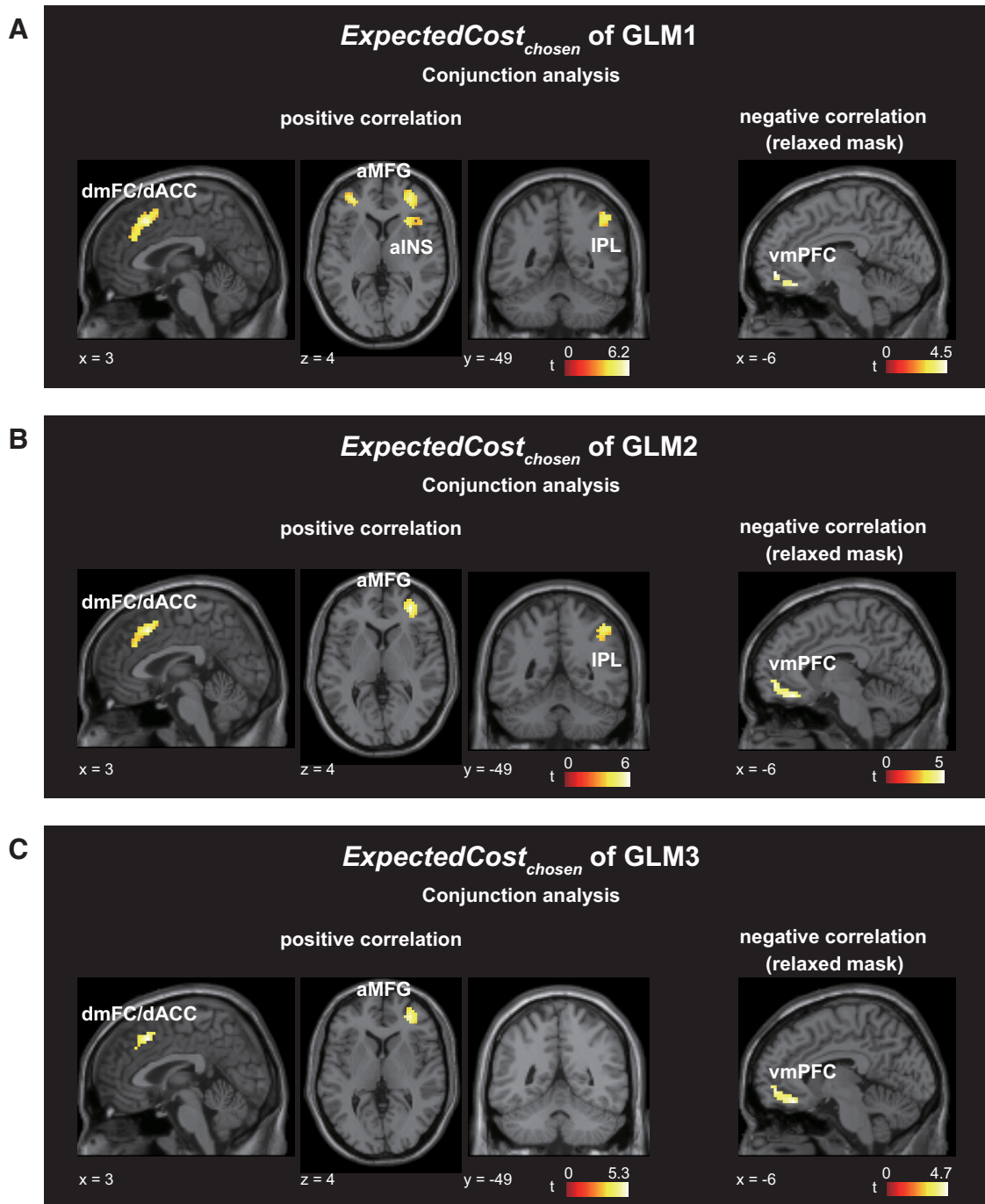


Figure 9. Neural correlates of the expected cost of mental effort for the chosen option ($ExpectedCost_{chosen}$) at the time of arrow-cue presentation common for both experiments. The results obtained by using GLM1 (**A**), GLM2 (**B**), and GLM3 (**C**). In each of these, the left three panels show the result of conjunction analysis with a binary mask consisting of common voxels between the positive correlations in Experiments 1 and 2 at cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$. In all the three cases, clusters in the right aMFG and dmFC/dACC were detected with the threshold of cluster-level uncorrected $p < 0.05$, voxel-level uncorrected $p < 0.001$, while additional cluster(s) were also detected in GLM1 (**A**) and GLM2 (**B**). The right panel shows the result of conjunction analysis with a binary mask consisting of common voxels between the negative correlations in Experiments 1 and 2 at the relaxed threshold, voxel-level uncorrected $p < 0.01$, revealing a cluster in the vmPFC in all the three cases.

precluding valid analysis. Therefore, it remained to be clarified whether *relative-cost*, *absolute-difference*, or *ExpectedCost_{unchosen}* was represented in addition to or instead of *ExpectedCost_{chosen}*.

The detection of *ExpectedCost_{chosen}*-correlated clusters in the dmFC/dACC and aMFG raises a further possibility. The dACC or nearby region has been suggested to encode/calculate the value of exploring alternative/nondefault options (Kolling et al., 2016b)

or the value of cognitive control (Shenhav et al., 2016a; see also Ebitz and Hayden, 2016 for the debate and Kolling et al., 2016a,b for presumable difference in the precise locations). While possible relation of our results to the latter proposal will be discussed (see Discussion), the former proposal raises a possibility that the dmFC/dACC activity correlated with *ExpectedCost_{chosen}* could possibly reflect an override of participants' default choice. Moreover, activity related to exploratory choices has also been re-

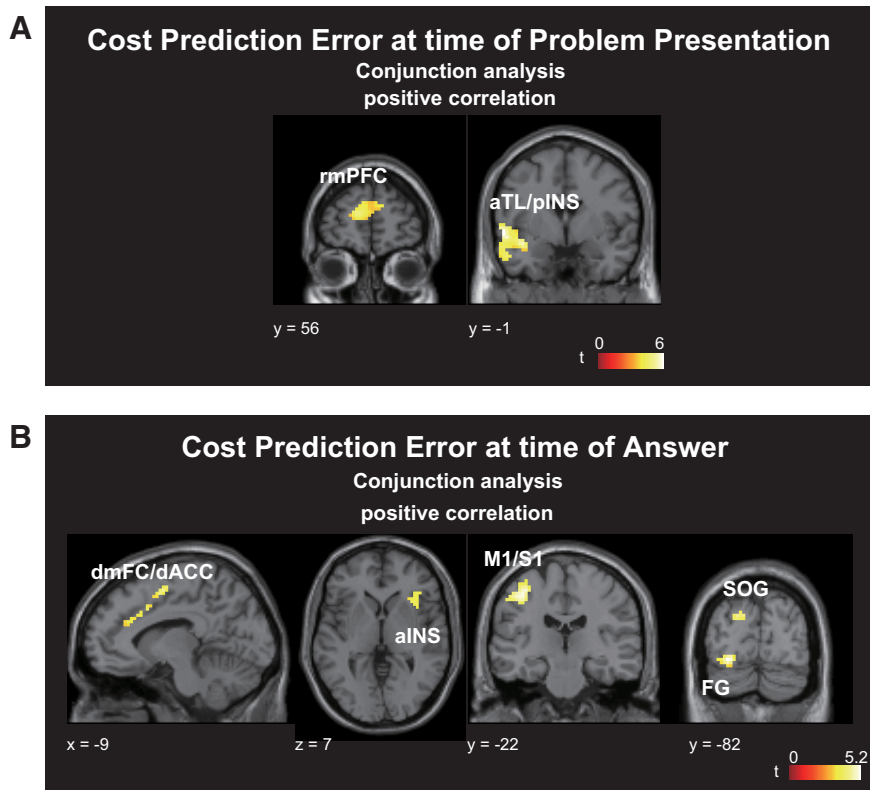


Figure 10. Neural correlates of the *CPE* at the times of problem presentation and answer. Results of conjunction analyses with a binary mask consisting of common voxels between the positive correlations of the *CPE* at the time of problem presentation (**A**: GLM1 was used) or answer (**B**: GLM3 was used) in Experiments 1 and 2 at cluster-level FWE corrected $p < 0.05$ and voxel-level uncorrected $p < 0.001$. **A**, Two clusters were detected in the rostromedial prefrontal cortex (rmPFC) and anterior temporal lobe/posterior insula (aTL/pINS). **B**, Five clusters were detected in the right anterior insula (aINS), bilateral dmFC/dACC, left primary motor cortex (M1)/primary somatosensory cortex (S1), left superior occipital gyrus (SOG), and left fusiform gyrus (FG).

ported in frontopolar regions that appear to be close to or overlap with our aMFG cluster (Daw et al., 2006). In our experiments, the rate of choosing the same option as in the previous trial was high as described before, and so participants could possibly regard it as a default choice and choosing the opposite option as the override of the default choice. Moreover, we found positive correlation between $ExpectedCost_{Chosen}$ and opposite choices (Exp. 1: $r = 0.55 \pm 0.04$; Exp. 2: 0.49 ± 0.03). We thus conducted analyses using another GLM, which differed from GLM1 in that the regressor at time of arrow cue was not set at the initial trial and was additionally parametrically modulated by opposite-versus-same choices (same choice as in the previous trial, 0; opposite choice, 1); VIF for $ExpectedCost_{Chosen}$ and opposite-versus-same choices was on average, across sessions and participants, < 5 . As a result, however, conjunction analysis with the strict mask revealed $ExpectedCost_{Chosen}$ -correlated clusters in the dmFC/dACC and the right aMFG that were similar to, albeit weaker than, those obtained in GLM1, whereas no cluster was detected as correlates of opposite-versus-same choices even with the relaxed mask (data not shown). Based on this result, it seems unlikely that our results for $ExpectedCost_{Chosen}$ are explained by an override of default choice of the same options. Another possibility related to nondefault/exploratory choices, which depends on the existence of $ExpectedCost$, is that choosing an option with higher $ExpectedCost$ could be regarded as nondefault/exploratory (cf. Daw et al., 2006). The rate of higher- $ExpectedCost$ choices (in Trials 2–180) was $20.2 \pm 2.2\%$ in Experiment 1 and $19.9 \pm 2.0\%$ in Experiment

2, and such choices were correlated with $ExpectedCost_{Chosen}$ (Exp. 1: $r = 0.59 \pm 0.02$; Exp. 2: 0.64 ± 0.02), and thus the override of the default choice in this sense could possibly contribute to the signal for $ExpectedCost_{Chosen}$.

Last, we report the correlates of *CPE*. Conjunction analyses with the strict mask revealed two clusters for positive correlations with *CPE* at time of problem presentation in GLM1 (Fig. 10A; Table 5C). For positive correlations with *CPE* at the midpoint of problem presentation and answer in GLM2, no cluster was detected by conjunction analysis with the strict mask, while analysis with the relaxed mask revealed a cluster (Table 6C). Meanwhile, five clusters for positive correlations with *CPE* were detected by conjunction analysis with the strict mask at time of answer in GLM3 (Fig. 10B; Table 7C). For negative correlations with *CPE* at any of the three times, conjunction analyses with the relaxed mask did not reveal any cluster. Comparing the revealed common positive correlates of *CPE* at time of problem presentation in GLM1 (Fig. 10A) or at time of answer in GLM3 (Fig. 10B) with the common positive correlates of $ExpectedCost_{Chosen}$ at time of arrow-cue presentation (Fig. 9), there appear to be possible overlap between the regions for *CPE* at time of answer and those for $ExpectedCost_{Choice}$ at time of arrow cue. We examined this possibility by using the correlates of $ExpectedCost_{Chosen}$ at the time of the arrow cue and *CPE* at time of answer obtained from the same GLM3, and found overlapping and neighboring regions in the right dmFC (overlap, seven voxels; Fig. 11).

The existences of common *CPE* correlates at time of both problem presentation and answer imply coexistence of the different possibilities mentioned before. Specifically, the *CPE* correlates at problem presentation imply existence of a system that registers the demand level itself as “actual cost” and calculates *CPE* before problem-solving. On the other hand, the *CPE* correlates at answer imply that *CPE*-dependent update of expected cost occurred at this time even though the demand-level itself was registered as “actual cost” and/or the actually experienced cost was in fact more closely approximated by binary variables than by the solve time. The result that there was overlap/adjacency between the common correlates of *CPE* at answer and the common correlates of $ExpectedCost_{Chosen}$ at arrow cue could then imply that *CPE* represented at answer, whether calculated from the demand level itself and/or the actually experienced cost, was used for update of expected cost. However, while these implications are intriguing, we note that this present work is limited by the difficulty in specifying the time of *CPE* generation, as well as the fact that in any case *CPE* generation was likely temporally overlapped with problem-solving. These points need to be addressed by using a different task design to isolate the time of *CPE* generation.

Table 5. Neural correlates of the cost prediction error at the time of problem presentation revealed by using GLM1

Area	Left, right, bilateral	cluster		Peak (<i>T</i>)	MNI coordinates (mm)		
		Size	<i>p</i> (FWE)		<i>x</i>	<i>y</i>	<i>z</i>
A, Experiment 1							
Positive correlation							
Rostromedial PFC	Bilateral	172	0.004	5.4	−9	56	13
Temporal pole	Right	106	0.027	5.1	33	17	−26
Temporal lobe/posterior insula/vmPFC/striatum	Left	1898	<0.001	8.2	−54	−4	−2
Hippocampus/parahippocampal gyrus	Right	94	0.040	5.9	27	−19	−23
Temporal lobe/posterior insula	Right	714	<0.001	7.0	66	−22	13
Cuneus	Bilateral	164	0.005	7.5	−3	−85	28
B, Experiment 2							
Positive correlation							
Rostromedial PFC/ventral anterior cingulate cortex/striatum	Bilateral	676	<0.001	8.9	0	17	−5
Temporal lobe/insula	Right	999	<0.001	7.2	−39	5	−17
Temporoparietal junction	Right	89	0.023	5.6	60	−58	16
Negative correlation							
MFG	Right	84	0.028	6.4	45	35	19
Precentral gyrus/posterior MFG	Left	101	0.014	5.4	−51	5	40
Precentral gyrus/posterior MFG	Right	150	0.002	4.9	27	−1	46
Inferior parietal sulcus	Right	423	<0.001	6.8	36	−46	46
Inferior parietal sulcus/occipital lobe	Left	667	<0.001	6.9	−27	−73	31
C, Conjunction analysis							
Positive correlation (strict mask ^a)							
Rostromedial PFC	Bilateral	108	0.031	4.8	−12	56	16
Anterior temporal lobe/posterior insula	Left	520	<0.001	5.9	−57	−1	−2

The brain area, laterality of hemisphere, cluster size, cluster *p* value, peak *T* value, and peak coordinate of the fMRI results for the neural correlates of the cost prediction error at the time of problem presentation obtained by using GLM1. Methods for analyses and statistical thresholds are described in Materials and Methods and Results.

^aA binary mask consisting of common voxels at cluster-level FWE corrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001.

Table 6. Neural correlates of the cost prediction error at the midpoint between problem presentation and answer revealed by using GLM2

Area	Left, right, bilateral	Cluster		Peak (<i>T</i>)	MNI coordinates (mm)		
		Size	<i>p</i> (FWE)		<i>x</i>	<i>y</i>	<i>z</i>
A, Experiment 1							
Positive correlation							
Parietal-temporal-occipital association area/superior temporal gyrus	Left	692	<0.001	5.5	−48	−76	28
Superior temporal gyrus/operculum	Right	328	<0.001	5.1	57	−16	7
B, Experiment 2							
Positive correlation							
Anterior temporal lobe/insula	Right	285	<0.001	8.5	42	−16	−17
Temporal lobe/subcallosal area/vmPFC	Bilateral	757	<0.001	7.7	−48	−19	−17
C, Conjunction analysis							
Positive correlation (relaxed mask ^a)							
Temporoparietal junction ^b	Left	82	0.079	4.2	−57	−52	7

The brain area, laterality of hemisphere, cluster size, cluster *p* value, peak *T* value, and peak coordinate of the fMRI results for the neural correlates of the CPE at the midpoint between problem presentation and answer obtained by using GLM2. Methods for analyses and statistical thresholds are described in Materials and Methods and Results.

^aA binary mask consisting of common voxels at voxel-level uncorrected *p* < 0.01.

^bCluster-level uncorrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001; with a binary mask consisting of common voxels at voxel-level uncorrected *p* < 0.01.

Discussion

Most participants learned to avoid higher cognitive demand in the changing environments, and their choices depended on the demand experienced during the preceding multiple trials; this could potentially be captured by the PE-based model assuming that the experienced demand level constituted actual cost. At the neural level, *ExpectedCost*_{Chosen} was positively correlated with the activity in the dmFC/dACC and aMFG, and with the relaxed mask, negatively correlated with vmPFC activity, commonly across the demand types. Further, we identified common positive correlates of CPE at time of problem presentation and answering the problem, the latter of which partially overlapped with or was in proximity with the positive correlates of *ExpectedCost*_{Chosen} at time of arrow cue in the dmFC.

Relation to previous studies on mental-effort avoidance

Previous studies have demonstrated that humans avoid cognitive demand/mental effort in carefully controlled experimental settings (Botvinick, 2007; Kool et al., 2010; McGuire and Botvinick, 2010; Risko et al., 2014; Schoupe et al., 2014). Our results have extended these findings by showing that humans adaptively learn to avoid higher cognitive demands under uncertain and nonstationary environments, with the choices depending on the demand on multiple preceding trials.

Previous studies have also explored neural substrates related to cognitive demand-avoidance by using fMRI (Botvinick et al., 2009; McGuire and Botvinick, 2010; Schoupe et al., 2014; Massar et al., 2015; Chong et al., 2017). One study (McGuire and Botvinick, 2010) reported that post-experience self-reports of the

Table 7. Neural correlates of the cost prediction error at the time of answer revealed by using GLM3

Area	Left, right, bilateral	Cluster		Peak (<i>T</i>)	MNI coordinates (mm)		
		Size	<i>p</i> (FWE)		<i>x</i>	<i>y</i>	<i>z</i>
A, Experiment 1							
Positive correlation							
Insula	Right	190	0.005	5.6	36	26	4
Anterior insula	Left	157	0.010	6.4	−33	23	−5
dmFC/dACC	Bilateral	529	<0.001	5.8	9	−1	55
Posterior insula/primary auditory cortex	Right	107	0.040	5.2	39	−22	4
Occipital lobe/parietal lobe	Bilateral	6393	<0.001	8.6	−18	−88	−14
B, Experiment 2							
Positive correlation							
dmFC/dACC/middle cingulate cortex	Bilateral	494	<0.001	8.0	−9	20	37
Anterior insula	Right	112	0.010	5.5	30	20	7
Primary motor cortex/primary somatosensory cortex/ supramarginal gyrus	Left	161	0.002	6.7	−36	−22	58
Fusiform gyrus	Right	165	0.001	6.4	30	−73	−14
Fusiform gyrus	Left	214	<0.001	5.5	−24	−73	−11
Occipital lobe	Left	94	0.020	5.9	−15	−91	25
Negative correlation							
Striatum	Right	74	0.047	6.2	21	8	19
C, Conjunction analysis							
Positive correlation (strict mask ^a)							
Anterior insula ^b	Right	65	0.144	4.4	33	20	7
dmFC/dACC	Bilateral	179	0.006	4.8	−9	−1	55
Primary motor cortex/primary somatosensory cortex	Left	148	0.013	5.3	−39	−22	52
Superior occipital gyrus ^b	Left	77	0.099	4.6	−18	−82	31
Fusiform gyrus	Left	110	0.037	5.5	−27	−82	−11

The brain area, laterality of hemisphere, cluster size, cluster *p* value, peak *T* value, and peak coordinate of the fMRI results for the neural correlates of the CPE at the time of answer obtained by using GLM3. Methods for analyses and statistical thresholds are described in Materials and Methods and Results.

^aA binary mask consisting of common voxels at cluster-level FWE corrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001.

^bCluster-level uncorrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001; with a binary mask consisting of common voxels at cluster-level FWE corrected *p* < 0.05 and voxel-level uncorrected *p* < 0.001.

Overlap/Adjacency

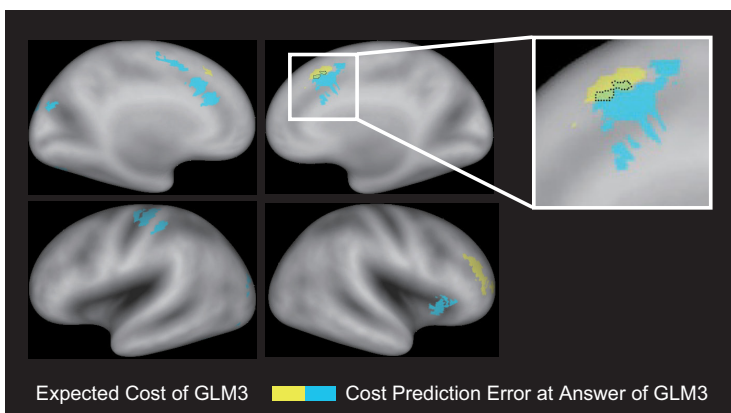


Figure 11. Overlap and adjacency between the correlates of the expected cost for the chosen option at the time of arrow-cue presentation and those of the CPE at the time of answer. Common positive correlates of the CPE at the time of answer obtained by using GLM3 (indicated by light blue color) and common positive correlates of the expected cost for the chosen option (*Expected-Cost_{choice}*) at the time of arrow-cue presentation obtained by using the same GLM3 (indicated by yellow color). The right panel shows an enlarged view. The overlapped region is enclosed by the black dashed line, which was drawn manually by the authors.

desire to avoid high demand were related with activity in regions in the lateral PFC, but not with a dmFC/dACC cluster. The apparent inconsistency between their results and ours can be explained given the differences between the studies: here we examined activations during demand-based selection, whereas in their work, no demand selection was made in the scanner and the avoidance ratings were made after experience of demands.

Another previous study (Schouppe et al., 2014) examined in-scanner choice of options with high or low expected cognitive demand in two conditions, where participants were instructed to

make either a voluntary but random choice, or a forced choice. Then, participants chose high-demand and low-demand options with almost equal rates in voluntary trials presumably because of the instruction to choose randomly, and thus brain activation during natural demand-avoidance was not examined.

Brain activity during effort-related choices has been also examined in the cases where reward values are discounted by mental (and physical) effort (Botvinick et al., 2009; Massar et al., 2015; Chong et al., 2017). A recent study investigated choices between two cues that explicitly represented a variable high-effort–high-reward option and a fixed low-effort–low-reward (baseline) option (Chong et al., 2017). Effort exertion was experienced before and after scans, but not during scans. This study showed that the activity in regions including the dmPFC/dACC was negatively correlated with the subjective

value difference between the chosen option and baseline, commonly across mental and physical effort tasks. The peak of the dmPFC/dACC cluster was within the region correlated with *ExpectedCost_{Chosen}* in our study. This seems reasonable given that the subjective value difference in their study could be negatively related to the *ExpectedCost_{Chosen}*. On the other hand, activity in the aMFG and vmPFC were detected in our study, but not in their study. This might reflect differences in our study from theirs, including the absence of reward manipulations and/or experi-

ence of in-scanner effort exertion. There exists much evidence that the vmPFC has common representations of values for use in decision making (Levy and Glimcher, 2012), negatively integrating the cost of monetary loss (Basten et al., 2010), delay (Prévost et al., 2010), or choice difficulty (Shenhav et al., 2016b). Therefore, the vmPFC might specifically serve for experience-based learned choices of values, as argued in the above-mentioned study (Chong et al., 2017). Meanwhile, the aMFG region might serve for mental-effort avoidance when experience-based learning occurs and/or when reward effects are absent. As for the latter, existence of such a specialized system for no reward-effect conditions seems in line with the suggestion that systems for appetitive and aversive learning can be separated to some extent (Seymour et al., 2004, 2005; Yacubian et al., 2006; Basten et al., 2010; Li et al., 2011; Roy et al., 2014; Scholl et al., 2017).

Implications for the mechanisms

We hypothesized the existence of neural representations of $ExpectedCost_{Chosen}$ which are updated according to CPE and used for decision making to avoid higher demand. Possible substrates of this could be captured in our finding that the cue-time activity of the dmFC/dACC and aMFG clusters and the answer-time activity of the clusters including a dmFC/dACC cluster, were correlated with $ExpectedCost_{Chosen}$ and CPE , respectively, both commonly across tasks, and these two correlates partially overlapped or were adjacent in the right dmFC. In reference to reinforcement learning (RL) theory (Daw et al., 2005), this mechanism could be called model-free RL based on the “cached cost” of options. Whereas RL of reward values (McClure et al., 2003; O’Doherty et al., 2003; Daw et al., 2006), pain (Seymour et al., 2004; Roy et al., 2014; Zhang et al., 2016), physical effort (Skvortsova et al., 2014), or sustained effort (selecting circles on the screen) concurrently with reward-learning (Scholl et al., 2015) has been well investigated, our current study presents for the first time an empirical indication that humans might also learn to avoid high cognitive demands, even without reward-learning, in an RL-like fashion, although a different decision strategy may also be used. An intriguing hypothesis (Kurzman et al., 2013) indicates that humans might avoid mental effort so as to minimize the opportunity cost of focusing on a particular task. Estimating opportunity cost by forward-reading may not always be possible, and thus caching mechanisms may be needed, possibly in line with the indication from our results. On the other hand, the detection of common CPE correlates at problem presentation in the rostromedial prefrontal cortex and the anterior temporal lobe/posterior insula implies that another, more explicit knowledge-based system might simultaneously operate. Specifically, error signal calculated from perceived demand level could possibly be used for learning of the probabilistic associations between the cues and demand levels.

Regarding the mechanisms of decision making, a circuit that selects lower expected-cost options in a softmax manner might exist. Alternatively, information about expected cost can be used to calculate the expected value, through a sign reversal by inhibitory connections, so that higher expected-value options are chosen in a softmax manner. The conjunction analysis with the relaxed mask suggested that $ExpectedCost_{Choice}$ was negatively correlated with vmPFC activity (Fig. 9; Tables 2C, 3C, 4C), potentially supporting the latter possibility. This possibility is also consistent with the suggestion that the vmPFC has common representations of values, and also that the vmPFC exhibits features of recurrent neural dynamics that can implement softmax selections (Hunt et al., 2012; Jocham et al., 2012). However, given a recent suggestion that reward-based choice emerges from com-

putations in distributed networks (Hunt and Hayden, 2017), choice might rather be made through interactions between the detected regions.

Activity in the dmFC, dACC, and the surrounding areas has been suggested to be related to the preparation of tasks or the allocation of control (Botvinick et al., 2001; Sakai, 2008). Extending the influential conflict monitoring theory (Botvinick et al., 2001) rooted in experimental findings (Pardo et al., 1990), Shenhav and colleagues recently proposed the Expected Value of Control (EVC) theory (Shenhav et al., 2013), which explicitly theorizes the relationship between mental effort, valuation, and the dACC (Shenhav et al., 2013, 2016a, 2017). EVC is defined as the expected total reward, given a control signal (identity and intensity), discounted by the cost associated with the signal, and the EVC theory proposes that the dACC receives motivation/valuation inputs from regions including vmPFC, calculates EVC, and allocates control through interactions with lateral PFC and other brain regions. Our finding that dmFC/dACC activity was correlated with $ExpectedCost_{Chosen}$ may be in line with this theory, although learning and selection based on cached costs have not been explicitly discussed. Moreover, from a perspective of the EVC theory, the $ExpectedCost_{Chosen}$ -correlated region in the vmPFC might send motivation/valuation inputs to the dmFC/dACC, while the aMFG region would communicate control allocation-related information with the dmFC/dACC.

There are further possibilities related to the EVC theory. In our tasks, when $ExpectedCost_{Chosen}$ is high (i.e., when a chosen problem is expected to be difficult), recruiting a high level of control is likely to be necessary to solve the forthcoming problem. In contrast, when $ExpectedCost_{Choice}$ is low, recruiting a low level of control is likely to be sufficient. $ExpectedCost_{Chosen}$ is thus considered to be positively correlated with the optimal intensity of control signal or a relative EVC (i.e., EVC of high control minus EVC of low control). Therefore, the correlates of $ExpectedCost_{Chosen}$ in our results potentially also reflected the relative EVC, expected cost of optimal control signal, or the optimal control signal itself. This last possibility indicates that the dmFC/dACC and aMFG activity could represent preparatory control instead of or in addition to the expected cost: an important issue that needs to be clarified in future work.

Note Added in Proof: The author contributions were accidentally incorrectly listed in the Early Release version published February 5, 2018. The author contributions have now been corrected.

References

- Basten U, Biele G, Heekeren HR, Fiebach CJ (2010) How the brain integrates costs and benefits during decision making. *Proc Natl Acad Sci U S A* 107:21767–21772. [CrossRef Medline](#)
- Blain B, Hollard G, Pessiglione M (2016) Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. *Proc Natl Acad Sci U S A* 113:6967–6972. [CrossRef Medline](#)
- Botvinick MM (2007) Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* 7:356–366. [CrossRef Medline](#)
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. *Psychol Rev* 108:624–652. [CrossRef Medline](#)
- Botvinick MM, Huffstetler S, McGuire JT (2009) Effort discounting in human nucleus accumbens. *Cogn Affect Behav Neurosci* 9:16–27. [CrossRef Medline](#)
- Chong TT, Apps M, Giehl K, Silence A, Grima LL, Husain M (2017) Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol* 15:e1002598. [CrossRef Medline](#)
- Daw ND (2011) Trial-by-trial data analysis using computational models. In: *Decision making, affect, and learning, attention and performance XXIII* (Delgado M, Phelps EA, Robbins TW, eds). Oxford: Oxford UP.

- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711. [CrossRef Medline](#)
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879. [CrossRef Medline](#)
- Duncan J (2010) The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci* 14:172–179. [CrossRef Medline](#)
- Ebitz RB, Hayden BY (2016) Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. *Nat Neurosci* 19:1278–1279. [CrossRef Medline](#)
- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862. [CrossRef Medline](#)
- Hunt LT, Hayden BY (2017) A distributed, hierarchical and recurrent framework for reward-based choice. *Nat Rev Neurosci* 18:172–182. [CrossRef Medline](#)
- Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, Behrens TE (2012) Mechanisms underlying cortical activity during value-guided choice. *Nat Neurosci* 15:470–476, S1–S3. [CrossRef Medline](#)
- Jocham G, Hunt LT, Near J, Behrens TE (2012) A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat Neurosci* 15:960–961. [CrossRef Medline](#)
- Kolling N, Behrens T, Wittmann MK, Rushworth M (2016a) Multiple signals in anterior cingulate cortex. *Curr Opin Neurobiol* 37:36–43. [CrossRef Medline](#)
- Kolling N, Wittmann MK, Behrens TE, Boorman ED, Mars RB, Rushworth MF (2016b) Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci* 19:1280–1285. [CrossRef Medline](#)
- Kool W, McGuire JT, Rosen ZB, Botvinick MM (2010) Decision making and the avoidance of cognitive demand. *J Exp Psychol Gen* 139:665–682. [CrossRef Medline](#)
- Krebs RM, Boehler CN, Roberts KC, Song AW, Woldorff MG (2012) The involvement of the dopaminergic midbrain and cortico-striatal-thalamic circuits in the integration of reward prospect and attentional task demands. *Cereb Cortex* 22:607–615. [CrossRef Medline](#)
- Kurzban R, Duckworth A, Kable JW, Myers J (2013) An opportunity cost model of subjective effort and task performance. *Behav Brain Sci* 36:661–679. [CrossRef Medline](#)
- Levy DJ, Glimcher PW (2012) The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22:1027–1038. [CrossRef Medline](#)
- Li J, Schiller D, Schoenbaum G, Phelps EA, Daw ND (2011) Differential roles of human striatum and amygdala in associative learning. *Nat Neurosci* 14:1250–1252. [CrossRef Medline](#)
- Mansouri FA, Egner T, Buckley MJ (2017) Monitoring demands for executive control: shared functions between human and nonhuman primates. *Trends Neurosci* 40:15–27. [CrossRef Medline](#)
- Massar SA, Libedinsky C, Weiyan C, Huettel SA, Chee MW (2015) Separate and overlapping brain areas encode subjective value during delay and effort discounting. *Neuroimage* 120:104–113. [CrossRef Medline](#)
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346. [CrossRef Medline](#)
- McGuire JT, Botvinick MM (2010) Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc Natl Acad Sci U S A* 107:7922–7926. [CrossRef Medline](#)
- Mumford JA, Poline JB, Poldrack RA (2015) Orthogonalization of regressors in fMRI models. *PLoS One* 10:e0126255. [CrossRef Medline](#)
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337. [CrossRef Medline](#)
- O'Doherty JP, Hampton A, Kim H (2007) Model-based fMRI and its application to reward learning and decision making. *Ann N Y Acad Sci* 1104:35–53. [CrossRef Medline](#)
- Pardo JV, Pardo PJ, Janer KW, Raichle ME (1990) The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc Natl Acad Sci U S A* 87:256–259. [CrossRef Medline](#)
- Prévost C, Pessiglione M, Météreau E, Cléry-Melin ML, Dreher JC (2010) Separate valuation subsystems for delay and effort decision costs. *J Neurosci* 30:14080–14090. [CrossRef Medline](#)
- Risko EF, Medimorec S, Chisholm J, Kingstone A (2014) Rotating with rotated text: a natural behavior approach to investigating cognitive offloading. *Cogn Sci* 38:537–564. [CrossRef Medline](#)
- Roy M, Shohamy D, Daw N, Jepma M, Wimmer GE, Wager TD (2014) Representation of aversive prediction errors in the human periaqueductal gray. *Nat Neurosci* 17:1607–1612. [CrossRef Medline](#)
- Sakai K (2008) Task set and prefrontal cortex. *Annu Rev Neurosci* 31:219–245. [CrossRef Medline](#)
- Scholl J, Kolling N, Nelissen N, Wittmann MK, Harmer CJ, Rushworth MF (2015) The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort. *J Neurosci* 35:11233–11251. [CrossRef Medline](#)
- Scholl J, Kolling N, Nelissen N, Browning M, Rushworth MF, Harmer CJ (2017) Beyond negative valence: 2-week administration of a serotonergic antidepressant enhances both reward and effort learning signals. *PLoS Biol* 15:e2000756. [CrossRef Medline](#)
- Schoupe N, Demanet J, Boehler CN, Ridderinkhof KR, Notebaert W (2014) The role of the striatum in effort-based decision-making in the absence of reward. *J Neurosci* 34:2148–2154. [CrossRef Medline](#)
- Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667. [CrossRef Medline](#)
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci* 8:1234–1240. [CrossRef Medline](#)
- Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79:217–240. [CrossRef Medline](#)
- Shenhav A, Straccia MA, Cohen JD, Botvinick MM (2014) Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat Neurosci* 17:1249–1254. [CrossRef Medline](#)
- Shenhav A, Cohen JD, Botvinick MM (2016a) Dorsal anterior cingulate cortex and the value of control. *Nat Neurosci* 19:1286–1291. [CrossRef Medline](#)
- Shenhav A, Straccia MA, Botvinick MM, Cohen JD (2016b) Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cogn Affect Behav Neurosci* 16:1127–1139. [CrossRef Medline](#)
- Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen JD, Botvinick MM (2017) Toward a rational and mechanistic account of mental effort. *Annu Rev Neurosci* 40:99–124. [CrossRef Medline](#)
- Skvortsova V, Palminteri S, Pessiglione M (2014) Learning to minimize efforts versus maximizing rewards: computational principles and neural correlates. *J Neurosci* 34:15621–15630. [CrossRef Medline](#)
- Sohn MH, Albert MV, Jung K, Carter CS, Anderson JR (2007) Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proc Natl Acad Sci U S A* 104:10330–10334. [CrossRef Medline](#)
- Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. *Science* 185:1124–1131. [CrossRef Medline](#)
- Vassena E, Silvetti M, Boehler CN, Achten E, Fias W, Verguts T (2014) Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS One* 9:e91008. [CrossRef Medline](#)
- Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C (2006) Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci* 26:9530–9537. [CrossRef Medline](#)
- Zhang S, Mano H, Ganesh G, Robbins T, Seymour B (2016) Dissociable learning processes underlie human pain conditioning. *Curr Biol* 26:52–58. [CrossRef Medline](#)