Value, confidence and deliberation: a functional partition of the medial prefrontal cortex across preference tasks

N. Clairis[1], M. Pessiglione[1]

[1]Motivation, Brain and Behavior team, Paris Brain Institute (ICM), Sorbonne University, Inserm, CNRS, Pitié-Salpêtrière Hospital, Paris, France

Abstract

Deciding about courses of action involves minimizing costs and maximizing benefits. Decision neuroscience studies have suggested a dissociation between the ventral and dorsal medial prefrontal cortex (vmPFC and dmPFC), which would provide estimates of goal value and action cost, respectively. However, this general idea of opponent reward and effort systems has been challenged by several studies that suggested alternative functional dissociations. These contradictions might reflect the diversity of tasks used to investigate the reward/effort tradeoff, and/or the confusion with a metacognitive tradeoff that minimizes the duration of deliberation while maximizing confidence in the behavioral response. Here, we used an original approach with the goal of identifying consistencies across several preference tasks, from likeability ratings to binary decisions involving both attribute integration and option comparison. FMRI results confirmed the vmPFC as a generic valuation system, its activity increasing with reward value and decreasing with effort cost. In contrast, more dorsal regions were not concerned with attributes of options but with metacognitive estimates, confidence level being computed in the mPFC and deliberation time in the dmPFC. Thus, assessing commonalities across preference tasks might help reaching a unified view of the neural mechanisms underlying the cost/benefit tradeoffs that drive human behavior.

Introduction

Cost/benefit arbitration is a general process involved in any decision that bears practical consequences. Standard decision theory assumes that selecting a course of action can be reduced to maximizing a net utility function, where expected benefits are discounted by expected costs. Numerous studies in decision neuroscience have implicated key regions of the medial prefrontal cortex in computing the net values of options during choice. While there is a general agreement for a functional dissociation between ventral and dorsal parts of the medial PFC (vmPFC, sometimes called medial OFC, versus dmPFC, sometimes called dACC), the specific roles of these subregions are still a matter of debate.

Some accounts insist on the opponency between costs and benefits (Pessiglione et al., 2018), implicating the vmPFC in the estimation of reward values and the dmPFC in the estimation of effort cost (Bartra et al., 2013; Botvinick et al., 2009; Clithero and Rangel, 2014; Klein-Flugge et al., 2016; Kurniawan et al., 2013; Lebreton et al., 2009; Skvortsova et al., 2014). A related view is that the vmPFC signals values in a space of goods, whereas the dmPFC encodes values in a space of actions. Although this view has received considerable empirical support (Padoa-Schioppa, 2011; Rangel and Hare, 2010; Rudebeck et al., 2008; Rushworth et al., 2012), it has also been contradicted by representations of effort cost found in vmPFC activity (Aridan et al., 2019; Arulpragasam et al., 2018; Gläscher et al., 2009; Hogan et al., 2019; Lopez-Gamundi et al., 2021; Seaman et al., 2018; Westbrook et al., 2019) and reward value in dmPFC activity (Cai and Padoa-Schioppa, 2012; Fouragnan et al., 2015; Klein-Flugge et al., 2016; Pisauro et al., 2017). The latter observations would rather suggest that both regions integrate the expected values of outcomes and the expected costs of actions.

Other accounts insist on the comparison between options that occurs during choice and suggest that both regions estimate decision values but in opposite fashion (Boorman et al., 2009; Hunt et al., 2012; Jocham et al., 2012; Wunderlich et al., 2009): the vmPFC would activate while the dmPFC would deactivate with value difference (chosen minus unchosen option value). This account is related to the views that the vmPFC activity signals the value of default behavior (or the reliability of current strategy), whereas the dmPFC signals the value of alternative/forgone options (or the need for exploration/switching). Again, although considerable evidence has been accumulated in favor of these views (Blanchard and Gershman, 2018; Domenech et al., 2020; Kolling et al., 2012; Lopez-Persem et al., 2016), contradictory results have also been reported. In particular, the correlation with chosen and unchosen values

is not always observed (Jocham et al., 2014; Lim et al., 2011; Lopez-Persem et al., 2016; Massar et al., 2015; Qin et al., 2011, 2011), which has raised the suspicion that other factors may better explain activity in these medial PFC regions. Indeed, the difference between chosen and unchosen option values is confounded with factors such as prior preference, choice confidence (or difficulty) and response time.

Here, we intend to take a step aside from these debates and propose a functional partition that would generalize beyond choice tasks. Indeed, neural correlates of values in the medial PFC have been found in many tasks that do not involve any choice, including likeability rating (De Martino et al., 2017; Lebreton et al., 2009; Lopez-Persem et al., 2020; Plassmann et al., 2010) or even passive viewing (Abitbol et al., 2015; Harvey et al., 2010; Kirk et al., 2011; Levy et al., 2011). We therefore reasoned that a general account of vmPFC and dmPFC functions in preference tasks should explain their pattern of activity during both rating and choice. This naturally excludes the idea that these regions are specifically wired to implement comparison between options, but could in principle be compatible with the opponency account, some regions being more sensitive to effort cost and others to reward value (during both rating and choice).

The new functional partition that we propose here is based on a metacognitive account (Lee and Daunizeau, 2021): the idea is that whatever the task, the brain invests effort in deliberation until it reaches a satisfactory level of confidence in the intended response. In other words, a second cost/benefit tradeoff would govern the meta-decision about when to make a response, the cost being the amount of time invested in deliberation and the benefit being the level of confidence attained. During this double cost/benefit arbitration, the brain would thus encode information at two levels: at the decisional level, it would evaluate the reward and effort values associated to options proposed for rating or choice, and at the metacognitive level, it would compute the expected confidence in the response and the required amount of deliberation.

The aim of the present study is to test whether this functional partition would better explain the activity of medial prefrontal regions. For this, we reversed the general logic of standard functional neuroimaging approach, which specifies the roles of brain regions with contrasts that isolate minimal differences between conditions. On the contrary, we intended to generalize our findings across various conditions and tasks, with the aim to reach more robust conclusions. Thus, we employed a series of preference tasks (also called 'value-based' tasks) that enable the investigation of 1) the reward value or effort cost of a single option, with likeability (or dislikeability) rating tasks, 2) the comparison between reward or effort attributes,

with one dimension – two options (1D-2O) choice tasks and 3) the integration of reward and effort attributes, with one option – two dimensions (1O-2D) choice tasks.

We then defined key variables of interest as the global stimulus value (Val), which increases with more appetitive reward and/or less aversive effort, the confidence in the response (Conf), which is higher for more extreme ratings and more likely choices, and deliberation time (DT), meaning duration of the effort invested in the valuation process so as to reach a satisfactory response. Importantly, the reward value and effort cost estimates did not imply any commitment to actual implementation in our tasks, as all options were hypothetical. On the contrary, confidence and deliberation time were variables associated to behavioral implementation, i.e. to the estimation of reward value and effort cost requested by the rating or choice task.

Our results support a functional partition within the medial prefrontal cortex, the global value of options (going up with reward and down with effort levels) being represented in vmPFC activity, while metacognitive variables (confidence in the response and duration of deliberation) in mPFC and dmPFC activity, respectively.

Results

Behavior

Participants (n=39 in total) first performed a series of ratings, divided into three fMRI sessions (Fig. 1A). Each session presented 72 items to be valuated one by one. Within a session, items were grouped into three blocks: one block with 24 reward items presented by text + image (R_{ti}), one block with 24 reward items presented by text only (R_t) and one block with 24 effort items presented by text only (E_t). The reason for varying the mode of presentation was to assess the generality of the neural valuation process across different inputs that require more or less imagination, according to previous study (Lebreton et al., 2013). For reward, participants were asked to rate how much they would like it, should they be given the item immediately after the experiment. Symmetrically, the instruction for effort was to rate how much they would dislike it, should they be requested to exert it immediately after the experiment. We included both food and non-food (goodies) reward items, and both mental and physical effort items. There was no number on the scale, just labels on endpoints, and ratings were pseudo-continuous, from ‘I would (dis-)like it not at all’ to ‘enormously’. Thus, the left endpoint corresponded to indifference and the right endpoint to extreme attraction or extreme aversion (Fig. 1).

The z-scored rating was taken as a proxy for stimulus value (Val) in this task, while the square of z-score rating was taken as a proxy for response confidence (Conf), following on previous studies (Lebreton et al., 2015; Lopez-Persem et al., 2020). Note that Val and Conf were orthogonal variables by construction. Deliberation time (DT) was defined as the time between item onset and the first button press used to move the cursor along the scale. DT was regressed against a linear model that included Val and Conf proxies (Fig. 2), in addition to nuisance factors (such as jitter duration, stimulus luminance, text length and trial index, see methods). Irrespective of stimulus type, we found a significant effect of both value (R_{ti}: β_{Val} = -0.208 ± 0.023, p = $4 \cdot 10^{-11}$; R_t: β_{Val} = -0.169 ± 0.019, p = $6 \cdot 10^{-11}$; E_t: β_{Val} = 0.261 ± 0.028, p = $2 \cdot 10^{-11}$) and confidence (R_{ti}: β_{Conf} = -0.174 ± 0.025, p = $3 \cdot 10^{-8}$; R_t: β_{Conf} = -0.188 ± 0.028, p = $7 \cdot 10^{-8}$; E_t: β_{Conf} = -0.130 ± 0.040; p = 0.0024). Thus, participants were faster to provide their rating when the item was more appetitive (or less aversive) and when they were more confident (going towards the extremes of the rating scale). Among the nuisance factors, we observed effects of jitter duration, stimulus luminance and text length, which were therefore included as regressors of no interest in subsequent analyses. However, there was no significant effect of trial index, which discards a possible contamination of DT by habituation or fatigue.

Then participants performed a series of binary choices, either 1D-2O choices or 1O-2D choices. The choice tasks were always performed after the rating tasks because the ratings were used to control the difficulty of choices (i.e., the difference in value between the two options). In the 1D-2O choice task (Fig. 1B), participants were asked to select the reward they would prefer to receive at the end of the experiment, if they were offered one of two options, or the effort they would prefer to exert, if they were forced to implement one of two options. Thus, the two options always pertained to the same dimension (reward or effort), and even to the same sub-category (food or good for reward, mental or physical for effort), to avoid shortcut of deliberation by general preference. The mode of presentation (text or image) was also the same for the two options, to avoid biasing the choice by a difference in salience. To obtain a same number of trials as in the rating task, each item was presented twice, for a total of 72 choices per stimulus type (R_{ti}, R_t, E_t) distributed over three fMRI sessions. Within a session, items were grouped into three blocks: one block with 24 choices between reward items presented with text + image (R_{ti}), one block with 24 choices between reward items presented with text only (R_t) and one block with 24 choices between effort items presented with text only (E_t). In the 1O-2D choice task (Fig. 1C), participants were asked whether they would be willing to exert an effort in order to obtain a reward, at the end of the experiment. Only items described with text were

retained for this task (since there was no picture for effort items), each item again appearing twice, for a total of 144 choices divided into three fMRI sessions of 48 trials each.

The 1D-2O choice task was meant to assess value comparison between the two options, within a same dimension. The decision value (DV) in this task was defined as the difference in (dis-)likeability rating between the two options. We checked with a logistic regression (Fig. 2) that DV was a significant predictor of choices, irrespective of stimulus type (R_{ti}: $\beta_{DV} = 3.383 \pm 0.273$, $p = 7 \cdot 10^{-15}$; R_t: $\beta_{DV} = 2.669 \pm 0.155$, $p = 2 \cdot 10^{-19}$; E_t: $\beta_{DV} = -2.278 \pm 0.157$, $p = 4 \cdot 10^{-17}$). The 1O-2D choice task was meant to assess value integration across two dimensions, for a single option. The decision value (or net value) in this task was defined as a linear combination of reward and effort ratings. Note that it would make no sense to fit an effort discounting function here, because such function is meant to capture the mapping from objective effort levels to subjective effort estimates, which we directly collected (with dislikeability ratings). We checked with a logistic regression that both reward and effort ratings were significant predictors of choice in this task ($\beta_R = 1.502 \pm 0.085$, $p = 6 \cdot 10^{-20}$; $\beta_E = -1.117 \pm 0.079$, $p = 1 \cdot 10^{-16}$).

To analyze DT (time between stimulus onset and button press) in choice tasks, we defined proxies for stimulus value and response confidence, as we did for the rating task. Stimulus value (Val) was defined as the addition of the likeability ratings assigned to the two stimuli on screen. In the 1D-2O choice task, this is simply the sum of the two item ratings. In the 1O-2D choice task, this is a weighted sum (with a scaling factor to adjust the unit of reward and effort ratings). In both cases, choice probability was calculated with the logistic regression model (softmax function of decision value). Response confidence (Conf) was defined, by analogy to the rating task, as the square of the difference between choice probability and mean choice rate. Linear regression (Fig. 2) showed that DT decreased with value in the 1D-2O choice task (R_{ti}: $\beta_{Val} = -0.061 \pm 0.010$, $p = 3 \cdot 10^{-7}$; R_t: $\beta_{Val} = -0.061 \pm 0.010$, $p = 3 \cdot 10^{-7}$; E_t: $\beta_{Val} = 0.048 \pm 0.013$, $p = 8 \cdot 10^{-4}$), albeit not in the 1O-2D choice task ($\beta_{Val} = 0.033 \pm 0.024$, $p = 0.172$). DT also decreased with confidence in both the 1D-2O choice task (R_{ti}: $\beta_{Conf} = -1.738 \pm 0.201$, $p = 2 \cdot 10^{-10}$; R_t: $\beta_{Conf} = -1.975 \pm 0.184$, $p = 4 \cdot 10^{-13}$; E_t: $\beta_{Conf} = -1.731 \pm 0.221$, $p = 2 \cdot 10^{-9}$) and the 1O-2D choice task ($\beta_{Conf} = -1.148 \pm 0.145$, $p = 1 \cdot 10^{-9}$). Thus, the relationship between DT and the two other factors of interest was similar in rating and choice tasks: participants were faster when the options were more appetitive (or less aversive) and when they were more confident (because of a strong preference for one response or the other).

Because we did not measure confidence in the present study, we verified that our proxy could predict confidence ratings in separate datasets. Note that this proxy has already been

validated for likeability rating tasks used in previous studies (De Martino et al., 2017; Lebreton et al., 2015; Lopez-Persem et al., 2020), a result that is reproduced in supplementary material (Fig. S1A). To test whether the same proxy could also predict confidence in choice tasks, we used another dataset from a published study (Lee and Daunizeau, 2020). In this study (Fig. S1B), participants provided confidence ratings about having selected the best option in binary choices (between food items presented two by two). Our confidence proxy could significantly predict confidence judgments not only in the rating task ($\beta = 0.074 \pm 0.027$; $p = 0.008$) but also in the choice task ($\beta = 4.157 \pm 0.715$; $p = 8 \cdot 10^{-7}$).

Neural activity

The aim of fMRI data analysis was to dissociate the first-level variables related to option attributes (reward and effort estimates) from the second-level variables related to metacognition (confidence and deliberation) across value-based tasks (rating and choice). To assess whether these variables can be dissociated on the basis of existing literature, we conducted a meta-analysis of fMRI studies using Neurosynth platform (Fig. 3A) with value, confidence and effort as keywords. Results show that the three keywords are associated to similar activation patterns, with clusters in both vmPFC and dmPFC. To better dissociate the neural correlates of these constructs in our dataset, we built a general linear model where stimulus onset events were modulated by our three parameters of interest - Val, Conf and DT (defined as in the behavioral data analysis). Nuisance parameters that were found to influence DT (jitter duration, stimulus luminance, text length) were also included as modulators of stimulus onset events, before the variables of interest. Note that by construction, Val and Conf regressors were orthogonal in our design, but DT was partially correlated with all the other regressors. To avoid confounding the interpretation, we therefore employed serial orthogonalization. Thus, the variables of interest were orthogonalized with respect to nuisance factors, and DT was made orthogonal to all other regressors, including Val and Conf.

After correction for multiple comparisons at the voxel level, we found only three significant clusters in the prefrontal cortex (Fig. 3B): Val was signaled in vmPFC activity (Table S1), Conf in mPFC activity (Table S2) and DT in dmPFC activity (Table S3). With a more lenient threshold (correction at the cluster level), we observed significant association with Val in other brain regions, such as the ventral striatum (vS), posterior cingulate cortex (pCC) and primary visual cortex (V1). Note that vS and pCC are standard components of the brain valuation system, whereas V1 activation is likely to be an artifact of gaze position on the rating

scale, as it was not observed in the choice tasks. Consistently, positive correlation with Val was found in right V1 activity, and negative correlation in left V1 activity (a pattern that was not observed with other clusters). To provide a more exhaustive depiction, we examined the distribution of regression estimates below statistical thresholds, along a path going from vmPFC to dmPFC within a medial plane (Fig. 3C). Results show that the three associations did not correspond to separate clusters (as was suggested by thresholded maps) but to gradients peaking at different positions along the line.

To assess whether the triple association between variables and clusters of interest was robust, we conducted a number of additional analyses using variants of the main GLM (Fig. S2). The same three clusters were significantly associated with the Val, Conf and DT regressors when 1) removing serial orthogonalization such that regressors could compete for variance and 2) replacing stick functions by boxcar functions extending from stimulus onset to behavioral response (showing a modulation of dmPFC activity by DT in amplitude and not just duration). In addition, we tested the triple association using different fMRI acquisition sequences in participants of the pilot study (n=15). The fMRI sessions acquired with multiband acceleration sequences (see methods) were not included in the main analysis, since they were not directly comparable to those using our standard EPI sequence. We separately regressed fMRI activity recorded during these sessions against our main GLM, and observed similar trends in this independent dataset. Due to a three times smaller sample, activations did not pass whole-brain corrected thresholds. However, using group-level significant clusters (from the main dataset) as ROI, we observed significant associations of Val and DT with vmPFC and dmPFC, respectively (Figure S3).

We further analyzed the relationship between computational variables and activity in the three medial prefrontal regions of interest (ROI) with post-hoc t-tests on regression estimates. To avoid any double-dipping issue, we used a leave-one-out procedure, such that clusters were defined from group-level analyses including all subjects but the one in whom regression estimates were extracted. We first verified that the three main associations were not driven by any particular task (Fig. 3D). Indeed, regression estimates were significant in both rating and choice tasks, more specifically for Val in vmPFC activity (rating: $\beta_{Val} = 0.688 \pm 0.130$, $p = 6 \cdot 10^{-6}$; choice: $\beta_{Val} = 0.469 \pm 0.099$, $p = 3 \cdot 10^{-5}$), for Conf in mPFC activity (rating: $\beta_{Conf} = 0.753 \pm 0.113$, $p = 8 \cdot 10^{-8}$; choice: $\beta_{Conf} = 0.314 \pm 0.103$, $p = 0.004$) and for DT in dmPFC activity (rating: $\beta_{DT} = 0.389 \pm 0.108$, $p = 9 \cdot 10^{-4}$; choice: $\beta_{DT} = 0.744 \pm 0.111$, $p = 7 \cdot 10^{-8}$). Note that our point was to generalize the associations across different tasks - comparing between

tasks would be meaningless because tasks were not designed to be comparable (any possible significant contrast could be due to many differences of no interest).

We also investigated whether each cluster of interest was better associated with the corresponding variable (across tasks), again using a leave-one-out procedure to avoid double dipping (Fig. 3D): Val was better reflected in vmPFC activity ($\beta_{Val/vmPFC} > \beta_{Val/mPFC}$: $p = 9{\cdot}10^{-8}$; $\beta_{Val/vmPFC} > \beta_{Val/dmPFC}$: $p = 4{\cdot}10^{-7}$), Conf in mPFC activity ($\beta_{Conf/mPFC} > \beta_{Conf/vmPFC}$: $p = 0.0043$; $\beta_{Conf/mPFC} > \beta_{Conf/dmPFC}$: $p = 3{\cdot}10^{-7}$) and DT in dmPFC activity ($\beta_{DT/dmPFC} > \beta_{DT/vmPFC}$: $p = 0.066$; $\beta_{DT/dmPFC} > \beta_{DT/mPFC}$: $p = 7{\cdot}10^{-4}$). However, the fact that vmPFC, mPFC and dmPFC better reflected Val, Conf and DT, respectively, does not imply that these regions were not affected by the other variables. In particular, vmPFC activity was also associated with Conf and DT, ($\beta_{Conf} = 0.257 \pm 0.097$, $p = 0.012$; $\beta_{DT} = 0.395 \pm 0.111$, $p = 0.001$), even if it was dominated by Val-related activity.

We next looked for further generalization of the valuation signal, not solely across tasks but also across stimuli. In the main analysis, fMRI time series were regressed against a GLM that separated stimulus types (Rti, Rt and Et) into different onset regressors, each modulated by corresponding ratings. Instead of testing the average regression estimates across stimulus categories, we tested regression estimates obtained for each category, separately (Fig. S4A and S4B). Regression estimates (extracted using leave-one-out procedure across rating and choice tasks) show that vmPFC activity was positively related to the subjective value of reward items, whether or not they are presented with an image (Rti: $\beta_{Val} = 0.487 \pm 0.134$, $p = 8{\cdot}10^{-4}$; Rt: $\beta_{Val} = 0.610 \pm 0.134$, $p = 5{\cdot}10^{-5}$), and negatively correlated to the subjective cost of effort items (Et: $\beta_{Val} = -0.349 \pm 0.134$, $p = 0.017$). Thus, the association between Val and vmPFC activity was independent of the presentation mode, and integrated costs as well as benefits.

On a different note, we questioned the validity of our Val proxy to capture value-related activity in choice tasks. Again, the reason for summing stimulus values in choice tasks instead of taking the difference between chosen and unchosen option values, as is often done, was that we wanted a proxy that could generalize to rating tasks, in which there is no notion of difference, since there is only one stimulus on screen. Nevertheless, we wondered whether vmPFC activity in choice tasks would be better captured by the difference ($V_c - V_{uc}$) than by the sum ($V_c + V_{uc}$). To test this, we simply replaced our partition (Val / Conf / DT) by V_c and V_{uc} regressors, and fitted the GLM to fMRI activity recorded during choice tasks only (Fig. S4C). The two regression estimates, extracted from the Val cluster in the main analysis, were significantly positive ($\beta_{Vc} = 0.415 \pm 0.115$, $p = 9{\cdot}10^{-4}$; $\beta_{Vuc} = 0.290 \pm 0.069$, $p = 2{\cdot}10^{-4}$), with no significant difference between the two ($p = 0.357$), therefore showing no evidence for a representation of

the difference. We completed this simple analysis by a comparison using Bayesian Model Selection at the group level, between two variants of the main GLM where Val was replaced by either the sum ($V_c + V_{uc}$) or the difference ($V_c - V_{uc}$), competing to explain choice-related activity in a vmPFC ROI defined from the literature (to avoid non-independence issues). Although not formally conclusive, the comparison showed that exceedance probability was in favor of the sum model (Fig. S4D), thus validating our Val proxy as most relevant to capture vmPFC activity, even during choices. Another advantage of this Val proxy is being orthogonal to confidence, whereas the difference between option values is not. The consequence is that the neural correlates of Conf were unaffected by introducing the Val regressor, or by serial orthogonalization (Fig. S1).

Importantly, no consistent association with reward value or effort cost was observed in putative opponent brain regions such as the dmPFC, which was instead systematically reflecting DT. Thus, it appeared that dmPFC activity reflected the metacognitive effort cost invested in the ongoing task (deliberation about the response) rather than the effort cost attached to the option on valuation. Importantly, the association with DT was observed despite the fact that DT was orthogonalized to both value and confidence, suggesting that the dmPFC represents the effort invested above and beyond that induced by the difficulty of value-based judgment or decision. The parametric modulation by DT was also obtained when dmPFC activation was fitted with a boxcar function extending from stimulus response (Fig. S2), suggesting a modulation in amplitude beyond prolonged activity.

However, DT is a very indirect proxy for the effort invested in solving the task, and could be affected by many other factors (such as distraction or mind-wandering). We therefore investigated the relationship between brain activity and another proxy that has been repeatedly related to effort: pupil size. Neural activity was extracted in each ROI by fitting a GLM containing one event (stimulus onset) per trial. Then pupil size at each time point was regressed across trials against a GLM that contained nuisance factors (luminance, jitter duration, text length), variables of interest (Val, Conf, DT) and neural activity (vmPFC, mPFC, dmPFC).

A positive association between pupil size and dmPFC activity was observed in both rating and choice tasks (Fig. 4), about one second before the response. This association was not an artifact of the trial being prolonged (and therefore of the response to luminance being cut at different durations), since it was observed both when locking time courses on stimulus onset and on motor response (button press). Finally, it was specific to the dmPFC ROI, and observed even if dmPFC was made independent (through serial orthogonalization) to all other variables (notably Val, Conf and DT). Thus, the association between dmPFC and pupil size was observed

above and beyond DT and factors that could affect DT. In contrast, there was no consistent association between vmPFC and pupil size before the response, suggesting that the correlates of DT observed in vmPFC were not related to effort but to some other factors affecting DT, such as mind-wandering.

Discussion

In this study, we investigated the neural correlates of variables that are common to different tasks involving valuation of stimuli and expression of preferences. We observed a functional partition within the medial prefrontal cortex: stimulus value, response confidence and deliberation time were best reflected in vmPFC, mPFC and dmPFC activity, respectively. These associations between regions and variables were consistent across rating and choice tasks, whether they involved likeability judgment, attribute integration or option comparison. They suggest that reward value and effort cost attached to choice options are integrated in a same brain region (vmPFC), while metacognitive variables such as response confidence and deliberation time are represented in distinct brain regions (mPFC and dmPFC).

Our results confirm the role attributed to the vmPFC as a generic valuation system (Bartra et al., 2013; Levy and Glimcher, 2012). The subjective value of reward items was reflected in vmPFC activity irrespective of the category (food versus goods), as was reported in many studies (Abitbol et al., 2015; Chib et al., 2009; Lebreton et al., 2009; Lopez-Persem et al., 2020). Also, vmPFC value signals were observed whether or not reward items were presented with images, suggesting that they can be extracted from both direct perceptual input or from text-based imagination which was shown to recruit episodic memory systems (Lebreton et al., 2013). Critically, our results show that the vmPFC also reflects the effort cost (whether mental or physical) attached to potential courses of actions. Therefore, they challenge previous suggestions that the vmPFC is involved in stimulus valuation, independently of action costs (Padoa-Schioppa, 2011; Pessiglione et al., 2018; Rangel and Hare, 2010). They rather suggest that the vmPFC might compute a net value, integrating reward benefit and effort cost, so as to prescribe whether or not an action is worth engaging. This idea is in line with early demonstrations that the vmPFC integrates costs such as potential loss or delay in reward delivery (Hare et al., 2009; Kable and Glimcher, 2007; Talmi et al., 2009; Tom et al., 2007) and with recent mounting evidence that it also integrates effort-related costs (Aridan et al., 2019; Hogan et al., 2019; Lopez-Gamundi et al., 2021; Westbrook et al., 2019). Consistent with these studies, we observed that vmPFC activity increases with potential benefit and decreases with potential cost, which is compatible with the idea of net value computation.

The other medial prefrontal clusters (mPFC and dmPFC) were not affected by reward values or effort costs attached to choice options, but by metacognitive variables related to providing a response in preference tasks, i.e. confidence and deliberation. We have previously argued that confidence is what participants maximize when performing these tasks (Lebreton et al., 2015). Following on instructions, they intend to give their best judgment, or make the

right choice, even if these responses have no material consequence. In a sense, confidence can be conceived as a value, but this value would be attached to the response, not to the options. By construction, our proxy for confidence, defined as absolute (squared) deviation from the mean response, was orthogonal to stimulus value, defined as the sum of (positive) likeability of reward items and (negative) likeability of effort items. This confidence proxy was found to elicit similar neural correlates as direct confidence ratings (De Martino et al., 2017; Lopez-Persem et al., 2020).

Confidence was the only variable significantly associated to mPFC activity, but was also reflected in vmPFC activity, as previously reported (Chua et al., 2006; De Martino et al., 2013; Gherman and Philiastides, 2018). Indeed, the addition of value and confidence signals in the vmPFC is a pattern that has been already observed in both fMRI and iEEG activity (Lebreton et al., 2015; Lopez-Persem et al., 2020). It has been argued that, in a binary choice, this pattern may denote the transition from the sum of option values to the difference between chosen and unchosen values (Hunt et al., 2012), this difference being globally positive and hence related to confidence. However, this explanation would not hold for the correlates of confidence observed outside choice tasks, for instance in likeability rating tasks. In addition, our analyses show that even in choice tasks, regression estimates are positive and similar for chosen and unchosen values, such that their sum captures most of the inter-trial variance in vmPFC activity. An interpretation compatible with these results (Pessiglione and Daunizeau, 2021) would be that, as a generic valuation device, the vmPFC computes both the value of options and the value of responses (i.e., confidence). The dissociation observed here, with value being better related to vmPFC activity and confidence better related to mPFC activity, is consistent with a general ventro-dorsal gradient from value to confidence representation that has been described previously (De Martino et al., 2017). On the contrary, dmPFC activity tended to decrease with confidence, but this trend did not survive significance threshold.

The variable that was robustly associated with dmPFC activity was deliberation time. This variable was not orthogonal to the others, since it decreased with both stimulus value and response confidence. The link between deliberation time and stimulus value might arise from an appetitive Pavlovian reflex, as suggested in previous studies (Oudiette et al., 2019; Shadmehr et al., 2019), since there was no reason to go faster when valuating better rewards, or slower when valuating worse efforts, in our design. The link between deliberation time and response confidence might relate to the difficulty of the task (Bang and Fleming, 2018; Kiani et al., 2014; Lee and Daunizeau, 2021; Pleskac and Busemeyer, 2010; Ratcliff et al., 2016; Yeung and Summerfield, 2012), i.e. the uncertainty about which rating or choice best reflects subjective

preference. In our analyses, deliberation time was post-hoc orthogonalized with respect to the other variables, meaning that the association with dmPFC activity was observed above and beyond the variance explained by stimulus value and response confidence.

This association alone would not yield a clear-cut interpretation, since many factors may affect response time. However, the systematic link observed between trial-wise dmPFC activation and the increase in pupil size just before the response hints that this association might reflect the cognitive effort invested in the task. Indeed, pupil size has been linked to the intensity of not only physical effort, such as handgrip squeeze (Zénon et al., 2014) but also mental effort, such as focusing attention to resolve conflict or overcome task difficulty (Alnaes et al., 2014; Kahneman and Beatty, 1966; van der Wel and van Steenbergen, 2018). By contrast, we did not observe this systematic link with pupil size during deliberation with vmPFC activity. The link between vmPFC and deliberation time might therefore reflect other sources of variance, such as mind-wandering (being slower because of some off-task periods).

Our dmPFC ROI overlaps with clusters that have been labeled dorsal anterior cingulate cortex, or sometimes pre-supplementary motor area, in previous studies (Kamiński et al., 2017; Kolling et al., 2016; Shenhav et al., 2013). The association with deliberation time is compatible with a role attributed to this region in conflict monitoring, or in signaling the need to exert cognitive control (Bhanji et al., 2010; Botvinick et al., 2001; Durston, 2003; Kerns et al., 2004; Shenhav et al., 2013; Sohn et al., 2007). This functional role would also be consistent with the negative association between dmPFC activity and our confidence proxy, which is opposite to a conflict or ambiguity signal. In binary choices, longer deliberation when options are close is often captured as a shallower drift rate in accumulation-to-bound models (Lopez-Persem et al., 2016; Ratcliff et al., 2016; Steverson et al., 2019). However, our results call for a more general theory, because the link between dmPFC activity and deliberation time was also observed in rating tasks.

To recapitulate, we have teased apart the neural correlates of likeability, confidence and deliberation in the medial prefrontal cortex, which have been confused in previous fMRI studies, as shown by meta-analytic maps. The key distinction operated here is perhaps between effort as an attribute of choice option and effort as a resource allocated to the preference task, or in other words, between valuation applied to effort (implicating the vmPFC) and effort invested in valuation (implicating the dmPFC). This dissociation is consistent with the idea that the vmPFC anticipates the aversive value of a potential effort, while the dmPFC represents the intensity of effort when it must be exerted (Hogan et al., 2019). At a metacognitive level, our results could be interpreted in the frame of a resource allocation model, where the effort or time

invested in the deliberation is meant to increase confidence in the response, whether a rating or a choice (Lee and Daunizeau, 2021). Yet it remains possible that dmPFC activation is not about demand for deliberation effort but about estimating the amount of deliberation already done, on the request of other brain systems. It could even be that dmPFC activation corresponds to the aversive feeling induced by effort exertion, without any implication in metacognitive regulation.

Crucially, we have shown here that the three associations hold across rating and choice tasks, and thus cannot be captured by models narrowly applied to the case of binary choice. However, this approach (looking for robust associations across tasks) also bears limitations. Notably, our design would not allow comparing between conditions, as is traditionally done in neuroimaging studies. One may want for instance to compare between tasks and test whether brain regions are more involved in one or the other, but this would be confounded by several factors, such as the order (choice tasks being performed after rating tasks). A significant contrast would not be interpretable anyway, because there is more than one minimal difference between tasks. Thus, the aim to generalize the role of brain regions across tasks carries the inherent drawback of a limited specificity, but also the promises of a more robust understanding of anatomo-functional relationships. We hope this study will pave the way to further investigations following a similar approach, assessing a same concept across several tasks in a single study, instead of splitting tasks over separate reports, with likely inconsistent conclusions.

Methods

Subjects

In total, 40 right-handed volunteers participated in this fMRI study, which was approved by the Pitié-Salpêtrière Hospital local ethics committee. Participants were recruited through the RISC (Relais d'Information en Sciences de la Cognition) online platform (https://www.risc.cnrs.fr/) and signed inform consent prior to participation in the study. All participants were screened for the use of psychotropic medications and drugs, history of psychiatric and neurologic disorders, and traumatic brain injury. One participant was excluded from all analyses because of a clear misunderstanding about task instructions, leaving n=39 participants for behavioral data analysis (22 females / 17 males, aged 25.4±4.1 years). Another participant was excluded from the fMRI analysis due to excessive movement inside the scanner (>3mm within-session per direction). Eleven additional participants were excluded from pupil size analysis, due to poor signal detection in at least one of the sessions (leaving a total of n=27 participants for pupil analysis).

All participants gave informed consent and were paid a fixed amount for their participation. The 15 first subjects were paid 60€ and the 25 other subjects were paid 75€. The difference in payoff corresponds to a difference in scanning protocols, although all participants performed the same tasks. The pilot protocol (n=15) aimed at comparing fMRI data acquisition sequences: regular EPI, EPI with multiband acceleration, EPI with multiband acceleration + multi-echo acquisition. The main protocol (n=25) aimed at addressing the neurocognitive question of interest with the best acquisition sequence. For this main protocol, we kept the regular EPI sequence for all sessions, as we saw no clear advantage for multiband acceleration or multi-echo acquisition in basic contrast images. Therefore, the analyses only include fMRI data using regular EPI acquisition (three sessions for the pilot protocol, all nine sessions for the main protocol).

Behavioral tasks

All tasks were programmed using the Psychtoolbox (Brainard, 1997) Psychtoolbox-3 running in Matlab (The MathWorks, Inc., Version 2012). Participants were given a 4-button box (fORP 932, Current Designs Inc, Philadelphia, USA) placed under their right hand to provide their responses. Stimuli were projected on a computer screen, their luminance being estimated using standard function of red-green-blue composition (0.299·red + 0.587·green + 0.114·blue, see http://www.w3.org/TR/AERT#color-contrast). Stimuli comprised 144 reward items (72 food and 72 goods) and 72 effort items (36 mental and 36 physical). Half the reward items

were presented with text only (R_t items), and the other half was presented with both text and image (R_{ti} items). All effort items were only described with text (E_t). For each task, fMRI sessions were preceded by a short training (not included in the analysis), for participants to familiarize with the sort of items they would have to valuate and with the button pad they would use to express their preferences.

Participants all started with a (dis-)likeability rating task (Fig. 1A), performed during the first three fMRI sessions, each divided into three 24-trial blocks corresponding to the three stimulus type (R_{ti}, R_t, E_t). The order of blocks within a session was counterbalanced across participants. The items were presented one by one, and participants rated them by moving a cursor along a visual analog scale. They used their index and middle fingers to press buttons corresponding to left and right movements, and validated the final position of the cursor by pressing a third button, which triggered the new trial. The initial position of the cursor, at the beginning of each trial, was randomly placed between 25 and 75% of the 0-100 rating scale. There was no mark on the scale, giving the impression of a continuous rating, although it was in practice discretized into 100 steps. The left and right extremes of the scale were labeled “I would not care” and “I would like it enormously” for reward items, “I would not mind” and “I would dislike it enormously” for effort items. Note that both reward and effort scales included indifference at one extremity, such that the two scales could form a continuum of increasing likeability from very aversive effort to very appetitive reward. In any case, the situations to be rated were hypothetical: the question was about how much they would like the reward (should it be given to them at the end of the experiment) and how much they would dislike the effort (should it be imposed to them at the end of the experiment). Should the timeout (10 s in rating tasks and 5s in choice tasks) be reached, the message ‘too slow’ would have been displayed on screen and the trial repeated later, but this remained exceptional.

After the three rating sessions, participants performed a series of binary choices. The 1D-2O left/right choice task (Fig. 1B) involved expressing a preference between two options of a same dimension, presented on the left and right of the screen. The two options were items presented in the rating task, drawn from the same category, regarding both the presentation mode (R_{ti} vs R_{ti}, R_t vs R_t, E_t vs E_t) and type of stimulus (food vs. food, goods vs. goods, mental vs mental, physical vs physical). Each item was presented twice, following two intermixed pairing schedules: one varied the mean rating (i.e., stimulus value) while controlling for distance (i.e., decision value or choice difficulty), whereas the other varied the distance in rating while controlling the mean. Participants selected the reward they would most like to obtain, or the effort they would least dislike to exert, by pressing the left or right button with their middle

or index finger. The chosen option was then highlighted with a red frame, so participants could check that their choice was correctly recorded. The fMRI sessions devoted to the 1D-2O choice task included three 24-trial blocks presenting the three types of options (R_{ti}, R_t, E_t), the order of blocks being counterbalanced across participants.

Then participants performed the 1O-2D yes/no choice task (Fig. 1C), which involved deciding whether to accept exerting a given effort in order to get a given reward. Thus, every trial proposed one option combining two dimensions (one R_t and one E_t item). Each item was presented twice, following two intermixed pairing schedules: one associating more pleasant reward with more painful effort (thus controlling for decision value or choice difficulty), the other associating more pleasant reward with less painful effort (thus varying choice difficulty). The mean net value was also balanced across fMRI sessions. Participants selected their response by pressing the button corresponding to 'yes' or 'no' with their index or middle finger. The left/right position of yes/no responses was counterbalanced across trials. To give participants a feedback on their choice, the selected option was highlighted with a red frame. The three fMRI sessions devoted to the 1O-2D choice task contained 48 trials each.

Behavioral data analysis

All data were analyzed using Matlab 2017a (The MathWorks, Inc., USA).

Choices were fitted with logistic regression models of decision value, with intercept and slope parameters.

For 1D-2O choices, the model was:

$$P(\,choice = left\,) = \frac{1}{1 + e^{\beta_0 + \beta_1 \cdot \Delta V}}$$

where ΔV is the decision value, i.e. the difference in likeability rating between left and right options, while β_0 and β_1 are the intercept and slope parameters capturing potential bias and choice consistency (inverse temperature).

For 1O-2D choices, the model was:

$$P(\,choice = accept\,) = \frac{1}{1 + e^{\beta_0 + \beta_R \cdot V_R + \beta_E \cdot V_E}}$$

where V_R and V_E are the likeability ratings provided for the reward and effort items. Thus, the decision value (or net value) here is a weighted sum of reward and effort likeability (one being positive and the other negative), the parameter weights β_R and β_E serving as both scaling factors and inverse temperature).

Deliberation time (DT) was defined across tasks as the time between stimulus onset and first button press. Trial-wise variations in DT were fitted with linear regression models,

including a session-specific intercept, nuisance factors - fixation cross display duration (Jitter), stimulus luminance (Lum), text length in number of words (Length) - and factors of interest - stimulus value (Val), response confidence (Conf). Thus, the model was:

$$DT = \beta_{s1} + \beta_{s2} + \beta_{s3} + \beta_{jit} \cdot Jitter + \beta_{lum}.Lum + \beta_{len}.Length + \beta_{val}.Val + \beta_{conf}.Conf$$

The Val and Conf regressors represented stimulus value (reward likeability plus effort likeability) and response confidence (squared distance from mean response). They were adapted to the task, as follows:

	Rating task	**1D-2O choice task**	**1O-2D choice task**
Val	V	$V_{left} + V_{right}$	$\beta_R \cdot V_R + \beta_E \cdot V_E$
Conf	$[V\text{-}mean(V)]^2$	$[P_{left} - mean(P_{left})]^2$	$[P_{accept} - mean(P_{accept})]^2$

In each case, P is the probability generated using the logistic model and V is either reward or effort likeability, provided by z-scored individual ratings. Note that, by construction, V is positive for reward items (which are liked) and negative for effort items (which are disliked).

fMRI data acquisition

Functional and structural brain imaging data was collected using a Siemens Magnetom Prisma 3-T scanner equipped with a Siemens 64 channel Head/Neck coil. Structural T1-weighted images were coregistered 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 BOLD contrast using the following parameters: repetition time TR = 2.01 seconds, echo time TE = 25 ms, flip angle = 78°, number of slices = 37, slice thickness = 2.5 mm, field of view = 200 mm. A tilted-plane acquisition sequence was used to optimize sensitivity to BOLD signal in the orbitofrontal cortex (44). Note that the number of volumes per session was not predefined, because all responses were self-paced. Volume acquisition was just stopped when the task was completed.

Most subjects (n=25) performed nine fMRI sessions (three per task) using this standard EPI sequence. The pilot subgroup (n=15) also performed nine fMRI sessions, but the fMRI data acquisition sequences were alternated between standard EPI, EPI with multi-band acceleration factor (TR = 1.20 s; TE = 25 ms; flip angle = 66°; number of slices = 44; slice thickness = 2.5 mm; acceleration factor = 2) and EPI with multi-band acceleration factor + multi-echo (TR = 1.28 s; TE = 11.00 ms and 29.89 ms; flip angle = 69°; number of slices = 44; slice thickness = 2.5 mm; acceleration factor = 2). The order of fMRI sequences was counterbalanced across participants. Preliminary analyses of basic contrast images were done using the pilot dataset to select the best acquisition sequence. As there was no clear benefit with the multi-band and

multi-echo add-ons, we retained the standard EPI for the main experiment. Results obtained with the other fMRI sequences are shown as supplementary material (Fig. S3).

fMRI data analysis

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

Preprocessed data were analyzed with a standard general linear model (GLM) approach at the first (individual) level and then tested for significance at the second (group) level. All GLM included the six movement regressors generated during realignment of successive scans. In our main GLM, stimulus onset was modeled by a stick function, modulated by the following regressors: 1) fixation cross duration, 2) luminance, 3) text length, 4) Val, 5) Conf, 6) DT. The first three were nuisance factors that were found to significantly impact DT in the linear regression analysis. The regressors of interest (Val, Conf and DT) were defined as explained in the behavioral data analysis section. The different blocks of the rating and 1D-2O choice tasks (presenting reward as text + image, reward as text and effort as text) were modeled in separate regressors. All regressors of interest were z-scored and convolved with the canonical hemodynamic response function and its first temporal derivative. All parametric modulators were serially orthogonalized. At the second level, correlates of Val, Conf and DT were obtained with contrasts tested across tasks of corresponding regression estimates against zero. Note that likeability ratings obtained for effort items were negative in all regressors (meaning that they can only decrease stimulus value).

Several alternative GLM were built to test variants of the main GLM. GLM2 was identical to GLM1 except that orthogonalization was removed such that all native regressors could compete to explain variance in fMRI time series. GLM3 was identical to GLM1, except that instead of a stick function, stimulus onsets were modeled with a boxcar function modeling periods from stimulus onset to first button press. Three additional GLM were built to further explore the choice tasks. In GLM4, the Val regressor (sum of option values) was replaced by the difference between option values ($V_c - V_{uc}$) for the two choice tasks. This GLM served to perform a group-level Bayesian model comparison to test which value regressor (sum or difference) best explains the fMRI time series during choice tasks. In GLM5, Conf and DT were removed and Val was replaced by two separate regressors for the chosen and unchosen

option values (V_c and V_{uc}). This GLM was used to test whether regressor estimates for chosen and unchosen values had the same sign (as in a sum) or opposite signs (as in a difference). In GLM6, reward and effort values were split in two separate regressors for all tasks (including the 1O-2D choice task). The purpose of this GLM was to distinguish between neural correlates of reward value and effort cost in brain valuation regions. Finally, a last GLM was built with one event per trial, modeled with a stick function, at the time of stimulus onset, with the aim to extract trial-by-trial activity levels in regions of interest, which then served as regressors to explain pupil size data (see next section).

Regions of interest (ROI) were defined as clusters in group-level statistical maps that survived significance threshold of $p < 0.05$ after family-wise error correction for multiple comparisons at the voxel level. To avoid double dipping (Kriegeskorte et al., 2009) in statistical tests, regression estimates were extracted from ROI re-defined for each participant through a leave-one-out procedure. Regarding Bayesian Model Selection, to avoid biasing the comparison in favor of one or the other GLM, an independent ROI was defined as the conjunction between the positive minus negative value contrast in a published meta-analysis (Bartra et al., 2013) and the bilateral medial orbitofrontal cortex region from the AAL atlas (Tzourio-Mazoyer et al., 2002). Additionally, we defined twelve 8-mm radius spherical ROI in the medial wall to illustrate the distribution of regression estimates for Val, Conf and DT. Parameter estimates were extracted from each voxel within these ROI and then averaged across voxels.

Meta-analysis of fMRI studies

The meta-analytic maps were extracted from the online platform Neurosynth (https://www.neurosynth.org/), using the keywords "value" (470 studies), "confidence" (79 studies) and "effort" (204 studies) for "uniformity test", which displays brain regions that are consistently activated across papers mentioning the target keyword. Each map was binarized to visualize clusters surviving a significance threshold of $p < 0.01$ after false discovery rate (FDR) correction for multiple comparisons.

Pupil size

Pupil diameter was recorded at a sampling rate of 1000Hz, 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. A cubicle interpolation was performed to compensate for any period of time when the pupil signal was lost due to blinking. The pupil size time series were subsequently band-pass filtered (1/128 to 1 Hz) and z-scored per session.

Within-trial variations in pupil size was baseline-corrected (by removing the mean signal over the 200 ms preceding stimulus onset) and time-locked either to stimulus onset or button press. Then trial-wise variations in pupil size were fitted with a linear regression model that included nuisance factors (an intercept per block, jitter duration, stimulus luminance and text length), variables of interest (Val, Conf and DT defined as in the behavioral data analysis section) and neural activity (extracted from vmPFC, mPFC and dmPFC ROI clusters). Within-trial individual time series of regression estimates were then smoothed using a 100ms kernel. Group-level significant time clusters were identified after correction for multiple comparisons estimated according to random field theory, using the RFT_GLM_contrast.m function of the VBA toolbox (available at http://mbb-team.github.io/VBA-toolbox/).

Acknowledgments

We would like to thank the CENIR (research neuroimaging center) staff for their help in fMRI data acquisition (particularly Stéphane Lehéricy, Romain Valabrègue and Mathieu Santin for the optimization of scanning sequences), Chen Hu for assistance in data collection, Jules Brochard for assistance in data analysis, Fabien Vinckier and Jean Daunizeau for insightful comments. The study was funded by a research grant from the “Fondation pour la Recherche Médicale” and by the “Investissements d’avenir” program (ANR-10- IBHU-0003).

Bibliography

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., and Pessiglione, M. (2015). Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans. Journal of Neuroscience *35*, 2308–2320.

Alnaes, D., Sneve, M.H., Espeseth, T., Endestad, T., van de Pavert, S.H.P., and Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. Journal of Vision *14*, 1–1.

Aridan, N., Malecek, N.J., Poldrack, R.A., and Schonberg, T. (2019). Neural correlates of effort-based valuation with prospective choices. Neuroimage *185*, 446–454.

Arulpragasam, A.R., Cooper, J.A., Nuutinen, M.R., and Treadway, M.T. (2018). Corticoinsular circuits encode subjective value expectation and violation for effortful goal-directed behavior. Proceedings of the National Academy of Sciences *115*, E5233–E5242.

Bang, D., and Fleming, S.M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. Proceedings of the National Academy of Sciences *115*, 6082–6087.

Bartra, O., McGuire, J.T., and 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.

Bhanji, J.P., Beer, J.S., and Bunge, S.A. (2010). Taking a gamble or playing by the rules: dissociable prefrontal systems implicated in probabilistic versus deterministic rule-based decisions. Neuroimage *49*, 1810–1819.

Blanchard, T.C., and Gershman, S.J. (2018). Pure correlates of exploration and exploitation in the human brain. Cogn Affect Behav Neurosci *18*, 117–126.

Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., and 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.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict monitoring and cognitive control. Psychol Rev *108*, 624–652.

Botvinick, M.M., Huffstetler, S., and McGuire, J.T. (2009). Effort discounting in human nucleus accumbens. Cognitive, Affective, & Behavioral Neuroscience *9*, 16–27.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spatial Vision *10*, 433–436.

Cai, X., and Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. J Neurosci *32*, 3791–3808.

Chib, V.S., Rangel, A., Shimojo, S., and 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.

Chua, E.F., Schacter, D.L., Rand-Giovannetti, E., and Sperling, R.A. (2006). Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. Neuroimage *29*, 1150–1160.

Clithero, J.A., and Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. Social Cognitive and Affective Neuroscience *9*, 1289–1302.

De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. Nature Neuroscience *16*, 105–110.

De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., and Love, B.C. (2017). Social Information Is Integrated into Value and Confidence Judgments According to Its Reliability. The Journal of Neuroscience *37*, 6066–6074.

Deichmann, R., Gottfried, J.A., Hutton, C., and Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. Neuroimage *19*, 430–441.

Domenech, P., Rheims, S., and Koechlin, E. (2020). Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. Science *369*, eabb0184.

Durston, S. (2003). Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. NeuroImage *20*, 2135–2141.

Fouragnan, E., Retzler, C., Mullinger, K., and Philiastides, M.G. (2015). Two spatiotemporally distinct value systems shape reward-based learning in the human brain. Nat Commun *6*, 8107.

Gherman, S., and Philiastides, M.G. (2018). Human VMPFC encodes early signatures of confidence in perceptual decisions. ELife *7*, e38293.

Gläscher, J., Hampton, A.N., and 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.

Hare, T.A., Camerer, C.F., and Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. Science *324*, 646–648.

Harvey, A.H., Kirk, U., Denfield, G.H., and Montague, P.R. (2010). Monetary favors and their influence on neural responses and revealed preference. J Neurosci *30*, 9597–9602.

Hogan, P.S., Galaro, J.K., and Chib, V.S. (2019). Roles of Ventromedial Prefrontal Cortex and Anterior Cingulate in Subjective Valuation of Prospective Effort. Cerebral Cortex *29*, 4277–4290.

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F.S., and Behrens, T.E.J. (2012). Mechanisms underlying cortical activity during value-guided choice. Nature Neuroscience *15*, 470–476.

Jocham, G., Hunt, L.T., Near, J., and Behrens, T.E.J. (2012). A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. Nat Neurosci *15*, 960–961.

Jocham, G., Furlong, P.M., Kröger, I.L., Kahn, M.C., Hunt, L.T., and Behrens, T.E.J. (2014). Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. Neuroimage *100*, 498–506.

Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. Nat Neurosci *10*, 1625–1633.

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. Science *154*, 1583–1585.

Kamiński, J., Sullivan, S., Chung, J.M., Ross, I.B., Mamelak, A.N., and Rutishauser, U. (2017). Persistently active neurons in human medial frontal and medial temporal lobe support working memory. Nature Neuroscience *20*, 590–601.

Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., and Carter, C.S. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. Science *303*, 1023–1026.

Kiani, R., Corthell, L., and Shadlen, M.N. (2014). Choice certainty is informed by both evidence and decision time. Neuron *84*, 1329–1342.

Kirk, U., Harvey, A., and Montague, P.R. (2011). Domain expertise insulates against judgment bias by monetary favors through a modulation of ventromedial prefrontal cortex. PNAS *108*, 10332–10336.

Klein-Flugge, M.C., Kennerley, S.W., Friston, K., and Bestmann, S. (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. Journal of Neuroscience *36*, 10002–10015.

Kolling, N., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2012). Neural Mechanisms of Foraging. Science *336*, 95–98.

Kolling, N., Wittmann, M.K., Behrens, T.E.J., Boorman, E.D., Mars, R.B., and Rushworth, M.F.S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. Nature Neuroscience *19*, 1280–1285.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., and Baker, C.I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci *12*, 535–540.

Kurniawan, I.T., Guitart-Masip, M., Dayan, P., and Dolan, R.J. (2013). Effort and Valuation in the Brain: The Effects of Anticipation and Execution. Journal of Neuroscience *33*, 6160–6169.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. Neuron *64*, 431–439.

Lebreton, M., Bertoux, M., Boutet, C., Lehericy, S., Dubois, B., Fossati, P., and Pessiglione, M. (2013). A Critical Role for the Hippocampus in the Valuation of Imagined Outcomes. PLoS Biology *11*, e1001684.

Lebreton, M., Abitbol, R., Daunizeau, J., and Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. Nature Neuroscience *18*, 1159–1167.

Lee, D., and Daunizeau, J. (2020). Choosing what we like vs liking what we choose: How choice-induced preference change might actually be instrumental to decision-making. PLoS ONE *15*, e0231081.

Lee, D., and Daunizeau, J. (2021). Trading mental effort for confidence in the metacognitive control of value-based decision-making. ELife *10*, e63282.

Levy, D.J., and Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. Current Opinion in Neurobiology *22*, 1027–1038.

Levy, I., Lazzaro, S.C., Rutledge, R.B., and Glimcher, P.W. (2011). Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level-Dependent Signals Obtained during Passive Viewing. Journal of Neuroscience *31*, 118–125.

Lim, S.-L., O'Doherty, J.P., and 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.

Lopez-Gamundi, P., Yao, Y.-W., Chong, T.T.-J., Heekeren, H.R., Herrero, E.M., and Pallares, J.M. (2021). The neural basis of effort valuation: A meta-analysis of functional magnetic resonance imaging studies. BioRxiv 2021.01.08.425909.

Lopez-Persem, A., Domenech, P., and Pessiglione, M. (2016). How prior preferences determine decision-making frames and biases in the human brain. ELife *5*.

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., et al. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. Nat Neurosci 1–12.

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

Oudiette, D., Vinckier, F., Bioud, E., and Pessiglione, M. (2019). A Pavlovian account for paradoxical effects of motivation on controlling response vigour. Scientific Reports *9*.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. Annu. Rev. Neurosci. *34*, 333–359.

Pessiglione, M., and Daunizeau, J. (2021). Bridging across functional models: The OFC as a value-making neural network. Behav Neurosci *135*, 277–290.

Pessiglione, M., Vinckier, F., Bouret, S., Daunizeau, J., and Le Bouc, R. (2018). Why not try harder? Computational approach to motivation deficits in neuro-psychiatric diseases. Brain *141*, 629–650.

Pisauro, M.A., Fouragnan, E., Retzler, C., and Philiastides, M.G. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. Nature Communications *8*, 15808.

Plassmann, H., O'Doherty, J.P., and Rangel, A. (2010). Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. J. Neurosci. *30*, 10799–10808.

Pleskac, T.J., and Busemeyer, J.R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. Psychol Rev *117*, 864–901.

Qin, S., Marle, H.J.F. van, Hermans, E.J., and 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.

Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. Current Opinion in Neurobiology *20*, 262–270.

Ratcliff, R., Smith, P.L., Brown, S.D., and McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. Trends in Cognitive Sciences *20*, 260–281.

Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., and Rushworth, M.F.S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. J Neurosci *28*, 13775–13785.

Rushworth, M.F., Kolling, N., Sallet, J., and Mars, R.B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Current Opinion in Neurobiology *22*, 946–955.

Seaman, K.L., Brooks, N., Karrer, T.M., Castrellon, J.J., Perkins, S.F., Dang, L.C., Hsu, M., Zald, D.H., and Samanez-Larkin, G.R. (2018). Subjective value representations during effort, probability and time discounting across adulthood. Social Cognitive and Affective Neuroscience *13*, 449–459.

Shadmehr, R., Reppert, T.R., Summerside, E.M., Yoon, T., and Ahmed, A.A. (2019). Movement Vigor as a Reflection of Subjective Economic Utility. Trends in Neurosciences *42*, 323–336.

Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. Neuron *79*, 217–240.

Skvortsova, V., Palminteri, S., and Pessiglione, M. (2014). Learning To Minimize Efforts versus Maximizing Rewards: Computational Principles and Neural Correlates. Journal of Neuroscience *34*, 15621–15630.

Sohn, M.-H., Albert, M.V., Jung, K., Carter, C.S., and Anderson, J.R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. PNAS *104*, 10330–10334.

Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K., and Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. Sci Rep *9*, 20053.

Talmi, D., Dayan, P., Kiebel, S.J., Frith, C.D., and Dolan, R.J. (2009). How humans integrate the prospects of pain and reward during choice. J. Neurosci. *29*, 14617–14626.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. Science *315*, 515–518.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and 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.

Weiskopf, N., Hutton, C., Josephs, O., Turner, R., and Deichmann, R. (2007). Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. Magnetic Resonance Materials in Physics, Biology and Medicine *20*, 39–49.

van der Wel, P., and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic Bulletin & Review *25*, 2005–2015.

Westbrook, A., Lamichhane, B., and Braver, T. (2019). The Subjective Value of Cognitive Effort is Encoded by a Domain-General Valuation Network. The Journal of Neuroscience 3071–18.

Wunderlich, K., Rangel, A., and 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.

Yeung, N., and Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. Philos Trans R Soc Lond B Biol Sci *367*, 1310–1321.

Zénon, A., Sidibé, M., and Olivier, E. (2014). Pupil size variations correlate with physical effort perception. Frontiers in Behavioral Neuroscience *8*.

Figures

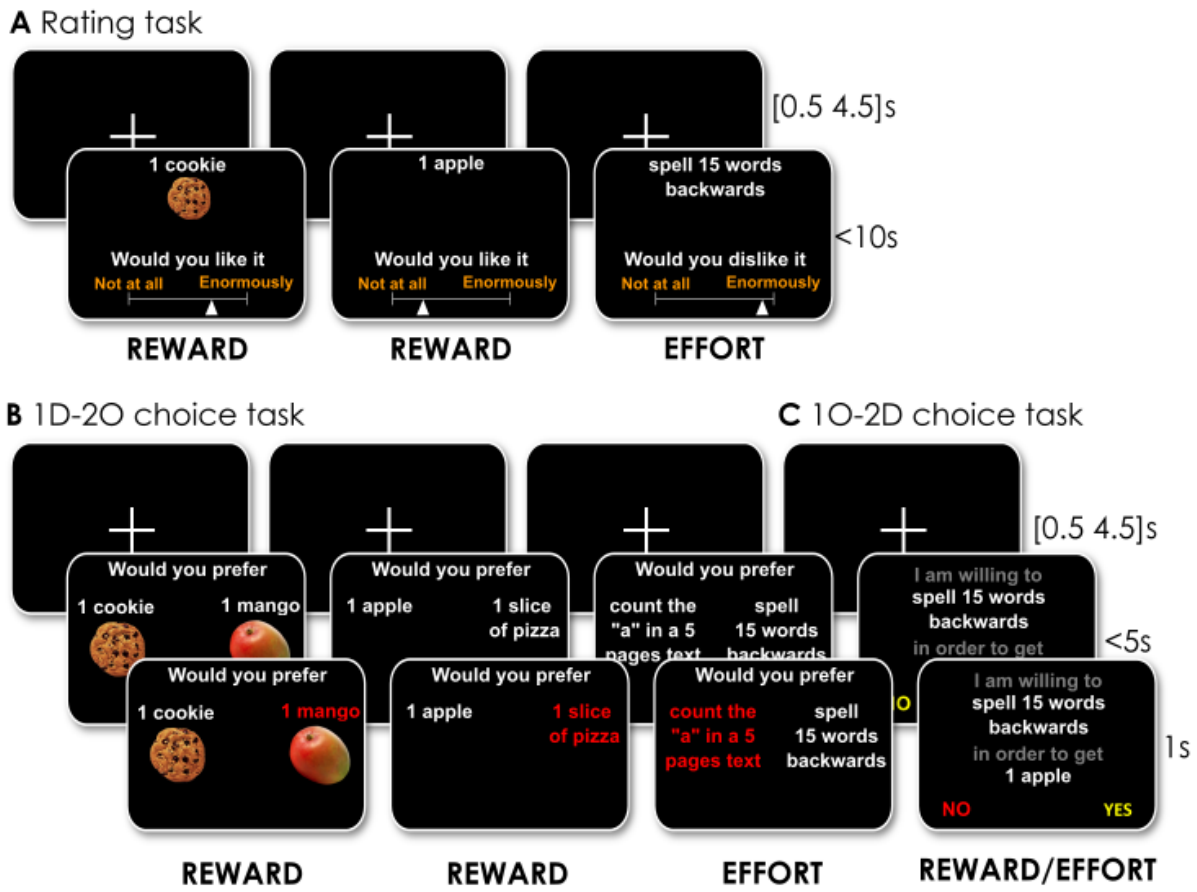

Figure 1. Behavioral tasks.

Example trials are illustrated as a succession of screenshots from top to bottom, with durations in seconds. Only the duration of fixation cross display at the beginning of trials is jittered. The duration of the response screen depends on deliberation time, as both rating and choice are self-paced.

A] Rating task. In every trial, subjects are shown an item that can be a reward described with both text and image (R_{ti}), a reward described with text only (R_t) or an effort described with text only (E_t). The task for subjects is to rate how much they would like receiving the proposed reward or dislike performing the proposed effort, should it occur, hypothetically, at the end of the experiment. They first move the cursor using left and right buttons on a pad to the position that best reflect their (dis)likeability estimate, then validate their response with a third button and proceed to the next trial.

B] 1D-2O choice task. In every trial, two options belonging to the same category are shown on screen and subjects are asked to select their favorite option, i.e. which reward they would prefer to receive if they were offered the two options (hypothetically) or which effort they would prefer to exert if they were forced to implement one of the two options (hypothetically). The choice is expressed by selecting between left and right buttons with the index or middle finger. The chosen option is then highlighted in red, and subjects proceed to the next trial.

C] 1O-2D choice task. In every trial, one option combining the two dimensions is shown on screen and subjects are asked to state whether they would be willing the exert the effort in order to receive the reward, if they were given the opportunity at the end of the experiment (hypothetically). They select their response (‘yes’ or ‘no’, positions counterbalanced across trials) by pressing the left or right button, with their index or middle finger.

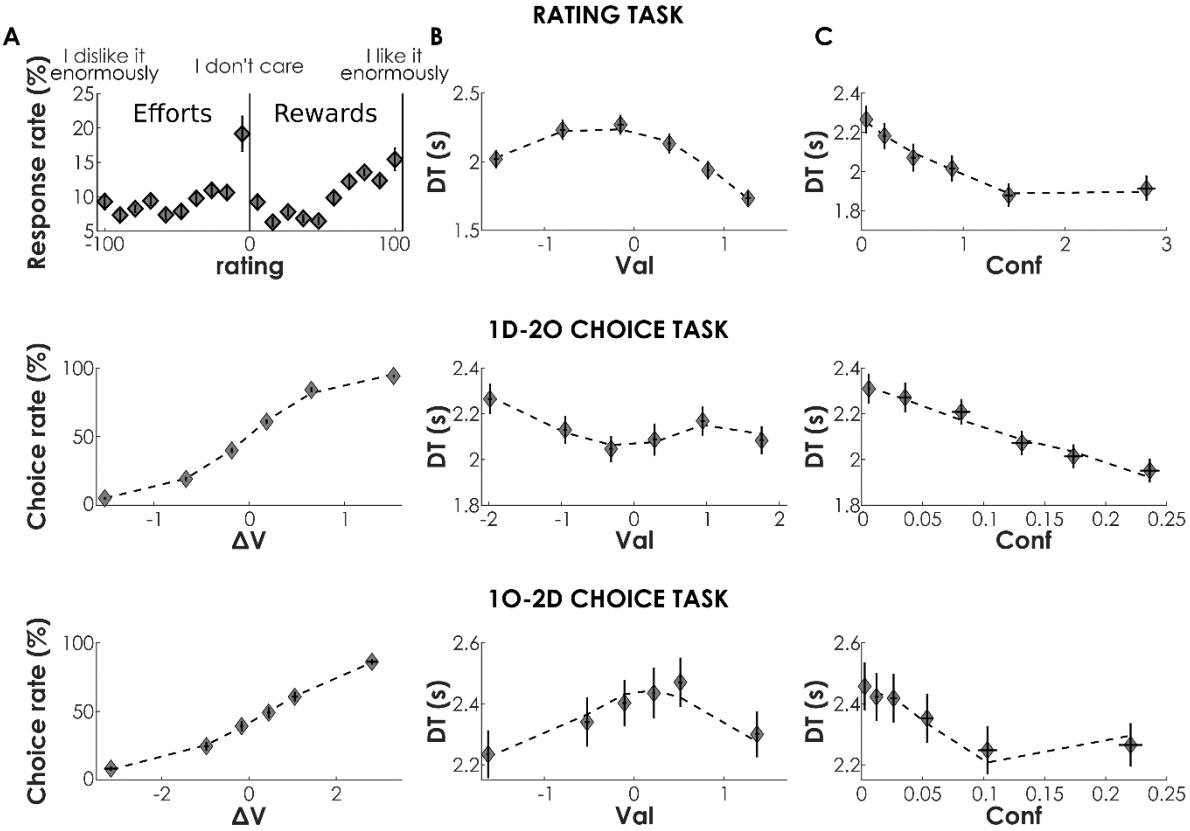

Figure 2: Behavioral results.

A] Response rate. For ratings, plots show the average response rate for each bin (portion of the rating scale). Effort items (on the left) are rated between bin 0 ('I would not mind') and bin -10 ('I would dislike it enormously'). Reward items (on the right) are rated between bin 0 ('I would not care') and bin +10 ('I would like it enormously'). Note that the x-axis has been reverted for effort ratings, compared to the visual scale presented in the task, such that it globally indicates increasing values (less aversive effort from -100 to 0 and more appetitive rewards from 0 to +100). For choices, the response rate is plotted as a function of binned decision value (ΔV). In the 1D-2O task, decision value is the difference in likeability rating between left and right options ($V_{left} - V_{right}$), and choice rate is the frequency of left option being selected. In the 1O-2D task, decision value is the addition of weighted reward and effort likeability ratings ($\beta_R \cdot V_R + \beta_E \cdot V_E$), which is equivalent to both stimulus value (Val) and to the value difference between yes and no options (net value minus zero). Lines show logistic regression fits.

B] Deliberation time as a function of stimulus value (Val), defined as the sum of centered likeability ratings assigned to items on screen, meaning V_R or V_E for the rating task, $V_{left} + V_{right}$ for the 1D-2O choice task, $\beta_R \cdot V_R + \beta_E \cdot V_E$ for the 1D-2O choice task.

C] Deliberation time as a function of confidence proxy (Conf), defined as the square of centered likeability ratings (V^2) for rating tasks and the square of centered choice likelihood (P^2) for choice tasks. Dots represent mean across participants, x and y error bars are inter-participant standard errors.

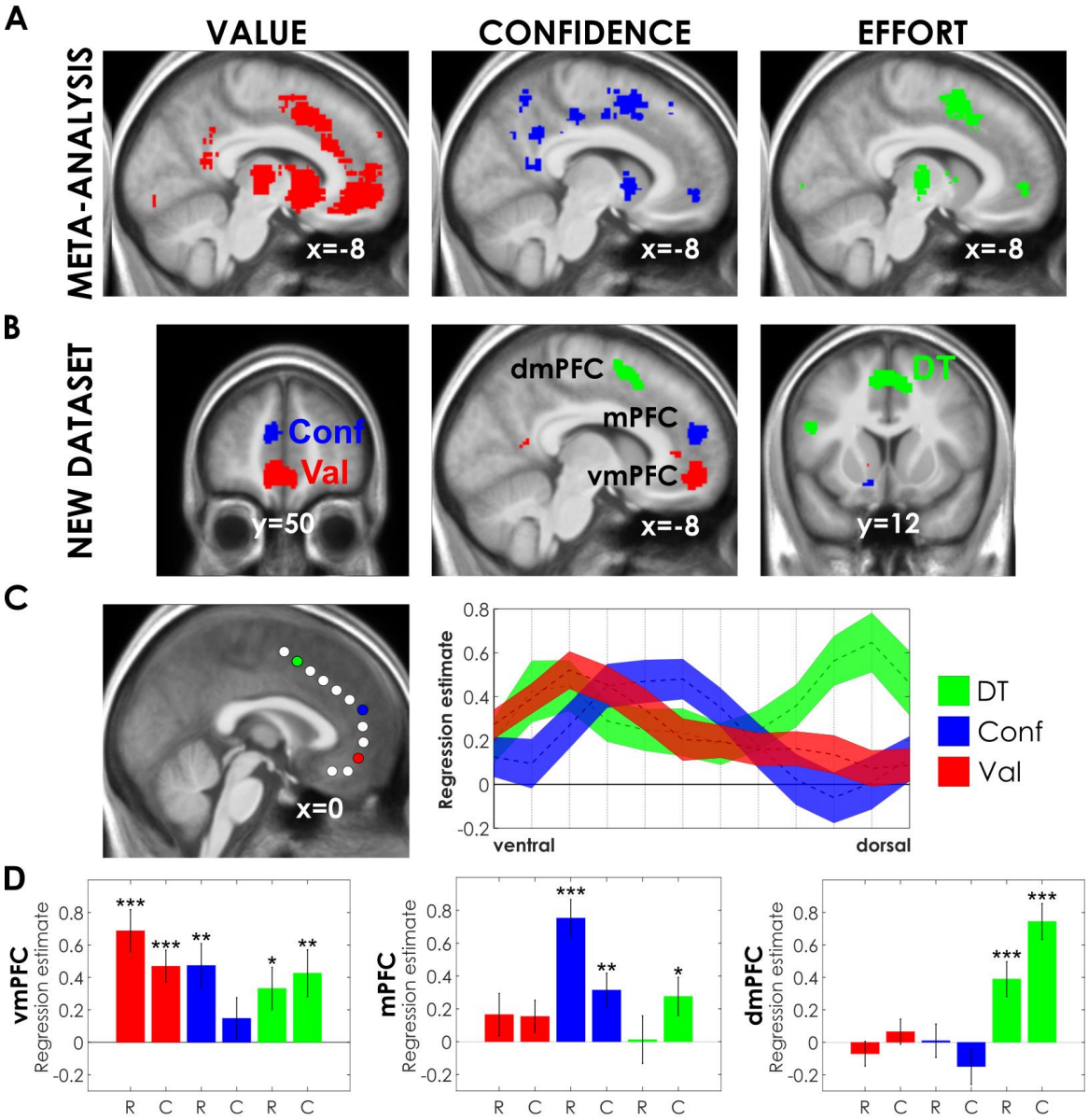

Figure 3: Neural results.

A] Meta-analysis of fMRI studies. Statistical maps (sagittal slices) were extracted from the Neurosynth platform with the 'value', 'confidence' and 'effort' keywords. Significant clusters in the medial prefrontal cortex are similar across keywords, being located in both ventral and dorsal regions.

B] Neural correlates of value, confidence and deliberation constructs in the present dataset. Statistical maps were obtained with a GLM including the different variables as parametric modulators of stimulus onset, across rating and choice tasks. Sagittal slice was taken at the same coordinates as the Neurosynth output, and superimposed on the average anatomical scan normalized to canonical (MNI) template. Coronal slices show the extent of the different medial prefrontal clusters. Statistical threshold was set at

$p < 0.05$ after family-wise error for multiple comparisons at the voxel level. For clusters outside the medial prefrontal cortex, surviving a more tolerant statistical threshold, see activations in Tables S1-3.

C] Distribution of regression estimates (inter-subject means ± standard errors) obtained for Val, Conf and DT variables along a ventro-dorsal line within the medial prefrontal cortex (sampled in each 8mm-radius shown on the average anatomical map). Colored circles show sampled spheres in which correlation with the corresponding variable was maximal (Val – red, Conf – blue and DT – green).

D] Decomposition of regression estimates obtained for each variable of interest, plotted separately for rating and choice tasks (noted R and C) and for the different ROI (vmPFC, mPFC, dmPFC). For the distinction between choice tasks (1D-2O vs. 10-2D, see Fig. S5).

Bars show mean across participants, error bars show inter-participant standard errors. Stars indicate significance of t-test against zero (*** $p < 0.001$, ** $p < 0.01$, * $p< 0.05$).

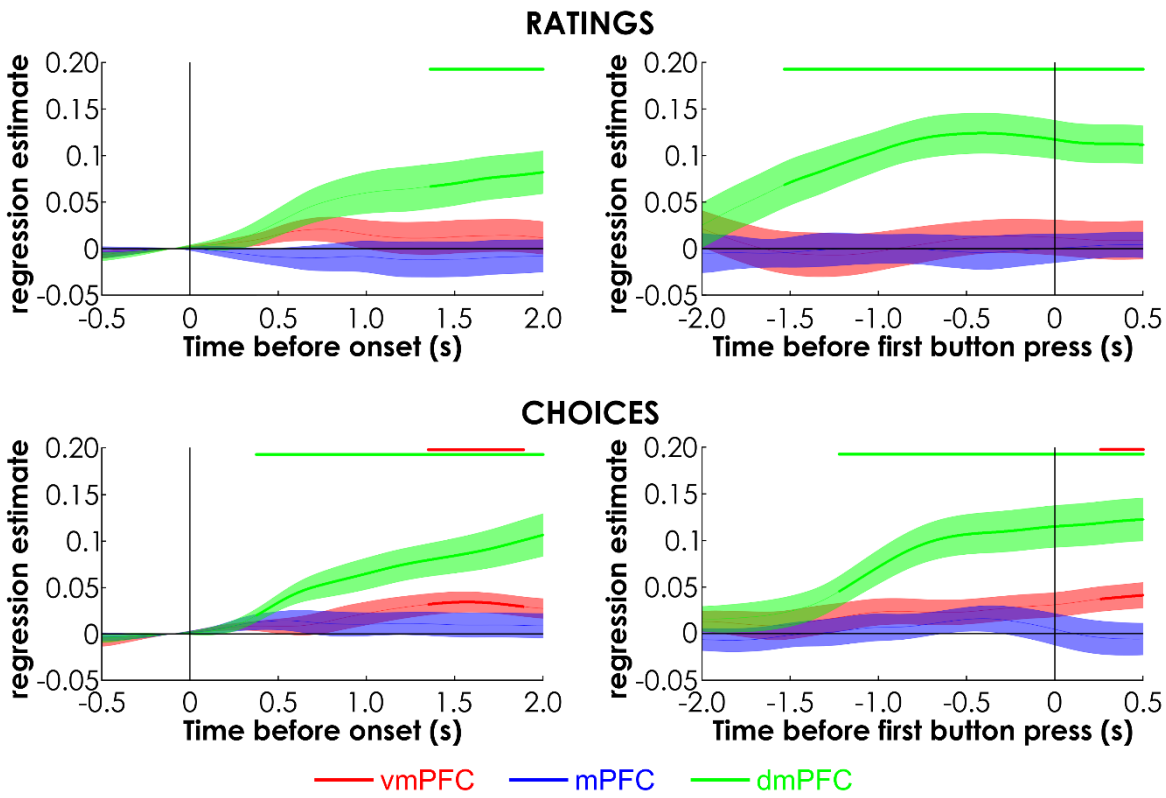

Figure 4: Pupillometric results.

Plots show the time course of regression estimates, obtained with a GLM built to explain pupil size. The GLM included nuisance regressors (jitter duration, stimulus luminance, text length), variables of interest (Val, Conf, DT) and activities in main ROI (vmPFC, mPFC, dmPFC, corresponding to red, blue and green traces, respectively). Each row corresponds to a different task (likeability rating, choice tasks). Left and right columns show time courses aligned onto stimulus onset and button press, respectively. Lines represent means across participants and shaded areas inter-participant standard errors. Horizontal bars indicate significant time clusters after correction for multiple comparisons using the random-field theory.