Value estimation versus effort mobilization: a general dissociation between ventromedial and dorsomedial prefrontal cortex

Nicolas Clairis, Mathias Pessiglione

Value estimation versus effort mobilization: a general dissociation between ventromedial and dorsomedial prefrontal cortex

Abbreviated title: Value vs. effort: vmPFC vs. dmPFC

Authors: Nicolas Clairis[1,2,3] and Mathias Pessiglione[1,2]

[1]Motivation, Brain and Behavior team, Paris Brain Institute (ICM), Paris, France

[2]Sorbonne Université, Inserm U1127, CNRS U7225, Paris, France

[3]Laboratory of Behavioral Genetics (LGC), Brain Mind Institute (BMI), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Corresponding author email address:

nicolas.clairis@protonmail.com; mathias.pessiglione@gmail.com

Conflict of interest statement: The authors declare no competing financial interests.

Acknowledgments: The study was funded by the Fondation pour la Recherche and an Investissements d'Avenir program (ANR-10-IBHU-0003). We thank the center for neuroimaging (CENIR) staff for help in fMRI data acquisition, particularly Stéphane Lehéricy, Romain Valabrègue and Mathieu Santin for the optimization of scanning sequences. We are also grateful to Jules Brochard for assistance in computational modeling and data analysis.

Key words: *Decision-making, learning, reward, effort, confidence, grip, Stroop, prefrontal cortex, fMRI, computational model*

Abstract

Deciding for a course of action requires both an accurate estimation of option values and a right amount of effort invested in deliberation to reach sufficient confidence in the final choice. In a previous study, we have provided evidence, across a series of judgement and choice tasks, for a dissociation between the ventromedial prefrontal cortex (vmPFC), which would represent option values, and the dorsomedial prefrontal cortex (dmPFC), which would represent the duration of deliberation. Here, we first replicate this dissociation and extend it to the case of an instrumental learning task, in which 24 human volunteers (13 women) choose between options associated with probabilistic gains and losses. According to fMRI data recorded during decision-making, vmPFC activity reflects the sum of option values generated by a reinforcement learning model, and dmPFC activity the deliberation time. To further generalize the role of the dmPFC in mobilizing effort, we then analyze fMRI data recorded in the same participants while they prepare to perform motor and cognitive tasks (squeezing a handgrip or making numerical comparisons) to maximize gains or minimize losses. In both cases, dmPFC activity is associated with the output of an effort regulation model, and not with response time. Taken together, these results strengthen a general theory of behavioral control that implicates the vmPFC in the estimation of option values and the dmPFC in the energization of relevant motor and cognitive processes.

Significance statement

The medial prefrontal cortex (mPFC) is known to represent key variables needed for choosing a course of action. We previously suggested a functional partition of this brain region: the expected values of choice options are signaled by the ventral part (vmPFC) and the effort invested in decision-making by the dorsal part (dmPFC). Here, we generalize this functional partition to various motor and cognitive tasks, using fMRI in healthy volunteers. Results show that vmPFC activity reflects the expected value of options generated by a reinforcement learning model (whether the goal is to maximize reward or avoid punishment), while dmPFC activity reflects the output of an effort regulation model (whether the task is to produce force or to compare digits).

Introduction

Standard economic decision theory assumes that choice options can be ordered on a common value scale. Functional neuroimaging studies have identified value signals in specific regions of the human brain, with the ventromedial prefrontal cortex (vmPFC) as a key node (Peters and Büchel, 2010; Bartra et al., 2013; Clithero and Rangel, 2014). Neural activity in the vmPFC reflects the values of stimuli belonging to different categories such as money, food, faces or paintings (Chib et al., 2009; Lebreton et al., 2009; Lopez-Persem et al., 2020; Tom et al., 2007), whatever the modality of presentation such as with text, image, taste or sound (Plassmann et al., 2007; Lebreton et al., 2015; Abitbol et al., 2015; Chang et al., 2021) and for different types of tasks such as rating or choice (Kable and Glimcher, 2009; Suzuki et al., 2017; Shenhav and Karmarkar, 2019; Clairis and Pessiglione, 2022). The vmPFC aggregates not only the positive features of expected rewards but also negative discounters such as risk (Levy et al., 2010; Schonberg et al., 2012; Seaman et al., 2018; Silston et al., 2021), delay (Economides et al., 2015; Jimura et al., 2013; Kable and Glimcher, 2007; Lee et al., 2021), and even physical and mental efforts (Aridan et al., 2019; Westbrook et al., 2019; Lopez-Gamundi et al., 2021; Clairis and Pessiglione, 2022). Thus, the vmPFC signal may provide a common neural currency, based on which options could be compared for making decisions (Levy and Glimcher, 2012).

The dorsomedial prefrontal cortex (dmPFC) has also been implicated in decision-making, but its precise role is more debated (Clairis and Lopez-Persem, 2023). Neural activity in this region (sometimes labeled dorsal anterior cingulate cortex, dACC) has been related to diverse variables, including negative net action value (Bartra et al., 2013; Pessiglione et al., 2018), choice uncertainty (Volz et al., 2005; Hogan et al., 2019), environment volatility (Behrens et al., 2007), exploration value (Kolling et al., 2012), model updating (Kolling et al., 2016; Fouragnan et al., 2018), etc. Thus, figuring out a single overarching function for this region remains a challenge. Some authors have noticed that the task features inducing an increase in dmPFC activity are often related to a higher demand in mental effort or cognitive control (Shenhav et al., 2013). Accordingly, dmPFC activity was shown to increase with the effort invested in the task, whether it is physical effort as when squeezing a handgrip (Kurniawan et al., 2021; Skvortsova et al., 2014), cognitive effort as when facing conflict (Pochon et al., 2008;

Shenhav et al., 2014) or simply deliberation time when making decisions (Clairis and Pessiglione, 2022; Grinband et al., 2011). The dmPFC may therefore be responsible for the mobilization of effort, defined as the investment of the physical or mental resources needed to attain a certain goal (Richter et al., 2016).

In a previous study (Clairis and Pessiglione, 2022), we dissociated the roles of the vmPFC and dmPFC in tasks involving an expression of subjective preference (rating and choice): while vmPFC activity reflected the value of options, dmPFC activity reflected the duration of deliberation. This suggests that the vmPFC signals an option value that integrates expected reward and effort, while the dmPFC signals the effort to invest in the deliberation process. Here, we first test whether this dissociation can be extended to the context of an instrumental learning task (adapted from (Pessiglione et al., 2006), in which cues are probabilistically associated with gain versus loss outcomes. The prediction was that dmPFC activity would reflect response time (RT), while vmPFC activity would represent option values estimated using a reinforcement learning model. We then test the functional interpretation that dmPFC activity reflects effort mobilization, using an incentive motivation task (adapted from (Schmidt et al., 2012), in which participants make either a physical effort (squeezing a handgrip as hard as possible) or a mental effort (doing numerical Stroop comparisons as fast as possible), to maximize gains and minimize losses. The prediction was that dmPFC activity would reflect the amount of exerted effort, estimated using an effort regulation model, irrespective of RT.

Methods

Subjects

In total, 24 volunteers (13 women), aged 25.9 ± 3.7 years (mean ± standard deviation) participated in this study, which was approved by the Pitié-Salpêtrière Hospital local ethics committee. Participants were recruited through the online RISC (Relais d'Information en Sciences de la Cognition) platform (https.//www.risc.cnrs.fr/). All participants were screened for the use of psychotropic medications and for any history of psychiatric or neurologic disorders. The inclusion criteria imposed being right-handed, fluent in French, between 20 and 39 years old, having normal or corrected-to-normal vision, not being pregnant and not wearing tattoos or metallic implants.

Participants were told that they would receive a fixed amount of 50€ for their participation, plus an additional bonus between 0 and 25€, depending on cumulative outcomes across tasks and sessions. In practice, all participants were paid the same amount in the end (75€). One participant was excluded from all analyses due to poor performance in all tasks. Another participant was excluded because of excessive movement inside the scanner (>5 mm) in all sessions. The dataset therefore includes a total of 22 participants (12 women), aged 25.6 ± 3.6 years (mean ± standard deviation). For the pupil data analysis, we had to remove 2 additional subjects because of poor-quality recordings (leaving n=20 participants).

Behavioral tasks

Subjects were given both written and oral instructions about the tasks, which were programmed using Psychtoolbox (http://psychtoolbox.org/) in Matlab 2012, see scripts at https://github.com/NicolasClairis/value_estimation_vs_effort_mobilization.

The learning task employed here (see Figure 1) was similar to that used in previous studies (Pessiglione et al., 2006; Palminteri et al., 2012). Participants were told that, in a given session, they would be confronted to 6 different visual cues (actually letters taken from the Agathodaimon alphabet). They had to find out, by trial and error, which ones they should select in order to maximize their payoff. There were 3 possible outcomes: a gain (+10€), nothing (0€), or a loss (-10€). The 6 cues of a session were grouped in 3 fixed pairs: one associated with gain (winning 10€ or 0€), one with neutral outcomes (0€ always) and one with loss (losing 10€ or 0€). Neutral pairs were useless for learning assessment but were nevertheless maintained to keep the number of cues

(hence the level of difficulty) included in the original task. Within each pair, the two cues were associated to the two possible outcomes with reciprocal probabilities (0.75/0.25 and 0.25/0.75). On each trial, one pair was randomly presented on screen. For each pair, the position of the cues on the left versus right side of the screen was counterbalanced across trials within a session. Gain and loss pairs were presented in 24 trials each, while neutral pairs only appeared in 12 trials. The 60 trials of a session were divided into 12 mini-blocks comprising two gain, one neutral and two loss trials. The order of conditions (gain, loss, neutral) within each mini-block was randomized. Participants were not informed about the conditions nor about the mini-block structure.

Participants completed four sessions of this learning task. The first session served as a training session and was performed on a laptop computer outside the scanner. It was repeated if performance was below 75% of correct choices, or if it appeared that the participant misunderstood the instructions. Each session comprised 3 novel pairs of cues. The associations between cues and outcomes were counterbalanced across participants (except for the training session). Participants were given a 4-button box (fORP 932, Current Designs Inc, Philadelphia, USA) placed under their right hand to make their choices. Once the cues appeared on screen, participants had 3 seconds to press a left button with their index finger for selecting the left option, or a right button with their major finger for selecting the right option. They were asked to keep pressing the button until the selected option appeared in red on the screen. The choice was considered valid only if the button was still being pressed at the end of the 3s delay, otherwise it was considered as a 'miss'. Participants were explained that a missed trial always resulted in the worse possible outcome of a given pair (0€ in the gain and neutral conditions, and -10€ in the loss condition). After the 3s delay, the chosen cue was framed in red and then the outcome (-10€/0€/10€) was displayed on screen. At the end of a session, participants were provided with feedback about their cumulative payoff. In order to maximize payoff, participants learned to choose the most rewarding cue in the gain condition and the less punishing cue in the loss condition.

The motor and cognitive performance tasks (i.e., grip and Stroop tasks, Figure 2) were similar to those used in previous studies (Pessiglione et al., 2007; Schmidt et al., 2008, 2009, 2012; Meyniel et al., 2013; Vinckier et al., 2022). Participants were told that their goal was to accumulate as much money as possible across trials. Every trial started with a fixation cross displayed at the center of the screen for 500ms. Then, the money

at stake for the current trial was displayed as a coin or banknote image for a jittered duration (1 to 4 s), which was either crossed for loss trials or not crossed for gain trials. There were 12 possible incentive levels: -20€, -5€, -1€, -0.5€, -0.2€, -0.01€ in the loss condition and 0.01€, 0.2€, 0.5€, 1€, 5€ or 20€ in the gain condition. Next, a graduate scale appeared on screen, which was the trigger for participants to perform the task (squeezing a handgrip or making numerical comparisons). Each graduation of the scale corresponded to 10% of the monetary incentive. The time window allocated to task performance was 5s for the grip task and 70% of calibration time for the Stroop task (see below). The trial ended by a screen providing feedback on the money gained or lost with the last performance and a cumulative total over all preceding trials of the current task. Feedback display lasted for a randomly jittered duration between 1 and 4 seconds.

Both motor and cognitive performance tasks comprised 60 trials per session, divided into 5 mini-blocks of 12 trials presenting each incentive level once, in a randomized order. Before scanning sessions, participants were trained on both tasks with a short 12-trial version. All tasks were performed with the right hand.

In the motor performance task, force was produced on a fMRI-compatible homemade power handgrip that has already been used in previous studies (Meyniel et al., 2013; Schmidt et al., 2009). The handgrip was composed of two plastic cylinders compressing an air tube when squeezed. The tube led to the control room, where it was connected to a transducer converting air pressure into voltage. Thus, grip compression resulted in the generation of a differential voltage signal, linearly proportional to the force exerted. The signal was fed to the stimuli presentation PC via a signal conditioner (CED 1401; Cambridge Electronic Design) and then read inside Matlab. Performance in the scanning sessions was normalized to the maximal force assessed during calibration, when participants were asked to squeeze the handgrip as hard as they could with their right hand. Maximal force was taken as the greatest peak reached over three calibration trials. Unbeknownst to participants, the top of the performance scale (100% of the incentive) in grip task trials was adjusted such that producing the maximal force observed during calibration would correspond to 75% of the incentive (gained or not lost). In case a higher peak was reached during task performance, the new maximal force was used to normalize the scale in the next session. Note that participants could not win more than the full incentive offered in a

given trial. During the 5-s performance window, participants could see a bar indicating the instantaneous force being produced on the handgrip. Participants were informed that payoff was based on peak force and not on the duration of squeezing, so they tended to produce short pulses. Because force was measured through air pressure, which can vary within a task session (with temperature for instance), performance was actually calculated as the difference between the peak reached (within the 5-s window) and a baseline signal (mean over the 500-ms fixation cross window), normalized by the maximal force.

In the cognitive performance task, participants were shown 10 pairs of digits aligned to the graduations of the scale dividing the incentive into 10% steps. To move up one step, participants had to indicate which digit was numerically higher, by pressing the button on the correct side with their right hand. The digits varied in both their numerical size (between 0 and 9) and their physical size (two possible fonts). In each pair, the two digits had both a different numerical size and a different physical size. Incongruent pairs, where the numerically bigger digit is not the same as the physically bigger digit, are known to generate a Stroop effect (Dadon and Henik, 2017). They therefore require allocation of attention to prevent interference and maintain accurate performance. There were 5 incongruent and 5 congruent pairs in each trial, displayed from bottom to top in a randomized order, the numerical distance between the two digits of a pair being varied from 1 to 5. The time given to participants was based on their performance during calibration. Before scanning sessions, participants performed three calibration trials in which they were to make 10 numerical comparisons as fast as possible. The shortest of the three calibration trials provided a duration that was used similarly to maximal force in the grip task. Unbeknownst to participants, the time window for Stroop task trials was set to 70% of the shortest duration measured during calibration. When participants made an error (pressing the button on the wrong side), digits turned red and the bar was frozen for 10% of the total time window. This time penalty for errors was meant to prevent participants from pressing both buttons at random.

Participants were trained on each task before going to the MRI scanner. During fMRI recording, they did 7 task sessions, with learning task in sessions 1, 4 and 7, and performance tasks in sessions 2-3 and 5-6, the order between grip and Stroop tasks being counterbalanced across participants.

Behavioral data analysis

All data were analyzed using MATLAB 2017a (The MathWorks), using scripts that can be found at https://github.com/NicolasClairis/value_estimation_vs_effort_mobilization. Dependent variables were choice (selected cue) and choice RT (from cue onset to button press) in the learning task. In the other tasks, the main dependent variable was performance, defined as the proportion of the incentive gained or not lost in both the grip task (where it corresponds to the proportion of maximal force) and the Stroop task (where it corresponds to the number of numerical comparisons correctly done). RT was defined as the latency at which produced force exceeded 1% of maximal force in the grip task and at which the first button press was made in the Stroop task. Dependent variables were analyzed using general linear models at the individual level followed by t-tests on regression estimates at the group level (as explained in the Results). More specific effects of experimental factors were tested using computational models.

Computational modeling

All computational models were inverted using Matlab VBA toolbox (available at http://mbb-team.github.io/VBA-toolbox/), which implements a variational Bayesian algorithm under the Laplace approximation (Daunizeau et al., 2014). The algorithm provides efficient and robust estimates of posterior distributions for the model free parameters, initially defined using Gaussian prior distributions.

Learning task. Choice behavior was fitted at the individual level using a standard "Q-learning" model (Watkins and Dayan, 1992), as was previously done with this task (Palminteri et al., 2012; Pessiglione et al., 2006). Each cue is associated to a Q-value that represents the expected reward (or punishment) if selected. As participants have no prior information when starting a learning session, all Q-values are initialized at zero. Q-values are then updated after every choice according to a delta rule adapted from the Rescorla and Wagner model:

$$\boldsymbol{Q_{CH}(t+1) = Q_{CH}(t) + \alpha \cdot (outcome(t) - Q_{CH}(t))} \quad (1)$$

where $Q_{CH}(t)$ is the expected value of the option chosen at trial t, $Q_{CH}(t+1)$ the expected value of the same option after updating, α a learning rate that adjusts the weight of the last observation relative to older ones, and $outcome(t)$ coded as 1 in

case of gain (+10 €), 0 in case of neutral feedback (0 €), and -1 in case of loss (-10 €). Note that the expression $outcome(t) - Q_{CH}(t)$ corresponds to prediction error, PE_{CH}.

To improve the fit, we integrated counterfactual reasoning (Ben-Artzi et al., 2023), using the same equations (with the same parameters) for updating the expected value of the cue that was not chosen. This implies that participants understood, during the training session, that the two cues of a given pair at a given trial yielded opposite outcomes (counterfactual outcome is 0 / 1 if actual outcome is 1 / 0 in the gain condition, and -1 / 0 if actual outcome is 0 / -1 in the loss condition).

Q-values were then used to derive the probabilities of selecting each option, according to the softmax formula:

$$\boldsymbol{p}(\boldsymbol{cue}) = \frac{\mathbf{1}}{\mathbf{1}+\boldsymbol{e}^{-\frac{\boldsymbol{DV}}{\boldsymbol{\beta}}}} \quad (2)$$

where β is a temperature parameter that controls the stochasticity of choices and DV the difference between the value of the considered cue and that of the other cue in the pair.

The same parameters α and β were used to fit choices made in all conditions and session. Prior distributions were centered on 0 for α and 1 for β (which was constrained to be strictly positive).

Performance tasks. In both grip and Stroop tasks, performance was fitted at the individual level using an effort regulation model, as was done previously with similar tasks (Le Bouc et al., 2016; Vinckier et al., 2022). This model is based on the principle that participants exert the amount of effort that maximizes a cost/benefit tradeoff. For each possible effort level, the benefit is the money gained or not lost. The cost is directly proportional to the effort invested, but may increase with fatigue across trials for a same fatigue level. The expected value function EV at a given trial t can be written as:

$$EV(E,t) = B(E,t) - C(E,t) \quad (3)$$

where $\boldsymbol{B}(\boldsymbol{E},\boldsymbol{t})$ and $\boldsymbol{C}(\boldsymbol{E},\boldsymbol{t})$ are benefit and cost expected at trial $\boldsymbol{t}$ if investing the amount of effort *E*.

The subcomponents can be decomposed as follows:

$$\boldsymbol{B(E,t) = P(E) + k_I \cdot (P(E) \cdot G(t) + (1 - P(E)) \cdot L(t))} \quad (4)$$

where $\boldsymbol{P(E)}$ is the performance expected if investing an amount of effort *E*, *G(t)* and *L(t)* the signed monetary incentive proposed at trial *t* (G=0 in a loss trial and L=0 in a gain trial). The first term accounts for performance increasing the benefit independently of financial outcomes (i.e., good performance is intrinsically valuable). The two other terms allow maximal performance to win the full incentive in gain trials, and to lose nothing in loss trials. Note that gain and loss terms were normalized to their maximum (i.e., divided by 20€).

$$\boldsymbol{C(E,t) = k_C \cdot (1 + k_T \cdot t) \cdot E} \quad (5)$$

where *t* is the normalized trial number (divided by 60, the total number of trials).

The k_X constants are weight parameters that control the sensitivity to the different factors (incentives and effort cost).

Finally, the impact of effort mobilization on performance was defined by a saturation function such that maximal performance is reached with infinite effort exertion:

$$\boldsymbol{P(E) = \frac{E}{\gamma + E}} \quad (6)$$

where γ is a (positive) parameter that controls the curvature of the *E* to *P* mapping.

The optimal effort *E** to be invested in a given trial *t* is obtained when the derivative $\frac{dEV}{dE}$ is null, which gives:

$$\boldsymbol{E(t) = \sqrt{\gamma \cdot \frac{1 + k_I \cdot (G(t) - L(t))}{k_C \cdot (1 + k_T \cdot t)}} - \gamma} \quad (7)$$

From this equation can be derived the optimal performance that represents the model prediction for trial *t*:

$$\boldsymbol{P(t) = 1 - \sqrt{\gamma \cdot \frac{k_C \cdot (1 + k_T \cdot t)}{1 + k_I \cdot (G(t) - L(t))}}} \quad (8)$$

All prior distributions were centered on one and all parameters were forced to be positive. The VBA_toolbox allowed us to obtain, for a given participant, the set of posterior means k_I, k_C, k_T and γ with which the model best matches the observed pattern of performance across trials. The two sessions of the same task were modeled

with the same parameters, but motor and cognitive performance tasks were modeled with different sets of parameters.

MRI data acquisition

Magnetic Resonance Imaging (MRI) was performed at the research neuroimaging center (CENIR) with a Siemens Magnetom Prisma 3-T scanner using a 64-channel head/neck coil. Structural T1-weighted images were co-registered to the mean echo planar image (EPI), segmented and normalized to the standard T1 template and then averaged across subjects for anatomical localization of group-level functional activation. Functional T2*-weighted EPIs were acquired with blood-oxygen-level dependent (BOLD) contrast using the following parameters: repetition time TR = 1.10 seconds, echo time TE = 25ms, flip angle = 60°, number of slices = 54, slice thickness = 2.0mm, field of view = 208mm, multiband accelerating factor: 3, voxel size: 2x2x2 mm. Note that the number of volumes per session was not predefined, because the time available for performance in the Stroop task varied across individuals. Volume acquisition was just stopped when the task session was completed. The number of volumes per session (mean±SD) was 369±5 in the learning task, 565±58 in the Stroop task, 592±6 in the grip task. Across participants, the total duration was between 17 and 27 minutes for the 3 sessions of the learning task, the 2 sessions of the grip task and the 2 sessions of the Stroop task.

fMRI data analysis

Functional images were preprocessed and analyzed using the SPM12 toolbox (Wellcome Trust Center for NeuroImaging) running in MATLAB 2017a. Preprocessing consisted of spatial realignment, normalization using the same transformation as structural images, and spatial smoothing using a Gaussian kernel with a full width at half maximum (FWHM) of 8 mm.

Preprocessed data were analyzed using generalized linear models (GLM) in SPM12 at the first (individual) level and then tested for significance at the second (group) level. All GLM included the 6 movement regressors generated during realignment of successive scans. Each task session was modeled separately.

For the learning task, the main GLM (GLM1) included a boxcar function encompassing the choice period (from cue onset to the end of the 3-s window), modulated by the following parametric regressors: (1) value (Val), (2) confidence (Conf), (3) deliberation

time (DT). Val was defined as the sum of cue values weighted by their choice probabilities ($p_{CH} \cdot Q_{CH} + p_{UC} \cdot Q_{UC}$), which has been referred to as state value in the reinforcement learning framework (e.g., (Palminteri et al., 2009). Following on our previous publication (Clairis and Pessiglione, 2022), Conf was defined as the squared distance from choice probability to chance level, normalized to a 0-1 range ($i.e., [2(p(left) - 0.5)]^2$). This is equivalent to taking the squared difference in choice probability between the left and right options ($i.e., [(p(left) - p(right))]^2$). DT was defined as the duration from cue onset to the start of button pressing. Gain and loss pairs of cues were modeled in a same regressor, but neutral pairs were modeled separately, with DT as a single parametric modulation, and were not included in the following analyses. The GLM also included a boxcar function encompassing the period from chosen option to outcome onset, and a stick function for the outcome itself, modulated by the prediction error generated by the model (PE_{CH}).

All regressors of interest were z-scored and convolved with the canonical hemodynamic response function (HRF). Parametric modulators were not orthogonalized in the main GLM so that they could compete for explaining the variance of fMRI time series. Several alternative GLM were built to test variants of GLM1. Two GLM were identical to GLM1, except that all regressors were serially orthogonalized, in either the Val/Conf/DT (GLM2) or the DT/Conf/Val (GLM3) order. Two other GLM were identical to GLM1, except that Val was defined as the sum of option values ($Q_{CH} + Q_{UC}$) in GLM4 and the difference between the chosen and unchosen option values ($Q_{CH} - Q_{UC}$) in GLM5.

For the performance tasks, the main GLM (GLM1) included a stick function for incentive onset, modulated by the following regressors (not orthogonalized): (1) the optimal effort E* generated by our computational model and (2) reaction time (RT). The performance time window (from scale onset to feedback display) was also modeled as a boxcar function, and the feedback onset as a stick function. As done for the learning task, we tested alternative GLM identical to GLM1, except that the two parametric regressors were orthogonalized, either in the E*/RT order (GLM2) or the RT/E* order (GLM3). A last alternative to GLM1 was built (GLM4) where, instead of modulating the time of incentive onset, E* and RT (not orthogonalized) were modulating the performance time window.

Note that images at the individual-level analysis were masked following the default SPM procedure, which removes any voxel with a signal intensity below 20% of the global mean, to exclude voxels outside the brain for group-level analyses. We verified that our main conclusions were still valid when using a white+grey matter mask (including all voxels with a probability to be in either grey or white matter above 5%, based on the average anatomical segmentation performed by SPM). Uncorrected maps obtained with this more inclusive mask can be found in Neurovault at the following repository address: https://neurovault.org/collections/15543/.

For the maps shown in the figures, we used an additional medial PFC (mPFC) inclusive mask (see Extended Data Figure 4-1) based on the aggregation of supplementary motor area, anterior and mid-cingulate area, gyrus rectus, middle frontal gyrus and superior frontal medial gyrus from the AAL atlas (Tzourio-Mazoyer et al., 2002). This masking procedure just filtered the voxels within the mPFC and was only used for display purposes; it did not impact statistical results, which were always calculated across the whole brain. In all figures and tables, the statistical threshold was set at $p<0.001$ uncorrected at the voxel level and $p<0.05$ family-wise error corrected for multiple comparisons at the cluster level. We defined our regions of interest as 8mm-radius spheres centered on the MNI coordinates of group-level clusters associated with Val (-10; 48; -12), Conf (-8; 52; 18) and DT (10; 12; 48) in our previous study (see Figure 3A). Violin plots were generated using the *violinplot* Matlab function developed by Bastian Bechtold (https://github.com/bastibe/Violinplot-Matlab, doi: 10.5281/zenodo.4559847).

Pupil size

Pupil diameter was recorded at a sampling rate of 1 kHz, using an EyeLink 1000 plus (SR Research) eye-tracker. The eye-tracker was calibrated before the start of fMRI sessions, once the subject was positioned inside the scanner. Interpolation was performed with Matlab *interp1* function, which implements the *pchip* cubic interpolation method to compensate for any period of time when the pupil signal was lost because of blinks. The pupil size time series were subsequently band-pass filtered (1/128-1 Hz) and z-scored per individual and per session.

Within-trial variations in pupil size were baseline-corrected by removing the mean signal over the 200ms preceding stimulus onset and time-locked to stimulus onset.

Then trial-wise variations in pupil size were fitted separately for the grip and for the Stroop task with a linear regression model that included factors of no interest (an intercept per block and stimulus luminance) and variables of interest (the effort E and the reaction time RT). Within-trial individual time series of regression estimates were then smoothed using a 200ms kernel. Group-level significant time clusters were identified after correction for multiple comparisons estimated according to random field theory, as implemented in the VBA toolbox (available at http://mbb-team.github.io/VBA-toolbox/). To complement this analysis, we also averaged the betas over a 5-second period following stimulus onset for each individual, and then performed a one-sample t-test against zero at the group-level.

Results

Behavior in the learning task

Participants (n=22) performed three sessions of a probabilistic instrumental learning task (Figure 1). They learned to select cues with high gain probability (75%) and low loss probability (25%), reaching an average correct choice of 88.51±2.98% in the gain condition and 80.13±2.15% in the loss condition (Figure 3A), which was significantly above chance level in both cases ($p = 2 \cdot 10^{-11}$ and $p = 4 \cdot 10^{-12}$, respectively). Learning curves were fitted using a standard Q-learning algorithm (with a balanced accuracy of 0.706±0.018), the two free parameters being adjusted individually (0.130±0.015 for the learning rate α and 0.164±0.017 for the choice temperature β). We used the cue values and choice probabilities provided by the fitted model to generate the constructs that we regressed against fMRI data. At each trial, value (Val) was defined as the sum of cue values weighted by choice probabilities, and confidence (Conf) as the squared difference between choice probability and chance level (0.5). Note that, by design (Figure 3B), Val and Conf were partially decorrelated in this task (r = 0.204±0.009), because while Conf always increased across trials (with learning), Val increased in the gain condition but decreased in the loss condition. There was also a modest but significant correlation between deliberation time (DT) and Val (r = -0.464±0.022), due to faster responses in the gain relative to loss condition, and between DT and Conf (r = -0.257±0.032), due to speed and accuracy improvement across trials in both conditions. Indeed, linear regression (Figure 3C) showed that DT was shorter both when Val was higher (β_{Val} = -0.452±0.039, $p = 1 \cdot 10^{-10}$) and when Conf was higher (β_{Conf} = -0.858±0.241, p = 0.0018). Together, Val and Conf explained 32.01±2.14% of the variance in DT, which may call for orthogonalization of these regressors (see below).

Neural activity in the learning task

As in our previous study (Clairis and Pessiglione, 2022), whole-brain maps (corrected for multiple comparisons) revealed a functional partition between the ventromedial, midmedial and dorsomedial regions of the prefrontal cortex (vmPFC, mmPFC and dmPFC), which respectively reflected the Val, Conf and DT constructs (Figure 4A). There was no negative association with any of these constructs that would survive correction at the whole-brain level. We systematically tested the links between all three

variables and all three regions of interest defined using the previous dataset to avoid non-independence issues (Figure 4B). The same 3 correlations were observed between Val and vmPFC activity ($\beta = 0.297 \pm 0.101$; $p = 0.008$), between Conf and mmPFC activity ($\beta = 0.252 \pm 0.073$; $p = 0.002$) and between DT and dmPFC activity ($\beta = 0.321 \pm 0.065$; $p = 7 \cdot 10^{-5}$). All 3 correlations remained significant (Extended Data Figure 4-2) when orthogonalizing regressors, whatever the order (Val/Conf/DT in GLM2 or DT/Conf/Val in GLM3). We note however that the orthogonalization might have generated spurious correlations. Indeed, when Val was orthogonalized to DT, the correlation between DT and vmPFC activity became significant, and reciprocally, when DT was orthogonalized to Val, the correlation between Val and dmPFC activity became significant. This is likely due to Val and DT sharing some common variance, which was attributed to one regressor or the other, depending on the order of serial orthogonalization. Apart from the 3 main ROI-variable associations, we also observed a correlation between Conf and vmPFC activity ($p<0.001$ in all GLM). Beyond the medial prefrontal cortex, significant correlation with DT (after cluster-wise correction for multiple comparisons) was observed in several other brain regions (see Extended Data Table 4-6), including the anterior insula and dorsolateral PFC, two brain regions classically involved in exerting effort and/or cognitive control.

While the link between vmPFC activity and subjective value is well established, there is still some debate about what values exactly are represented during a binary choice. We have therefore regressed vmPFC activity against alternative GLM in which the weighted sum of option values (GLM1) was replaced either by the straight sum (GLM4) or the difference between chosen and unchosen option values (GLM5). Significant regression estimates (see Extended Data Figure 4-3) were observed with the sum (GLM4: $\beta = 0.284 \pm 0.098$; $p = 0.009$), but not with the difference (GLM5: $\beta = 0.132 \pm 0.115$; $p = 0.263$). We kept the weighted sum regressor because it explained more variance in vmPFC activity (GLM1: $\beta = 0.297 \pm 0.101$; $p = 0.008$), but we conclude that any regressor modeling a positive correlation with option values would also provide a significant fit.

Behavior in the performance tasks

Between learning sessions, participants performed two sessions of the motor and cognitive performance tasks (Figure 2). These tasks required the allocation of either force (grip task) or attention (Stroop task) in order to maximize the monetary payoff,

which depended on both the incentive level and the performance achieved in a particular trial. We verified that, as intended, performance improved with incentive level (Figure 5) in both grip and Stroop tasks, and both gain and loss conditions. Indeed, higher performance was achieved when unsigned incentives (*i.e.* stakes), were greater in both the grip task ($\beta_{|I|}$ = 1.164±0.128; p = $9 \cdot 10^{-9}$) and the Stroop task ($\beta_{|I|}$ = 0.190±0.053; p = 0.0016). This effect of incentive motivation on performance was similar in gain and loss trials. Also, performance decreased with trial number in both the grip task (β_T = -0.125±0.033; p = 0.0011) and the Stroop task (β_T = -0.077±0.021; p = 0.0015), probably reflecting the emergence of fatigue. This pattern was not observed with response time (RT), defined as the start of force production in the grip task and first button press in the Stroop task (Figure 5). Although not significant in all cases, the trend was the opposite: RT tended to be shorter with higher incentive levels (grip: $\beta_{|I|}$ = -0.0030±0.003; p = 0.331; Stroop: $\beta_{|I|}$ = -0.0035±0.0009; p = $8 \cdot 10^{-4}$), and longer with higher trial number (grip: β_T = 0.0016±0.0007; p = 0.041; Stroop: β_T = 0.0007±0.0004; p = 0.075).

To fit performance, we developed a computational model (see Methods) adapted from previous studies using similar tasks (Le Bouc et al., 2016; Vinckier et al., 2022). The model with fitted parameters was then used to generate the best possible proxy for the effort invested in motor and cognitive performance, so we could use it to identify the underlying neural activity.

Neural activity in the performance tasks

Whole-brain maps (corrected for multiple comparisons) highlighted dmPFC as showing a positive association between activity triggered by incentive display and the optimal effort E* that was estimated by the computational model fitted to the behavioral data. This was true across motor and cognitive performance tasks, as it was significant in a conjunction analysis (Figure 6). Several other significant clusters (see Extended Data Table 6-2) were observed outside the medial prefrontal cortex in this conjunction analysis (even after voxel-wise correction for multiple comparisons), notably in the striatum, a brain region that has been involved in incentive motivation. There was no negative association with effort E* that would survive correction at the whole-brain level. When testing the dmPFC ROI identified in our previous study (Clairis and Pessiglione, 2022), activity at incentive display was correlated with E* in both grip and Stroop tasks (Figure 6, both p<0.005). Note that RT did not yield any significant

positive correlation in whole-brain analysis ($p<0.001$, uncorrected for multiple comparisons) when pooling grip and Stroop tasks together. Moreover, the correlation with E* (but not RT) also held when variables were orthogonalized in serial order (Extended Data Figure 6-1), either following the E*/RT order (GLM2) or the RT/E* order (GLM3). Also, when replacing E* by unsigned incentive level, regression coefficients were significantly lower ($\beta = 1.850 \pm 0.273$ vs. 1.952 ± 0.266; $p = 0.0235$), indicating that dmPFC activity better reflected effort than stakes. In GLM4, which focuses on the performance time window (squeezing the handgrip or making numerical comparisons), the association between dmPFC activity and optimal effort E* was no longer significant (grip task: $b = 0.009 \pm 0.055$; $p = 0.868$; Stroop task: $b = -0.083 \pm 0.078$; $p = 0.299$), suggesting that dmPFC activity was reflecting an antecedent more than a consequence of effort exertion.

Pupil dilation in the performance tasks

As another marker of effort exertion, we investigated pupil dilation in the grip and Stroop task (Figure 7). Over the 0-5s time window, the correlation with pupil diameter was globally positive for optimal effort E* (grip: $\beta_{E^*} = 0.094 \pm 0.038$, $p = 0.022$; Stroop: $\beta_{E^*} = 0.042 \pm 0.031$, $p = 0.184$) and negative for RT (grip: $\beta_{RT} = -0.069 \pm 0.017$, $p = 6 \cdot 10^{-4}$; Stroop: $\beta_{RT} = -0.072 \pm 0.013$, $p = 3 \cdot 10^{-5}$). In the grip task, pupil size was significantly correlated (after correction for multiple comparisons) with effort E* from 3.11s to 6.26s following scale onset. The trend was similar in the Stroop task but there was no time point at which correlation between E* and pupil size would survive correction for multiple comparisons. Nevertheless, these results support the notion that more effort is associated to shorter RT in these performance tasks, in contrast with what was observed in the deliberation tasks (i.e., during rating, choice and learning).

Conjunction across learning and performance tasks

Finally, we tested the conjunction between activity associated with DT in the learning task and E* in the grip and Stroop tasks. The conjunction was significant in a dmPFC cluster (Figure 8A), together with clusters in the anterior insula and dorsolateral PFC (Extended Data Table 8-1). Thus, the same dmPFC region reflected the time invested in deliberation and the effort invested in motor and cognitive performance. We also overlapped this dmPFC cluster with the dmPFC cluster associated with DT in our

previous study (Clairis and Pessiglione, 2022). The overlap (Figure 8B) confirmed that a common dmPFC region was also reflecting the time invested in expressing preference (during rating and choice tasks). Together, these results support the implication of the dmPFC in effort mobilization across preference, learning and performance tasks.

Discussion

In this study, we first replicate, in the context of instrumental learning, the functional partition of the mPFC that was initially demonstrated across choice and rating tasks (Clairis and Pessiglione, 2022). When values are generated by a reinforcement learning model, instead of expressed as subjective preferences, we still observe that option values (Val), choice confidence (Conf) and deliberation time (DT) are respectively reflected in the activity of vmPFC, mmPFC and dmPFC during decision-making. We then strengthen our functional interpretation of the correlation with DT as signaling effort mobilization, in performance tasks where participants maximize a tradeoff between reward prospect and effort cost. During preparation of both motor and cognitive performance, we find that dmPFC activity reflects the optimal effort to be exerted according to an effort regulation model.

The functional partition was based on a theoretical analysis of judgment and decision processes. In brief, we argue that what is maximized in rating and choice tasks is the confidence in the eventual response (Lebreton et al., 2015; Lee and Daunizeau, 2021). Thus, on top of the first-level decision process that estimates option values, a second-level metacognitive process arbitrates the tradeoff between an expected gain in confidence and the time invested in deliberation. There is therefore a need for representing these three types of variables in brain activity. A difficulty for the dissociation of these variables in the analysis of fMRI data is that they are more or less correlated, depending on the task. Here, we take advantage of gain and loss conditions to decorrelate value and confidence: while confidence increases across trials with learning, values increase in the gain condition but decrease in the loss condition. To infer both value and confidence from choice behavior, we use a classical reinforcement learning model that was already validated as providing a good account of choice behavior in this task (Palminteri et al., 2012; Pessiglione et al., 2006). We find option value representations in more ventromedial regions, and choice confidence representations in more dorsomedial regions. This is reminiscent of the dissociation previously reported (Clairis and Pessiglione, 2022), where confidence was decorrelated from the option value (in rating tasks) or the sum of option values (in choice tasks). It is also consistent with previous demonstration that value-to-confidence representation follows a ventral-to-dorsal gradient in the PFC (De Martino et al., 2017).

The dissociation observed here is only partial, as activity in our ventromedial ROI is significantly associated with both option values and choice confidence. Although the vmPFC has been identified in meta-analyses of fMRI studies as a valuation hub (Bartra et al., 2013; Clithero and Rangel, 2014), it has also been shown to signal confidence in non-valuation tasks (e.g., (Gherman and Philiastides, 2018; Rouault et al., 2023)). A similar overlap of value and confidence representations in the vmPFC has been reported in previous fMRI studies using both rating and choice tasks (De Martino et al., 2012; Lebreton et al., 2015; Shapiro and Grafton, 2020). This is also consistent with a MEG study that observed both the sum and difference of option values being reflected in the vmPFC low-frequency activity (Hunt et al., 2012), because our confidence construct is close to the difference between chosen and unchosen option values. Unsurprisingly, many studies have reported correlations between vmPFC activity and chosen option value, which is correlated with both the sum and the difference (e.g., (Baram et al., 2021; Gershman et al., 2009; Gläscher et al., 2009; Seaman et al., 2018; Wunderlich et al., 2009)). The correlation with value difference is inconsistent studies, being significant in some (Boorman et al., 2009; Chau et al., 2014) but not in others (Jocham et al., 2014; Lim et al., 2011; Lopez-Persem et al., 2016; Qin et al., 2011; Ting et al., 2023; Van der Laan et al., 2012). We suspect that the issue might relate, at least in some cases, to the potential confound between value difference and choice confidence. Indeed, when comparing different combinations of option values, we observe that vmPFC activity correlates with either the straight sum or the sum weighted by choice probability, but not with the difference (when included together with confidence in a same GLM).

Yet the correlation with the sum does not tell whether the vmPFC signals the overall value of the option set, as previously suggested (Shenhav and Karmarkar, 2019), or the value of each option, independently. The sum of option values weighted by their choice probability could represent a state value, as defined in reinforcement learning theory (Sutton and Barto, 1998), but also a succession of option value estimates modulated by attention, as proposed in some versions of sequential sampling models (Krajbich et al., 2010). A time-resolved recording technique, such as intracerebral EEG, coupled with an eye-tracking device, might help address this issue. Finally, we acknowledge that the identification of confidence representation is limited by the absence of trial-by-trial confidence rating, which would have provided a finer

633 estimation than the approximation generated by our model. The function that we use
634 here is clearly not the only one possible, but it was previously validated as an accurate
635 proxy for how confidence ratings vary with option values, on average (Clairis and
636 Pessiglione, 2022).

637 Contrary to value and confidence, we find correlates of DT in the dorsomedial part of
638 the prefrontal cortex. This is a direct replication of the correlation previously observed
639 in rating and choice tasks (Clairis and Pessiglione, 2022) and also reported in several
640 fMRI studies (e.g., (Grinband et al., 2011; McGuire and Botvinick, 2010; Wu et al.,
641 2018). Because DT is correlated with both value and confidence in our task, we have
642 tested whether GLMs with and without orthogonalization would account for dmPFC
643 activity. The Val-vmPFC and DT-dmPFC associations are significant whether or not
644 regressors are orthogonalized, and whatever the order of orthogonalization. Yet the
645 correlation cannot tell whether dmPFC activity reflects some relevant cognitive
646 variable that would determine DT, such as the presence of conflict (Yeung et al., 2011)
647 or the need for cognitive control (Shenhav et al., 2013). Indeed, variations in RT are
648 susceptible to many factors (such as distraction, fatigue or mind wandering) that may
649 induce changes in dmPFC activity. Alternatively, the increase in dmPFC activity with
650 DT might not reflect a cognitive antecedent but a by-product of any process lasting
651 longer (Alexander and Brown, 2011; Grinband et al., 2011; Weissman and Carp,
652 2013).

653 We therefore test our initial interpretation that the correlation with DT denotes a
654 representation of the effort invested in deliberation, across two other tasks assessing
655 motor and cognitive performance (grip force production and Stroop numerical
656 comparison). Although incentives are varied to manipulate motivation, these tasks do
657 not involve proper valuation processes, because coins and banknotes have an
658 obvious monetary value. There is no clear need for an estimation of confidence either,
659 because there is no uncertainty in how much reward each performance level would
660 bring (this is explicitly indicated on the screen). We therefore dropped the value and
661 confidence constructs and focused on effort mobilization. The amount of effort exerted
662 can be inferred from behavioral performance by fitting a computational model that was
663 previously validated using the same tasks (Le Bouc et al., 2023; Vinckier et al., 2022).
664 Here, we further validate the theoretical effort output by the model as being positively
665 associated with pupil dilation, which can arguably be considered a measure of physical

and mental effort exertion (Hess and Polt, 1964; Kahneman and Beatty, 1966; van der Wel and van Steenbergen, 2018; Zénon et al., 2014).

Critically, dmPFC activity reflects this effort proxy but not RT, during the preparation period but not during effort exertion. This result provides a strong argument for the idea that dmPFC activation is not a mechanical by-product of RT prolongation but a reflection of effort mobilization for the upcoming task. This idea might be more general than triggering cognitive control, since the association between the effort proxy and dmPFC activity was observed for both motor and cognitive performance. It is consistent with previous observations that reward and effort are both represented in dmPFC activity (often called dACC), whatever the types of reward and effort (Le Bouc et al., 2022; Pessiglione et al., 2018; Shenhav et al., 2013). Finally, the conjunction between preference, learning and performance tasks shows that the same dmPFC cluster reflects DT during decision and theoretical effort during motor and cognitive preparation. We cannot rule out that the correlation with DT and E* might arise while the dmPFC cluster would still serve different functions in the different tasks, but a more parsimonious interpretation would involve the dmPFC in mobilizing effort in all tasks, as suggested by a recent meta-analysis of fMRI studies (Lopez-Gamundi et al., 2021).

In conclusion, we provide here evidence that implicate ventromedial regions of the PFC in the estimation of option value and response confidence, and dorsomedial regions in the adjustment of effort mobilization for an appropriate performance. These results extend previous findings and thereby contribute to establishing functional specifications of brain regions that are robust across a variety of behavioral tasks. However, many questions remain unaddressed. Notably, we have dissociated the neural representations of value, effort, and confidence, but have not brought any insight into the mechanisms that must link these representations, such that investing effort would enable refining value estimates to gain confidence in the response. Also, we have used a computational definition of effort mobilization, but have not contributed to elucidating what effort might represent at the biological level. It could be related to autonomic arousal, given the link observed with pupil dilation, and the known connections between dmPFC regions and the autonomic nervous system (Amiez and Procyk, 2019; Beissner et al., 2013; Critchley et al., 2003). It could also be related to metabolic support within the brain, as suggested by studies that examined the cost of cognitive effort (Holroyd, 2016; Wiehler et al., 2022; Zénon et al., 2019). Further

699 research is needed to bridge the informational and biophysical notions of effort
700 mobilization.

References

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., Pessiglione, M., 2015. Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. Journal of Neuroscience 35, 2308–2320. https://doi.org/10.1523/JNEUROSCI.1878-14.2015

Alexander, W.H., Brown, J.W., 2011. Medial prefrontal cortex as an action-outcome predictor. Nat Neurosci 14, 1338–1344. https://doi.org/10.1038/nn.2921

Amiez, C., Procyk, E., 2019. Midcingulate somatomotor and autonomic functions. Handb Clin Neurol 166, 53–71. https://doi.org/10.1016/B978-0-444-64196-0.00004-2

Aridan, N., Malecek, N.J., Poldrack, R.A., Schonberg, T., 2019. Neural correlates of effort-based valuation with prospective choices. Neuroimage 185, 446–454. https://doi.org/10.1016/j.neuroimage.2018.10.051

Baram, A.B., Muller, T.H., Nili, H., Garvert, M.M., Behrens, T.E.J., 2021. Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. Neuron 109, 713-723.e7. https://doi.org/10.1016/j.neuron.2020.11.024

Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. Nature Neuroscience 10, 1214–1221. https://doi.org/10.1038/nn1954

Beissner, F., Meissner, K., Bar, K.-J., Napadow, V., 2013. The Autonomic Brain: An Activation Likelihood Estimation Meta-Analysis for Central Processing of Autonomic Function. Journal of Neuroscience 33, 10503–10511. https://doi.org/10.1523/JNEUROSCI.1103-13.2013

Ben-Artzi, I., Kessler, Y., Nicenboim, B., Shahar, N., 2023. Computational mechanisms underlying latent value updating of unchosen actions. Science Advances 9, eadi2704. https://doi.org/10.1126/sciadv.adi2704

Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., Rushworth, M.F.S., 2009. How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. Neuron 62, 733–743. https://doi.org/10.1016/j.neuron.2009.05.014

Chang, L.J., Jolly, E., Cheong, J.H., Rapuano, K.M., Greenstein, N., Chen, P.-H.A., Manning, J.R., 2021. Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. Sci Adv 7, eabf7129. https://doi.org/10.1126/sciadv.abf7129

Chau, B.K.H., Kolling, N., Hunt, L.T., Walton, M.E., Rushworth, M.F.S., 2014. A neural mechanism underlying failure of optimal choice with multiple alternatives. Nat Neurosci 17, 463–470. https://doi.org/10.1038/nn.3649

Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. Journal of Neuroscience 29, 12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009

Clairis, N., Lopez-Persem, A., 2023. Debates on the dorsomedial prefrontal/dorsal anterior cingulate cortex: insights for future research. Brain awad263. https://doi.org/10.1093/brain/awad263

Clairis, N., Pessiglione, M., 2022. Value, Confidence, Deliberation: A Functional Partition of the Medial Prefrontal Cortex Demonstrated across Rating and Choice Tasks. J. Neurosci. 42, 5580–5592. https://doi.org/10.1523/JNEUROSCI.1795-21.2022

Clithero, J.A., Rangel, A., 2014. Informatic parcellation of the network involved in the computation of subjective value. Social Cognitive and Affective Neuroscience 9, 1289–1302. https://doi.org/10.1093/scan/nst106

Critchley, H.D., Mathias, C.J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B.-K., Cipolotti, L., Shallice, T., Dolan, R.J., 2003. Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. Brain 126, 2139–2152. https://doi.org/10.1093/brain/awg216

Dadon, G., Henik, A., 2017. Adjustment of control in the numerical Stroop task. Mem Cognit 45, 891–902. https://doi.org/10.3758/s13421-017-0703-6

Daunizeau, J., Adam, V., Rigoux, L., 2014. VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. PLoS Comput Biol 10, e1003441. https://doi.org/10.1371/journal.pcbi.1003441

De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., Love, B.C., 2017. Social Information Is Integrated into Value and Confidence Judgments According to Its Reliability. The Journal of Neuroscience 37, 6066–6074. https://doi.org/10.1523/JNEUROSCI.3880-16.2017

De Martino, B., Fleming, S.M., Garrett, N., Dolan, R.J., 2012. Confidence in value-based choice. Nature Neuroscience 16, 105–110. https://doi.org/10.1038/nn.3279

Economides, M., Guitart-Masip, M., Kurth-Nelson, Z., Dolan, R.J., 2015. Arbitration between controlled and impulsive choices. Neuroimage 109, 206–216. https://doi.org/10.1016/j.neuroimage.2014.12.071

Fouragnan, E., Retzler, C., Philiastides, M.G., 2018. Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. Hum. Brain. Mapp. 39, 2887–2906. https://doi.org/10.1002/hbm.24047

Gershman, S.J., Pesaran, B., Daw, N.D., 2009. Human Reinforcement Learning Subdivides Structured Action Spaces by Learning Effector-Specific Values. J. Neurosci. 29, 13524–13531. https://doi.org/10.1523/JNEUROSCI.2469-09.2009

Gherman, S., Philiastides, M.G., 2018. Human VMPFC encodes early signatures of confidence in perceptual decisions. eLife 7, e38293. https://doi.org/10.7554/eLife.38293

Gläscher, J., Hampton, A.N., O'Doherty, J.P., 2009. Determining a Role for Ventromedial Prefrontal Cortex in Encoding Action-Based Value Signals During Reward-Related Decision Making. Cerebral Cortex 19, 483–495. https://doi.org/10.1093/cercor/bhn098

Grinband, J., Savitskaya, J., Wager, T.D., Teichert, T., Ferrera, V.P., Hirsch, J., 2011. The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. NeuroImage 57, 303–311. https://doi.org/10.1016/j.neuroimage.2010.12.027

Hess, E.H., Polt, J.M., 1964. Pupil Size in Relation to Mental Activity during Simple Problem-Solving. Science 143, 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Hogan, P.S., Galaro, J.K., Chib, V.S., 2019. Roles of Ventromedial Prefrontal Cortex and Anterior Cingulate in Subjective Valuation of Prospective Effort. Cerebral Cortex 29, 4277–4290. https://doi.org/10.1093/cercor/bhy310

Holroyd, C.B., 2016. The waste disposal problem of effortful control, in: Motivation and Cognitive Control, Frontiers of Cognitive Psychology. Routledge/Taylor & Francis Group, New York, NY, US, pp. 235–260.

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F.S., Behrens, T.E.J., 2012. Mechanisms underlying cortical activity during value-guided choice. Nature Neuroscience 15, 470–476. https://doi.org/10.1038/nn.3017

Jimura, K., Chushak, M.S., Braver, T.S., 2013. Impulsivity and self-control during intertemporal decision making linked to the neural dynamics of reward value representation. J Neurosci 33, 344–357. https://doi.org/10.1523/JNEUROSCI.0919-12.2013

Jocham, G., Furlong, P.M., Kröger, I.L., Kahn, M.C., Hunt, L.T., Behrens, T.E.J., 2014. Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. Neuroimage 100, 498–506. https://doi.org/10.1016/j.neuroimage.2014.06.005

Kable, J.W., Glimcher, P.W., 2009. The Neurobiology of Decision: Consensus and Controversy. Neuron 63, 733–745. https://doi.org/10.1016/j.neuron.2009.09.003

Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10, 1625–1633. https://doi.org/10.1038/nn2007

Kahneman, D., Beatty, J., 1966. Pupil diameter and load on memory. Science 154, 1583–1585. https://doi.org/10.1126/science.154.3756.1583

Kolling, N., Behrens, T.E.J., Mars, R.B., Rushworth, M.F.S., 2012. Neural Mechanisms of Foraging. Science 336, 95–98. https://doi.org/10.1126/science.1216930

Kolling, N., Wittmann, M.K., Behrens, T.E.J., Boorman, E.D., Mars, R.B., Rushworth, M.F.S., 2016. Value, search, persistence and model updating in anterior cingulate cortex. Nature Neuroscience 19, 1280–1285. https://doi.org/10.1038/nn.4382

Krajbich, I., Armel, C., Rangel, A., 2010. Visual fixations and the computation and comparison of value in simple choice. Nat. Neurosci. 13, 1292–1298. https://doi.org/10.1038/nn.2635

Kurniawan, I.T., Grueschow, M., Ruff, C.C., 2021. Anticipatory Energization Revealed by Pupil and Brain Activity Guides Human Effort-Based Decision Making. J. Neurosci. 41, 6328–6342. https://doi.org/10.1523/JNEUROSCI.3027-20.2021

Le Bouc, R., Borderies, N., Carle, G., Robriquet, C., Vinckier, F., Daunizeau, J., Azuar, C., Levy, R., Pessiglione, M., 2023. Effort avoidance as a core mechanism of apathy in frontotemporal dementia. Brain 146, 712–726. https://doi.org/10.1093/brain/awac427

Le Bouc, R., Borderies, N., Carle, G., Robriquet, C., Vinckier, F., Daunizeau, J., Azuar, C., Levy, R., Pessiglione, M., 2022. Effort avoidance as a core mechanism of apathy in frontotemporal dementia. Brain awac427. https://doi.org/10.1093/brain/awac427

Le Bouc, R., Rigoux, L., Schmidt, L., Degos, B., Welter, M.-L., Vidailhet, M., Daunizeau, J., Pessiglione, M., 2016. Computational Dissection of Dopamine Motor and Motivational Functions in Humans. Journal of Neuroscience 36, 6623–6633. https://doi.org/10.1523/JNEUROSCI.3078-15.2016

Lebreton, M., Abitbol, R., Daunizeau, J., Pessiglione, M., 2015. Automatic integration of confidence in the brain valuation signal. Nature Neuroscience 18, 1159–1167. https://doi.org/10.1038/nn.4064

Lebreton, M., Jorge, S., Michel, V., Thirion, B., Pessiglione, M., 2009. An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. Neuron 64, 431–439. https://doi.org/10.1016/j.neuron.2009.09.040

Lee, D., Daunizeau, J., 2021. Trading mental effort for confidence in the metacognitive control of value-based decision-making. eLife 10, e63282. https://doi.org/10.7554/eLife.63282

Lee, S., Yu, L.Q., Lerman, C., Kable, J.W., 2021. Subjective value, not a gridlike code, describes neural activity in ventromedial prefrontal cortex during value-based decision-making. Neuroimage 237, 118159. https://doi.org/10.1016/j.neuroimage.2021.118159

Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. Current Opinion in Neurobiology 22, 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W., 2010. Neural Representation of Subjective Value Under Risk and Ambiguity. Journal of Neurophysiology 103, 1036–1047. https://doi.org/10.1152/jn.00853.2009

Lim, S.-L., O'Doherty, J.P., Rangel, A., 2011. The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. Journal of Neuroscience 31, 13214–13223. https://doi.org/10.1523/JNEUROSCI.1246-11.2011

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

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., Pessiglione, M., 2020. Four core properties of the human brain valuation system demonstrated in intracranial signals. Nat Neurosci 23, 664–675. https://doi.org/10.1038/s41593-020-0615-9

Lopez-Persem, A., Domenech, P., Pessiglione, M., 2016. How prior preferences determine decision-making frames and biases in the human brain. Elife 5, e20317. https://doi.org/10.7554/eLife.20317

McGuire, J.T., Botvinick, M.M., 2010. Prefrontal cortex, cognitive control, and the registration of decision costs. Proc Natl Acad Sci USA 107, 7922–7926. https://doi.org/10.1073/pnas.0910662107

Meyniel, F., Sergent, C., Rigoux, L., Daunizeau, J., Pessiglione, M., 2013. Neurocomputational account of how the human brain decides when to have a break. Proceedings of the National Academy of Sciences 110, 2641–2646. https://doi.org/10.1073/pnas.1211925110

Palminteri, S., Boraud, T., Lafargue, G., Dubois, B., Pessiglione, M., 2009. Brain hemispheres selectively track the expected value of contralateral options. J. Neurosci. 29, 13465–13472. https://doi.org/10.1523/JNEUROSCI.1500-09.2009

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., Durr, A., Pessiglione, M., 2012. Critical Roles for

Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning. Neuron 76, 998–1009. https://doi.org/10.1016/j.neuron.2012.10.017
Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R.J., Frith, C.D., 2007. How the Brain Translates Money into Force: A Neuroimaging Study of Subliminal Motivation. Science 316, 904–906. https://doi.org/10.1126/science.1140459
Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442, 1042–1045. https://doi.org/10.1038/nature05051
Pessiglione, M., Vinckier, F., Bouret, S., Daunizeau, J., Le Bouc, R., 2018. Why not try harder? Computational approach to motivation deficits in neuro-psychiatric diseases. Brain 141, 629–650. https://doi.org/10.1093/brain/awx278
Peters, J., Büchel, C., 2010. Neural representations of subjective reward value. Behav Brain Res 213, 135–141. https://doi.org/10.1016/j.bbr.2010.04.031
Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions. Journal of Neuroscience 27, 9984–9988. https://doi.org/10.1523/JNEUROSCI.2131-07.2007
Pochon, J.-B., Riis, J., Sanfey, A.G., Nystrom, L.E., Cohen, J.D., 2008. Functional Imaging of Decision Conflict. Journal of Neuroscience 28, 3468–3473. https://doi.org/10.1523/JNEUROSCI.4195-07.2008
Qin, S., Marle, H.J.F. van, Hermans, E.J., Fernández, G., 2011. Subjective Sense of Memory Strength and the Objective Amount of Information Accurately Remembered Are Related to Distinct Neural Correlates at Encoding. J. Neurosci. 31, 8920–8927. https://doi.org/10.1523/JNEUROSCI.2587-10.2011
Richter, M., Gendolla, G.H.E., Wright, R.A., 2016. Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don't Know, in: Elliot, A.J. (Ed.), Advances in Motivation Science. Elsevier, pp. 149–186. https://doi.org/10.1016/bs.adms.2016.02.001
Rouault, M., Lebreton, M., Pessiglione, M., 2023. A shared brain system forming confidence judgment across cognitive domains. Cereb Cortex 33, 1426–1439. https://doi.org/10.1093/cercor/bhac146
Schmidt, L., Clery-Melin, M.-L., Lafargue, G., Valabregue, R., Fossati, P., Dubois, B., Pessiglione, M., 2009. Get Aroused and Be Stronger: Emotional Facilitation of Physical Effort in the Human Brain. Journal of Neuroscience 29, 9450–9457. https://doi.org/10.1523/JNEUROSCI.1951-09.2009
Schmidt, L., d'Arc, B.F., Lafargue, G., Galanaud, D., Czernecki, V., Grabli, D., Schüpbach, M., Hartmann, A., Lévy, R., Dubois, B., Pessiglione, M., 2008. Disconnecting force from money: effects of basal ganglia damage on incentive motivation. Brain 131, 1303–1310. https://doi.org/10.1093/brain/awn045
Schmidt, L., Lebreton, M., Cléry-Melin, M.-L., Daunizeau, J., Pessiglione, M., 2012. Neural Mechanisms Underlying Motivation of Mental Versus Physical Effort. PLoS Biology 10, e1001266. https://doi.org/10.1371/journal.pbio.1001266
Schonberg, T., Fox, C.R., Mumford, J.A., Congdon, E., Trepel, C., Poldrack, R.A., 2012. Decreasing ventromedial prefrontal cortex activity during sequential risk-taking: an FMRI investigation of the balloon analog risk task. Front Neurosci 6, 80. https://doi.org/10.3389/fnins.2012.00080
Seaman, K.L., Brooks, N., Karrer, T.M., Castrellon, J.J., Perkins, S.F., Dang, L.C., Hsu, M., Zald, D.H., Samanez-Larkin, G.R., 2018. Subjective value

representations during effort, probability and time discounting across adulthood. Social Cognitive and Affective Neuroscience 13, 449–459. https://doi.org/10.1093/scan/nsy021
Shapiro, A.D., Grafton, S.T., 2020. Subjective value then confidence in human ventromedial prefrontal cortex. PLoS ONE 15, e0225617. https://doi.org/10.1371/journal.pone.0225617
Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. Neuron 79, 217–240. https://doi.org/10.1016/j.neuron.2013.07.007
Shenhav, A., Karmarkar, U.R., 2019. Dissociable components of the reward circuit are involved in appraisal versus choice. Sci Rep 9, 1958. https://doi.org/10.1038/s41598-019-38927-7
Shenhav, A., Straccia, M.A., Cohen, J.D., Botvinick, M.M., 2014. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nature Neuroscience 17, 1249–1254. https://doi.org/10.1038/nn.3771
Silston, B., Wise, T., Qi, S., Sui, X., Dayan, P., Mobbs, D., 2021. Neural encoding of perceived patch value during competitive and hazardous virtual foraging. Nat Commun 12, 5478. https://doi.org/10.1038/s41467-021-25816-9
Skvortsova, V., Palminteri, S., Pessiglione, M., 2014. Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. Journal of Neuroscience 34, 15621–15630. https://doi.org/10.1523/JNEUROSCI.1350-14.2014
Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT Press.
Suzuki, S., Cross, L., O'Doherty, J.P., 2017. Elucidating the underlying components of food valuation in the human orbitofrontal cortex. Nat. Neurosci. 20, 1780–1786. https://doi.org/10.1038/s41593-017-0008-x
Ting, C.-C., Salem-Garcia, N., Palminteri, S., Engelmann, J.B., Lebreton, M., 2023. Neural and computational underpinnings of biased confidence in human reinforcement learning. Nat Commun 14, 6896. https://doi.org/10.1038/s41467-023-42589-5
Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The Neural Basis of Loss Aversion in Decision-Making Under Risk. Science 315, 515–518. https://doi.org/10.1126/science.1134239
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289. https://doi.org/10.1006/nimg.2001.0978
Van der Laan, L.N., De Ridder, D.T.D., Viergever, M.A., Smeets, P.A.M., 2012. Appearance Matters: Neural Correlates of Food Choice and Packaging Aesthetics. PLoS ONE 7, e41738. https://doi.org/10.1371/journal.pone.0041738
van der Wel, P., van Steenbergen, H., 2018. Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic Bulletin & Review 25, 2005–2015. https://doi.org/10.3758/s13423-018-1432-y
Vinckier, F., Jaffre, C., Gauthier, C., Smajda, S., Abdel-Ahad, P., Le Bouc, R., Daunizeau, J., Fefeu, M., Borderies, N., Plaze, M., Gaillard, R., Pessiglione, M., 2022. Elevated effort cost identified by computational modeling as a distinctive feature explaining multiple behaviors in patients with depression. Biological

Psychiatry: Cognitive Neuroscience and Neuroimaging. https://doi.org/10.1016/j.bpsc.2022.07.011
Volz, K.G., Schubotz, R.I., Cramon, D.Y. von, 2005. Variants of uncertainty in decision-making and their neural correlates. Brain Research Bulletin 67, 403–412. https://doi.org/10.1016/j.brainresbull.2005.06.011
Watkins, C.J.C.H., Dayan, P., 1992. Q-learning. Mach Learn 8, 279–292. https://doi.org/10.1007/BF00992698
Weissman, D.H., Carp, J., 2013. The congruency effect in the posterior medial frontal cortex is more consistent with time on task than with response conflict. PLoS One 8, e62405. https://doi.org/10.1371/journal.pone.0062405
Westbrook, A., Lamichhane, B., Braver, T., 2019. The Subjective Value of Cognitive Effort is Encoded by a Domain-General Valuation Network. The Journal of Neuroscience 3071–18. https://doi.org/10.1523/JNEUROSCI.3071-18.2019
Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., Pessiglione, M., 2022. A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. Curr Biol S0960-9822(22)01111–3. https://doi.org/10.1016/j.cub.2022.07.010
Wu, T., Dufford, A.J., Egan, L.J., Mackie, M.-A., Chen, Cong, Yuan, C., Chen, Chao, Li, X., Liu, X., Hof, P.R., Fan, J., 2018. Hick–Hyman Law is Mediated by the Cognitive Control Network in the Brain. Cerebral Cortex 28, 2267–2282. https://doi.org/10.1093/cercor/bhx127
Wunderlich, K., Rangel, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. Proc Natl Acad Sci U S A 106, 17199–17204. https://doi.org/10.1073/pnas.0901077106
Yeung, N., Cohen, J.D., Botvinick, M.M., 2011. Errors of interpretation and modeling: a reply to Grinband et al. Neuroimage 57, 316–319. https://doi.org/10.1016/j.neuroimage.2011.04.029
Zénon, A., Sidibé, M., Olivier, E., 2014. Pupil size variations correlate with physical effort perception. Frontiers in Behavioral Neuroscience 8. https://doi.org/10.3389/fnbeh.2014.00286
Zénon, A., Solopchuk, O., Pezzulo, G., 2019. An information-theoretic perspective on the costs of cognition. Neuropsychologia 123, 5–18. https://doi.org/10.1016/j.neuropsychologia.2018.09.013

Figures

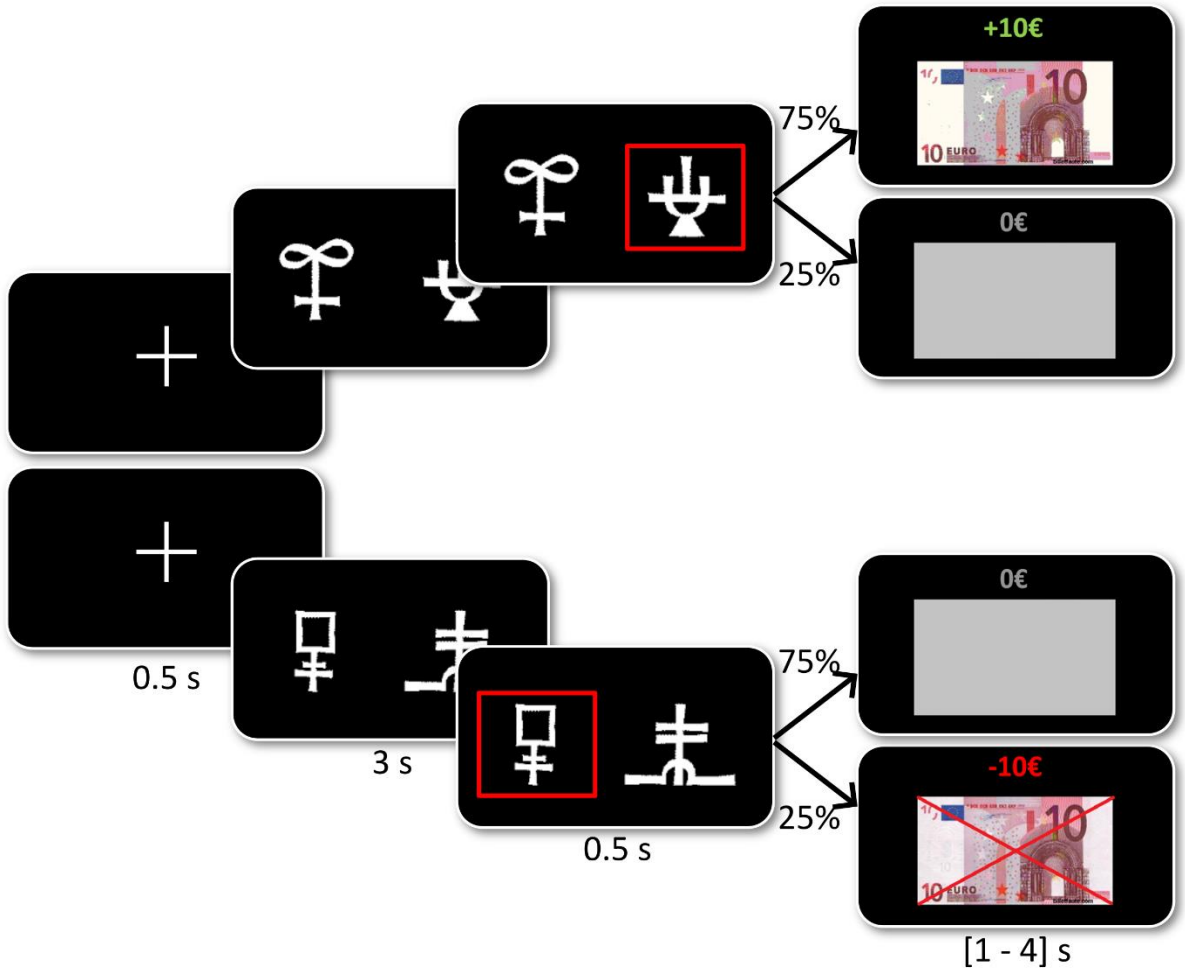


Figure 1. The learning task.

Screenshots presented in example trials are presented from left to right, with their duration in seconds indicated below. Every trial started with the display of a fixation cross. Participants chose between two visual cues and then observed the outcome of their choice. In a given learning session, there were only 6 cues always displayed in pairs. The gain pair provided either a reward or a neutral outcome, with different probabilities (25/75 or 75/25%) depending on which cue was chosen (top row). The loss pair provided either a neutral outcome or a loss outcome, with different probabilities (25/75 or 75/25%) depending on which cue was chosen (bottom row). In the examples, choices are correct (selected cues are associated with 75% probability of winning / not losing). The choice was recorded and shown on screen (with red frame) at the end of a fixed 3-s delay, depending on which button (left versus right) was being pressed. Outcomes (10€ banknote for gain, grey rectangle for zero, crossed 10€ banknote for loss) were last presented on screen with a jittered duration.

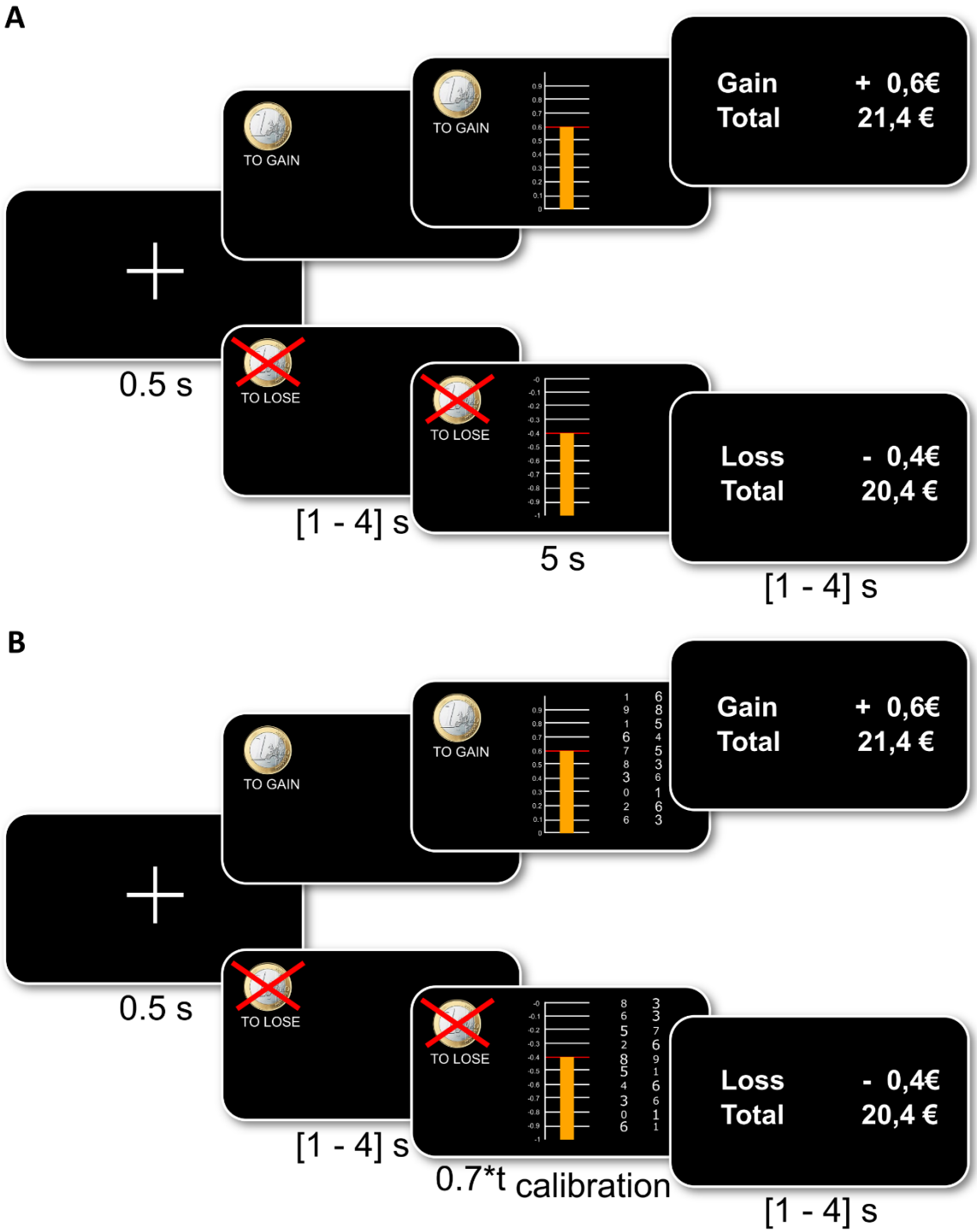


Figure 2. The performance tasks (grip and Stroop).

Screenshots presented in example trials are presented from left to right, with their duration in seconds indicated below. Every trial started with the display of a fixation cross. Then the incentive (among 6 possible amounts: 0.01, 0.20, 0.50, 1.00, 5.00, or 20.00€) was displayed with a cue for the condition (gain vs. loss trial). Real-time visual feedback on performance was provided as a bar that moved up within a scale graduated from 0 to maximum. In gain trials, participants received a fraction of the incentive proportional to their performance (e.g., 60 cents if they reached 60% of the scale for a 1€ incentive). In loss trials, participants avoided losing the fraction of the incentive proportional to their performance (i.e., they would only lose

40 cents in the example as they reached 60% of the scale). The money gained or lost in the current trial, and the cumulative total over all preceding trials, were shown in a last screen.

A] The grip task. Participants had to squeeze the handgrip as hard as they could. Performance was defined as the peak of the force pulse, expressed as a percentage of maximal force produced during calibration. The scale was adjusted such that the participant's maximal force corresponded to 75% of the incentive.

B] The Stroop task. Participants had to make as many numerical comparisons as they could, starting with the first pair of digits at the bottom. Performance was the number of correct numerical comparisons made within a predefined time window (70% of the time taken to complete all 10 comparisons during calibration). Note that in half the pair of digits, font size and numerical size were incongruent, creating a Stroop effect.

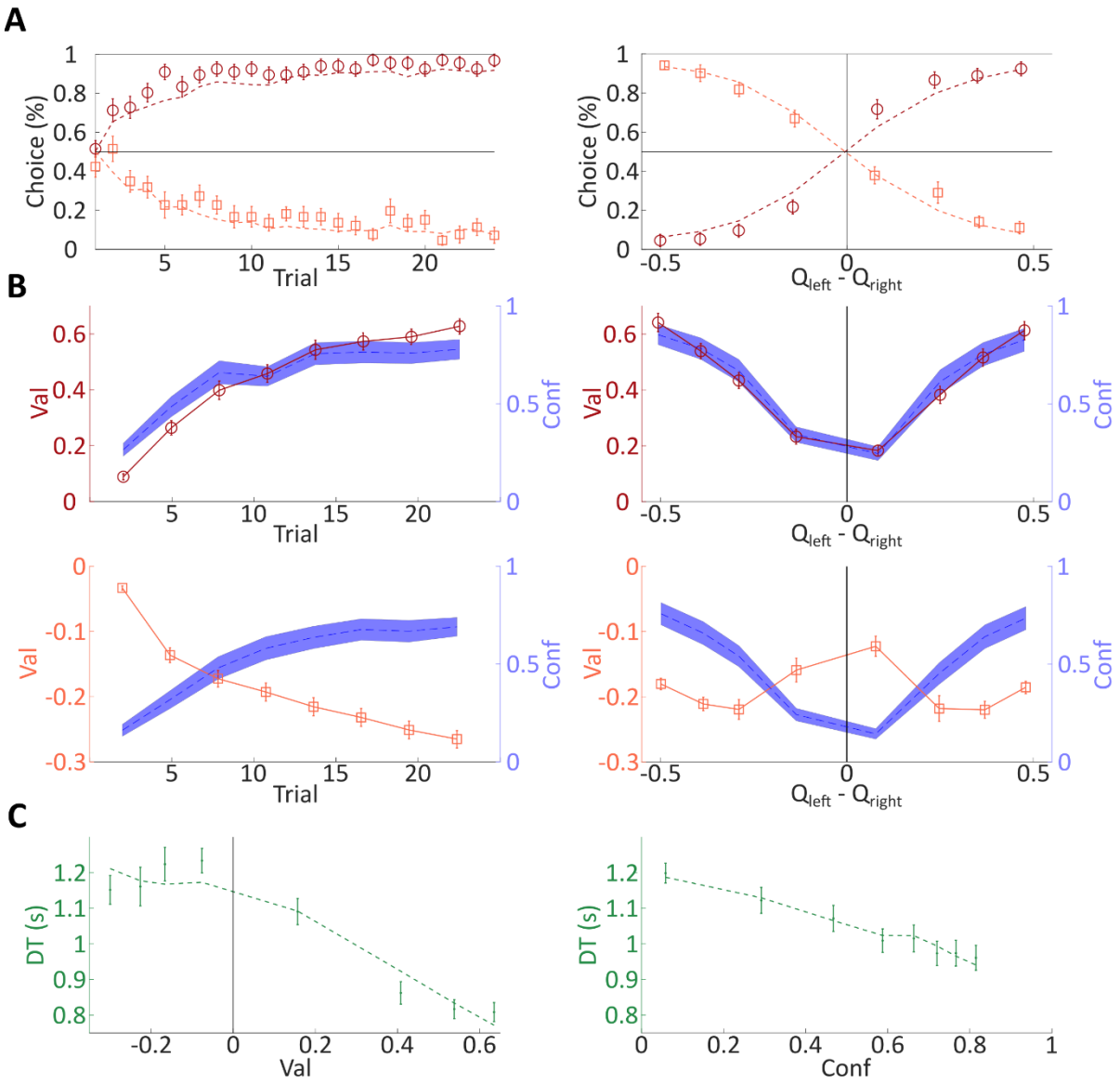


Figure 3. Behavior in the learning task.

A] Choice behavior. Left panel: percentage of correct choice (for gains) and incorrect choice (for losses) is shown as a function of the trial number within a session. Depending on the condition (i.e., the pair of cues), correct choice means selecting the cue with 75% chance of winning or 25% chance of losing 10€. Right panel: percentage of left choices (for gains) and right choices (for losses) is shown as a function of decision value (difference between left and right option values). The two conditions (gain and loss pairs) are plotted separately (dark red circles and light red squares, respectively). Choice data were fitted with a Q-learning model. Both observed and modeled choice data have been averaged across sessions and participants.

B] Val and Conf variables. Graphs show how our constructs for value (sum of option values weighted by their selection probabilities) and confidence (squared difference between selection probability and chance level) vary with trial number (left panels) and decision value (right panels), separately for the gain (top panels) and the loss (bottom panels) conditions.

C] Deliberation time (DT). The plots show how DT (time from option display to button press) varies with the Val (left panel) and Conf (right panel) constructs, pooling over gain and loss conditions. DT was fitted with a linear combination of Val and Conf variables.

In all figures, data points and error bars indicate the mean and standard error of the mean (SEM) across participants. Dotted lines indicate mean model fits.

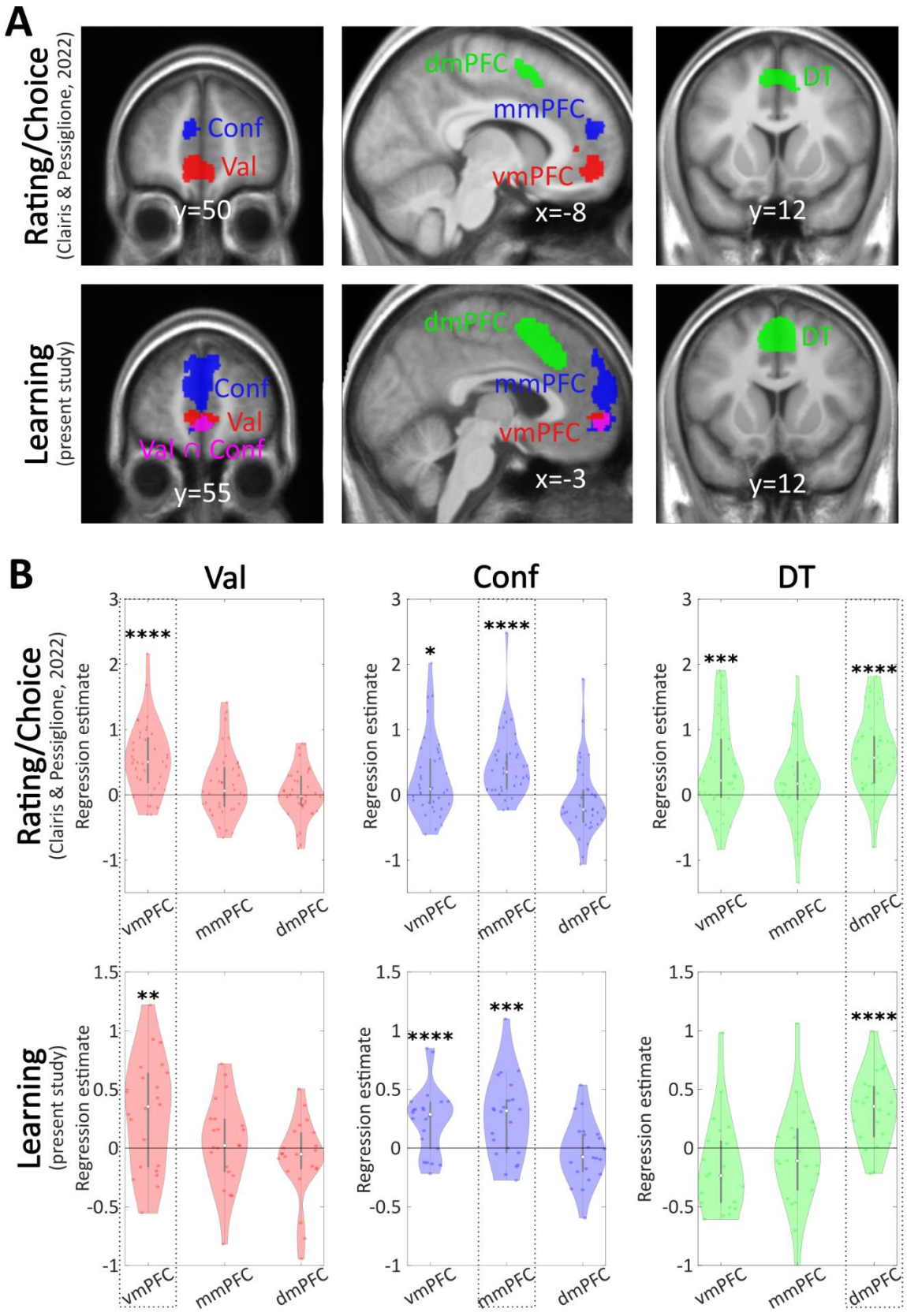


Figure 4. Neural activity during learning (versus rating and choice).

A] Activation maps. Colored voxels show group-level clusters within the medial prefrontal cortex mask (see Extended Data Figure 4-1) that were significantly associated with Val (red), Conf (blue) and DT (green) during rating and choice tasks (top panels) in our previous study

(Clairis and Pessiglione, 2022) and during the choice period of the learning task (bottom panels) in the present study. The overlap between Val and Conf clusters is shown in purple. The statistical thresholds were whole-brain FWE-corrected for multiple comparisons at the voxel level for rating / choice and at the cluster level for learning, due to a difference in statistical power between the two studies (n=38 vs. n=22). Clusters are overlaid on the average anatomical scan across participants of each study, normalized to the canonical Montreal Neurological Institute (MNI) template. They are labeled vmPFC, mmPFC and dmPFC for ventromedial, midmedial and dorsomedial prefrontal cortex. The 3 corresponding whole-brain activation tables for Val, Conf and DT can be found in Extended Data Tables 4-4, 4-5 and 4-6, respectively.

B] Region-of-interest analysis. In the previous study (top panels), regression estimates of Val, Conf and DT were extracted with a leave-one-out procedure to avoid double-dipping. In the current study (bottom panels), regression estimates of Val, Conf and DT were extracted from spheres positioned on the peaks of group-level significant clusters obtained in the previous study (Clairis and Pessiglione, 2022)**.** The three regressors were not orthogonalized in the main GLM used to fit neural activity during the choice period of the learning task. However, the same associations between the 3 ROI and the 3 variables are observed when the regressors are orthogonalized (see Extended Data Figure 4-2). Other combinations of option values (notably, chosen minus unchosen) have also been tested as possible definitions for Val (see Extended Data Figure 4-3). Dots are individual regression estimates. Error bars indicate the mean and standard error of the mean (SEM) across individuals. Stars denote significance of t-test against zero: ****p<0.001; ***p<0.005; **p<0.01; *p<0.05. Abbreviations: vmPFC, mmPFC and dmPFC designate ventromedial, midmedial and dorsomedial prefrontal cortex.

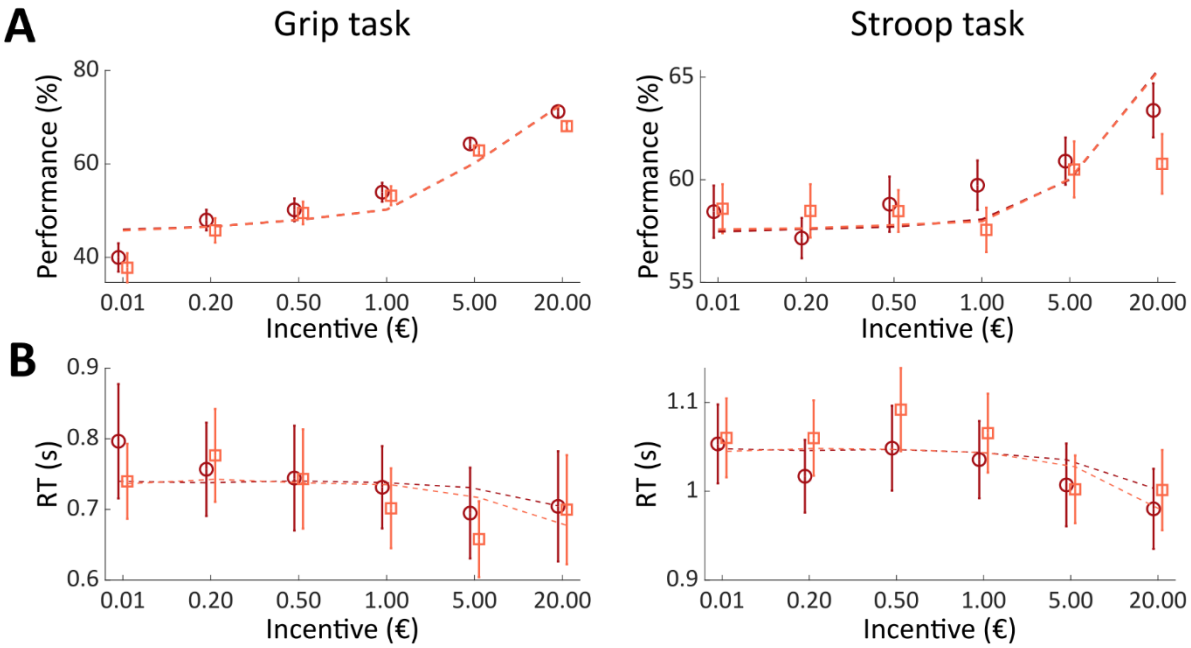


Figure 5. Behavior in motor and cognitive performance tasks.

A] Performance is the height reached within the scale, which is proportional to peak force in the grip task (left panel) and to the number of correct numerical comparisons in the Stoop task (right panel).

B] Reaction time is the latency at which force exceeded 1% of maximum in the grip task (left panel) and at which the first button press was made in the Stroop task (right panel).

Gain and loss conditions are shown in dark red circles and in light red squares, respectively. Error bars represent the mean and standard error of the mean (SEM) across participants for each incentive level. Dotted lines indicate the fit of the effort regulation model for performance and the fit of a multiple linear regression model for response time.

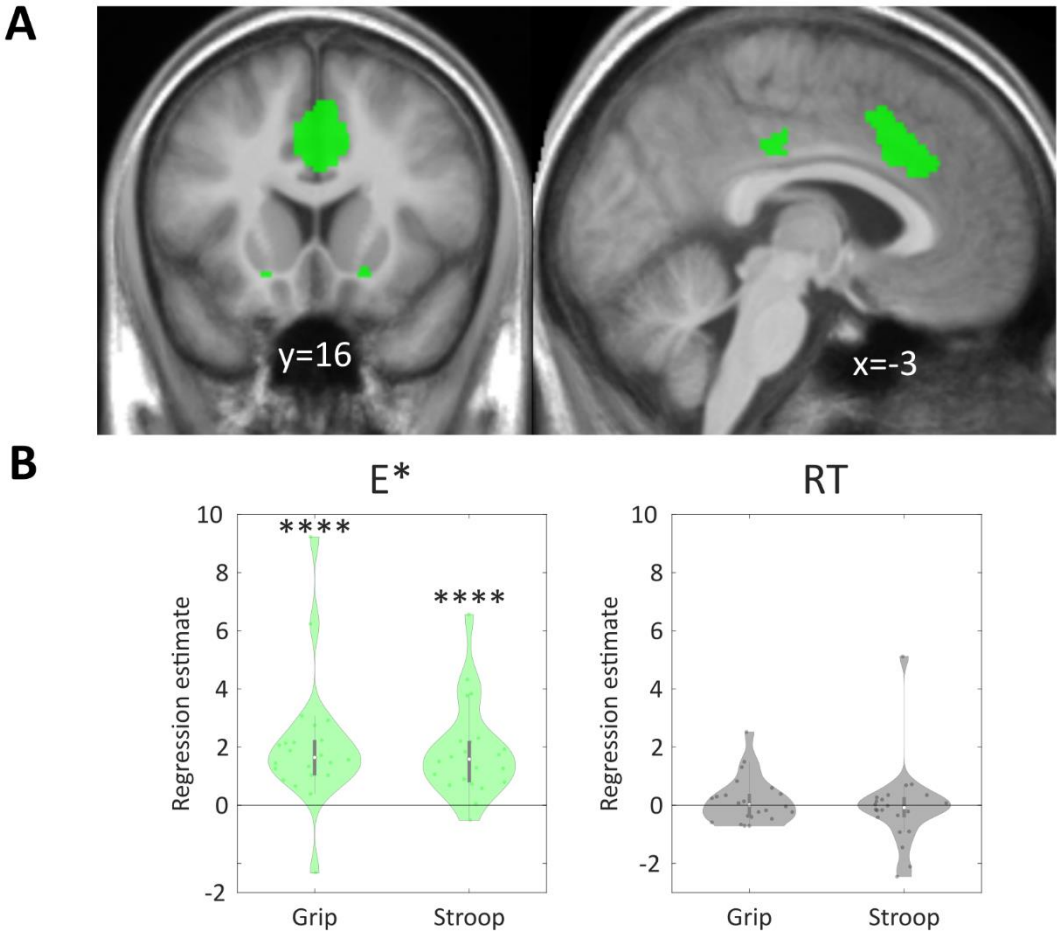


Figure 6. Neural activity during motor and cognitive performance tasks.

A] Activation maps. Colored voxels show group-level clusters within the medial prefrontal cortex mask (see Extended Data Figure 4-1) significantly associated ($p < 0.05$ after whole-brain family-wise error correction for multiple comparisons at the voxel level) with the theoretical effort exerted E* (generated with our computational model), in a conjunction between grip and Stroop tasks, at the time of incentive display. The corresponding whole-brain activation table is displayed in Extended Data Table 6-2. Clusters are overlaid on the average anatomical scan across participants, normalized to the canonical Montreal Neurological Institute (MNI) template. Sections are taken at the peak in the dmPFC, which stands for dorsomedial prefrontal cortex.

B] Regression estimates of effort exerted E* and reaction time RT were extracted from spheres positioned on the peaks of group-level significant clusters obtained in the previous study (Clairis and Pessiglione, 2022). The two regressors were not orthogonalized in the main GLM used to fit neural activity evoked by incentive display in both the grip and Stroop tasks. However, the main results are stable, even when the regressors are orthogonalized (see Extended Data Figure 6-1). On the opposite, the correlation between dmPFC and E* only holds when modeled during the incentive period but not if modeled during the performance period (see Extended Data Figure 6-1). Dots are individual regression estimates. Error bars indicate the mean and standard error of the mean (SEM) across individuals. Stars denote significance of t-test against zero: ****p<0.001; ***p<0.005; **p<0.01; *p<0.05. Abbreviations: vmPFC and dmPFC designate ventromedial and dorsomedial prefrontal cortex.

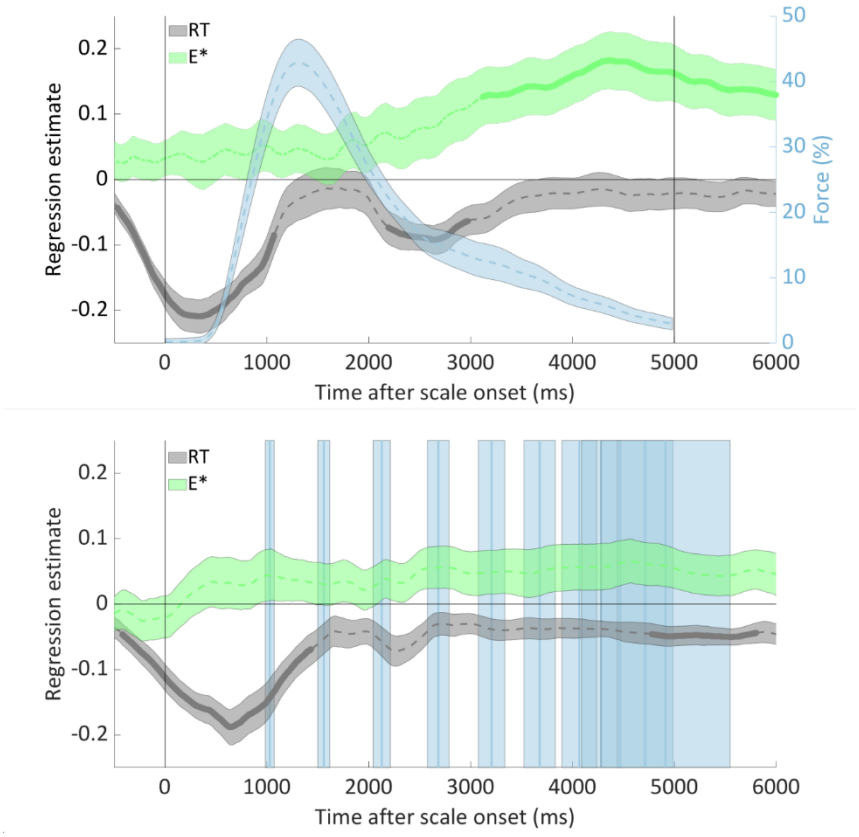


Figure 7. Pupil dilation in the motor and cognitive performance tasks.

Plots represent the time course of regression estimates, obtained with a GLM built to explain pupil size, on the grip and Stroop tasks (top and bottom graphs). The GLM included factors of no interest (session number and stimulus luminance, not shown) and variables of interest (theoretical effort E* and reaction time RT, shown in green and grey). Movements are indicated in blue (force produced in the grip task and button press in the Stroop task). Lines indicate means across participants and shaded areas the inter-participant standard error of the mean (SEM).

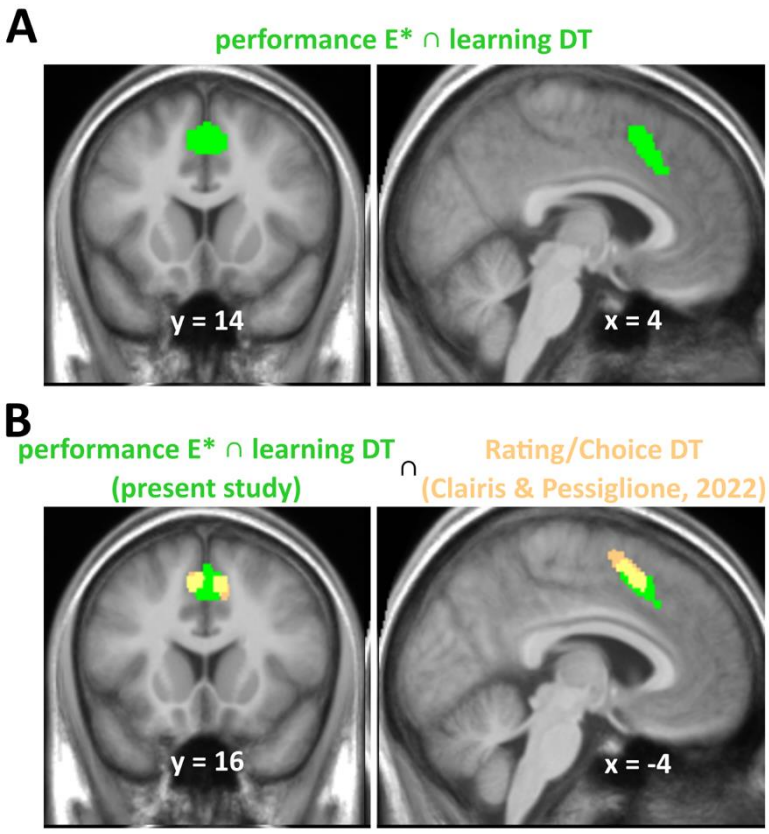
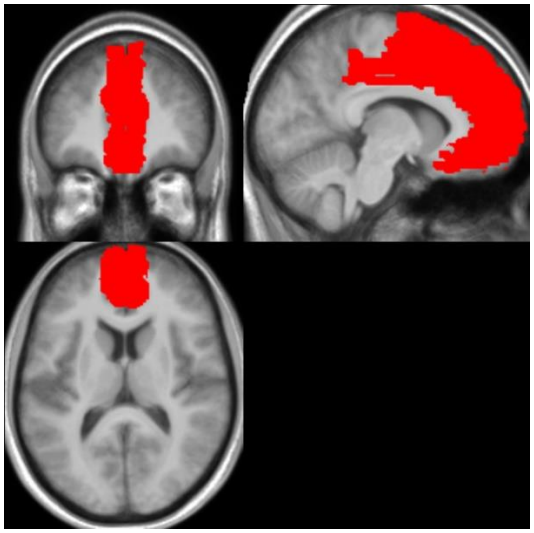


Figure 8. Global conjunction of effort-related activity across all tasks.

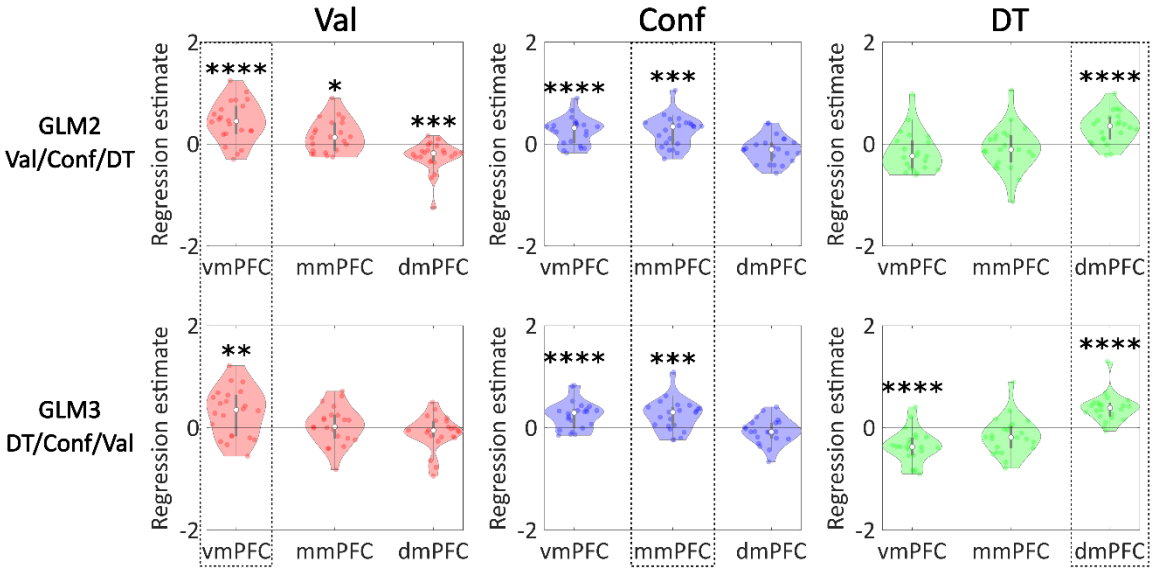
A] Colored voxels show the significant cluster (voxel-wise threshold: $p<0.05$ corrected for multiple comparisons), within the medial prefrontal cortex mask (see Extended Data Figure 4-1), resulting from the conjunction between two contrasts: DT in the learning task and E* in performance tasks. The corresponding whole-brain activation table is displayed in Extended Data Table 8-1.

B] Overlap (in yellow) between the significant cluster (in green) displayed in [A] and the dmPFC cluster (in orange) associated with rating/choice DT in our previous study (Clairis and Pessiglione, 2022). Clusters are overlaid on the average anatomical scan across participants, normalized to the canonical Montreal Neurological Institute (MNI) template.

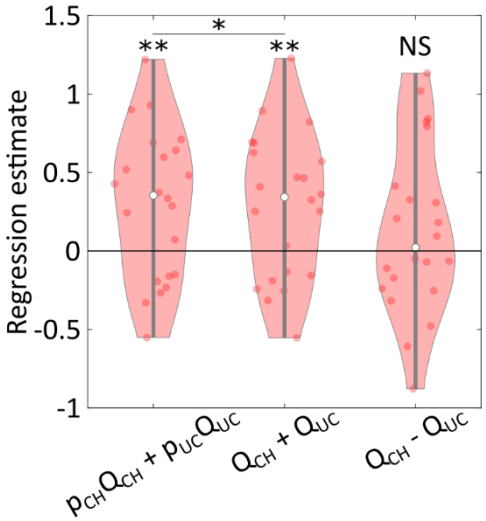
Extended Data



Extended Data Figure 4-1. Mask of the medial prefrontal cortex. This mask was built by merging brain regions (see Methods) of the AAL atlas parcellation (Tzourio-Mazoyer et al., 2002) and used for the display of Val, Conf and DT neural correlates in Figure 4 and 6. It is overlaid on the average anatomical scan of the 22 subjects included in the present study.



Extended Data Figure 4-2. Neural correlates of Val, Conf and DT in the learning task after orthogonalization. The regions of interest have been defined as spheres positioned on the peaks of group-level significant clusters obtained in the previous study (Clairis and Pessiglione, 2022). As in GLM1 (no orthogonalization), Val, Conf and DT are parametric modulators of choice-related activity, which have now been serially orthogonalized either in the Val/Conf/DT order (GLM2) or in the DT/Conf/Val order (GLM3). In all figures, dots are individuals, error bars show the mean and standard error of the mean (SEM) across individuals and stars indicate significance level of t-test against zero: ****p<0.001; ***p<0.005; **p<0.01; *p<0.05. Abbreviations: vmPFC, mmPFC and dmPFC designate ventromedial, midmedial and dorsomedial prefrontal cortex.



Extended Data Figure 4-3. Testing different associations between option values and vmPFC activity. Regression estimates were extracted from a sphere positioned on group-level activation peak observed with the Val regressor in our previous study (Clairis and Pessiglione, 2022). Graphs show regression estimates for the weighted sum of option values (GLM1), the straight sum (GLM4), and the difference between option values (GLM5). Dots are individuals, error bars show the mean and standard error of the mean (SEM) across individuals, and stars indicate significance level of t-test against zero: ****p<0.001; ***p<0.005; **p<0.01; *p<0.05.

Region	P cluster	Peak x	Peak y	Peak z	No. of Voxels
vmPFC	0.036	4	58	-4	255

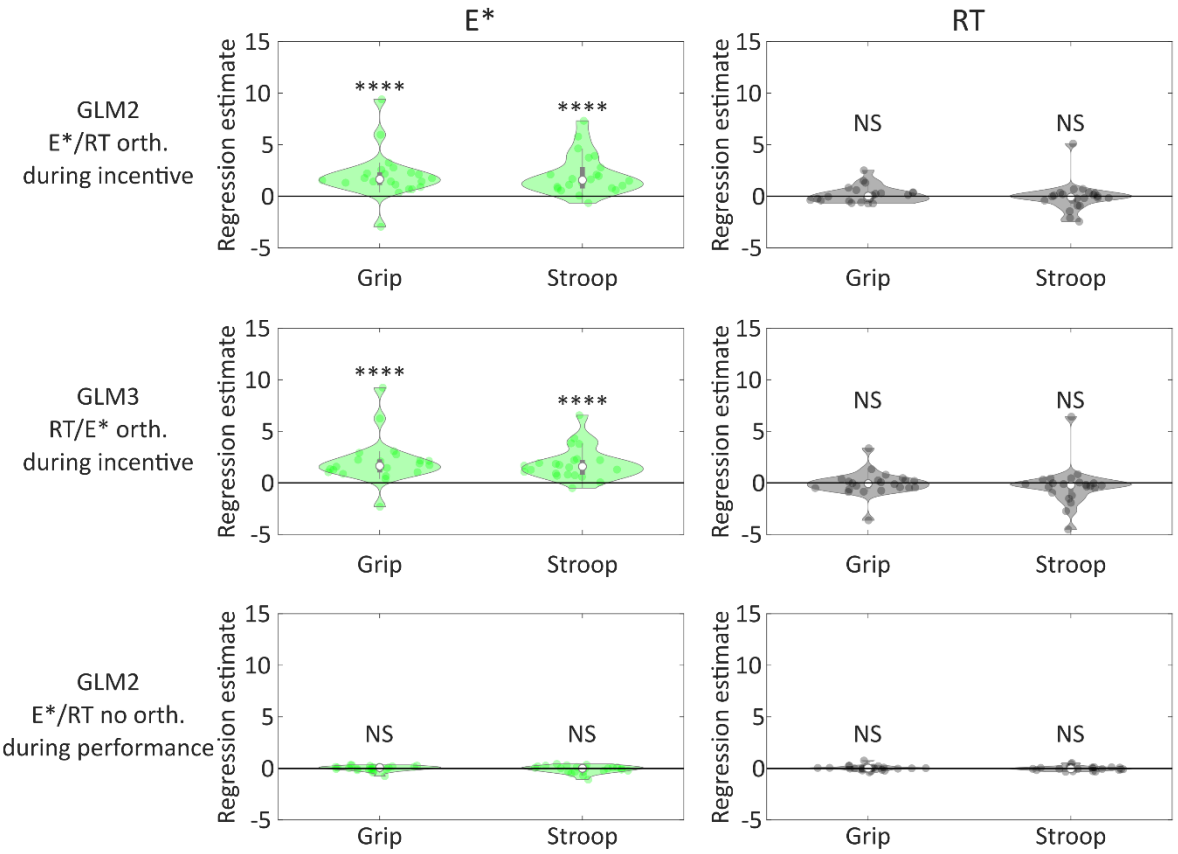
Extended Data Table 4-4: Whole-brain neural correlates of Val in the learning task (voxel-wise threshold: p<0.001 uncorrected; cluster-wise threshold: p<0.05 FWE corrected for multiple comparisons). The table shows the positive correlations in a t-test against zero; there was no negative correlation surviving the correction.

Region	P cluster	Peak x	Peak y	Peak z	No. of Voxels
Right precuneus	$2 \cdot 10^{-8}$	4	-46	54	1285
mmPFC	$3 \cdot 10^{-8}$	-4	64	16	1264
Left superior temporal gyrus	0.028	-50	-38	16	202
Left superior temporal gyrus	0.011	-34	10	-24	253

Extended Data Table 4-5: Whole-brain neural correlates of Conf in the learning task (voxel-wise threshold: p<0.001 uncorrected; cluster-wise threshold: p<0.05 FWE corrected for multiple comparisons). The table shows the positive correlations in a t-test against zero; there was no negative correlation surviving the correction.

Region	P cluster	Peak x	Peak y	Peak z	No. of Voxels
Left inferior occipital gyrus	$6 \cdot 10^{-12}$	-44	-58	-12	3049
dmPFC	$1 \cdot 10^{-10}$	-4	8	54	2558
Left superior parietal gyrus	$9 \cdot 10^{-8}$	-22	-60	46	1605
Left middle frontal gyrus	$1 \cdot 10^{-5}$	-38	36	36	1025
Right fusiform gyrus	$2 \cdot 10^{-5}$	42	-52	-16	957
Right middle frontal gyrus	$2 \cdot 10^{-4}$	28	40	34	729
Left anterior insula	$6 \cdot 10^{-4}$	-34	14	6	585
Right anterior insula	$9 \cdot 10^{-4}$	32	22	4	557
Right angular gyrus	0.006	26	-62	46	382

Extended Data Table 4-6: Whole-brain neural correlates of DT in the learning task (voxel-wise threshold: $p<0.001$ uncorrected; cluster-wise threshold: $p<0.05$ corrected for multiple comparisons). The table shows the positive correlations in a t-test against zero; there was no negative correlation surviving the correction.



Extended Data Figure 6-1: Neural correlates of E* and RT in the performance tasks, with orthogonalization or at a later time.

The regions of interest have been defined as spheres positioned on the peaks of group-level significant clusters obtained in the previous study (Clairis and Pessiglione, 2022). Compared to GLM1, where E* (left graphs) and RT (right graphs) are parametric modulators of incentive display in both grip and Stroop tasks with no orthogonalization, the parametric regressors have been orthogonalized in the E*/RT (GLM2, top row) or RT/E* (GLM3, middle row) order, or moved to the performance time window without orthogonalization (GLM4, bottom row). In all figures, dots are individuals, error bars show the mean and standard error of the mean (SEM) across individuals and stars indicate significance level of t-test against zero: ****p<0.001; ***p<0.005; **p<0.01; *p<0.05. Abbreviations: vmPFC and dmPFC designate ventromedial and dorsomedial prefrontal cortex.

Region	P cluster	Peak x	Peak y	Peak z	No. of Voxels
Right caudate	$8 \cdot 10^{-7}$	18	16	2	1154
Right calcarine	$6 \cdot 10^{-6}$	8	-90	2	864
dmPFC	$1 \cdot 10^{-5}$	-2	16	38	738
Right cerebellum	$2 \cdot 10^{-5}$	20	-52	-24	720
Left putamen	$3 \cdot 10^{-5}$	-14	10	0	621
middle cingulate cortex	$6 \cdot 10^{-4}$	6	-22	40	287
Right supramarginal gyrus	0.002	56	-44	24	180
Left cerebellum	0.002	-32	-58	-26	175
Ventral posterolateral thalamus	0.005	-14	-18	10	117
Right middle frontal gyrus	0.006	34	42	34	100
Right inferior frontal gyrus, opercular part	0.007	40	10	30	87
Right cuneus	0.015	18	-68	30	44

Extended Data Table 6-2: Whole-brain neural correlates of E* in a conjunction between grip and Stroop tasks (voxel-wise threshold: $p<0.05$ corrected for multiple comparisons, cluster-wise threshold: minimum of 33 voxels, i.e. the volume of the Gaussian kernel used for smoothing fMRI data). The table shows the positive correlations in a t-test against zero; there was no negative correlation surviving the correction.

Region	P cluster	Peak x	Peak y	Peak z	No. of Voxels
dmPFC	$6 \cdot 10^{-8}$	4	14	50	505
Right anterior insula	$7 \cdot 10^{-5}$	32	22	4	170
Left anterior insula	$2 \cdot 10^{-4}$	-34	18	6	125
Right middle frontal gyrus	0,002	38	42	32	61
Left superior Frontal gyrus	0,005	-24	-2	62	35

Extended Data Table 8-1: Whole-brain neural correlates of the conjunction between DT in the learning task and E* in performance tasks (voxel-wise threshold: $p<0.05$ corrected for multiple comparisons, cluster-wise threshold: minimum of 33 voxels, i.e. the volume of the Gaussian kernel used for smoothing fMRI data). The table shows the positive correlations in a t-test against zero; there was no negative correlation surviving the correction.