Value, confidence and deliberation in preference tasks: a triple dissociation in the medial prefrontal cortex

N. Clairis[1,2], M. Pessiglione[1,2]

[1]Motivation, Brain and Behavior team, Institut du Cerveau (ICM),

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

Abstract

Deciding about courses of action involves an estimation of costs and benefits. Decision neuroscience studies have suggested a dissociation between the ventral and dorsal medial prefrontal cortex (vmPFC and dmPFC), which would process reward value and effort cost, respectively. However, several results appeared inconsistent with this general idea of opponent reward and effort systems. These contradictions might reflect the diversity of tasks used to investigate the trade-off between effort cost and reward value. They might also reflect the confusion with a meta-decision process about the amount of deliberation needed to reach a sufficient confidence in the reward/effort estimates. Here, we used fMRI to examine the neural correlates of reward and effort estimates across several preference tasks, from (dis-)likeability ratings to binary decisions involving attribute integration and option comparison. Results confirm the role of the vmPFC as a generic valuation system, across the different tasks (likeability rating or binary decision) and attributes (the activity increasing with reward value and decreasing with effort cost). However, meta-decision variables were represented in more dorsal regions, with confidence in the mPFC and deliberation time in the dmPFC. These findings suggest that assessing commonalities across preference tasks and distinguishing between decision and meta-decision variables might help reaching a unified view of how the brain chooses a course of action.

Introduction

Cost-benefit arbitrage is a pervasive 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. In decision neuroscience, a general debate has been engaged between opponency theories, where pursuit of benefits and avoidance of costs are mediated by separate brain systems, and integration theories, where costs and benefits are combined in a single brain system (Boureau and Dayan, 2011; Cisek and Kalaska, 2010; Padoa-Schioppa, 2011; Pessiglione and Delgado, 2015; Rangel and Hare, 2010).

A central dilemma that has been scrutinized in decision neuroscience is between the resource to be invested in the action (i.e., the effort) and the value of the action outcome (i.e., the reward). Many fMRI studies have documented the existence of a brain valuation system (BVS), with the ventromedial prefrontal cortex (vmPFC) as a core component, that would signal the value of various categories of rewards (Bartra et al., 2013; Levy and Glimcher, 2012). Although the generality of the vmPFC value signal across reward categories is well accepted, whether or not this value signal integrates effort costs is a controversial issue. While some studies have reported that effort costs result in a decreased vmPFC value signal (Arulpragasam et al., 2018; Seaman et al., 2018), others have suggested the existence of an opponent system that would signal effort costs with increasing activity, such as the dorsomedial prefrontal cortex (dmPFC, sometimes referred to as 'dorsal anterior cingulate cortex') and the anterior insula (Kurniawan et al., 2013; Skvortsova et al., 2014).

One explanation for theses discrepancies is the diversity of behavioral tasks employed to probe the neural correlates of reward and effort estimates. Here, 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 (dis-)likeability 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.

Another explanation for inconsistencies in the neural correlates of reward and effort is the possible confusion with meta-decision variables. Indeed, reward value and effort cost estimation can itself be performed with various amounts of deliberation effort. Intuitively, everyone knows that some decisions are made on the fly, without thinking too much, while others demand a long and painful deliberation, in order to weigh the pros and cons. This intuition has been captured in a theory postulating that the investment of cognitive control is based on a cost-benefit arbitrage (Shenhav et al., 2013). Applied to decision-making, the cost is essentially the time that the agent spends considering the different options and their attributes, while the benefit is essentially the confidence that the agent gains in making the right choice. Therefore, in addition to the reward and effort levels that are attributes of choice options, we examined constructs related to the meta-choice: confidence and deliberation time. Importantly, the reward and effort estimates did not imply actual and immediate implementation, as all choices were hypothetical in our tasks. On the contrary, confidence and time were variables associated to current behavior, i.e. to the ongoing estimation of reward value and/or effort cost.

Our results suggest a triple dissociation within the medial prefrontal cortex that is stable across preference tasks, with option attributes (reward and effort levels) being represented in vmPFC activity, whereas meta-decision variables (confidence and deliberation time) being represented in mPFC and dmPFC activity, respectively.

Results

Behavior

Participants (n=39 in total) first performed a series of ratings, divided into three blocks (Fig. 1A). Each block presented 72 items one by one: reward items presented with text + image (R_{ti}), reward items presented with text only (R_t) and effort items presented with 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 demand more or less imagination, according to previous study (Lebreton et al., 2013). The order of the R_{ti}, R_t and E_t blocks was counterbalanced across participants. For rewards, 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 efforts was to rate how much they would dislike it, should they be requested to implement 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. 2A).

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). Deliberation time (DT) was defined in this task 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, in addition to nuisance factors (such as jitter duration, stimulus luminance, text length and trial index, see methods). We found in all blocks (Fig. 2B) a significant effect of both value (R_{ti}: $\beta_{Val} = -0.208 \pm 0.023$ (mean ± standard error of the mean across participants), $p = 4{\cdot}10^{-11}$; R_t: $\beta_{Val} = -0.169 \pm 0.019$, $p = 6{\cdot}10^{-11}$; E_t: $\beta_{Val} = 0.261 \pm 0.028$, $p = 2{\cdot}10^{-11}$) and confidence (R_{ti}: $\beta_{Conf} = -0.174 \pm 0.025$, $p = 3{\cdot}10^{-8}$; R_t: $\beta_{Conf} = -0.188 \pm 0.028$, $p = 7{\cdot}10^{-8}$; E_t: $\beta_{Conf} = -0.130 \pm 0.040$; $p = 0.0024$). 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 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 three blocks (R_{ti},R_t, E_t) of 72 choices. Again, the order of the R_{ti}, R_t and E_t blocks was counterbalanced across participants. 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, each item again appearing twice, for a total of 144 choices divided into three blocks.

The 1D-2O choice task (Fig. 1C) 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 that DV was a significant predictor of choices (Fig. 2E) in all blocks (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. 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), we defined proxies for stimulus value and response confidence, as we did for the rating task. Stimulus value (Val) was defined as a linear integration 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 the net value (sum of reward and effort ratings weighted by a scaling factor). 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 showed that DT decreased with value (Fig. 2C) 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 (Fig. 2D) 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).

Neural activity

The aim of fMRI data analysis was to dissociate the variables related to option valuation (reward and effort estimates) from the variables related to the meta-decision (confidence and deliberation) across value-based tasks (rating and choice). A meta-analysis of fMRI studies using Neurosynth platform (Fig. 3A) shows that value, confidence and effort keywords yield to similar activation patterns with clusters in both vmPFC and dmPFC. To better dissociate the neural correlates of these constructs, we built a general linear model where stimulus onset events were modulated by three parameters of interest - Val, Conf and DT (defined as in the behavior 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. Thus, due to serial orthogonalization, the variables of interest were orthogonalized with respect to nuisance factors, and deliberation time was made orthogonal to all other regressors, including stimulus value and response confidence.

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, Conf in mPFC activity and DT in dmPFC activity. 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. This was not the case in the other clusters, which were either medial or bilateral.

We further analyzed the relationship between computational variables and activity in the three medial prefrontal regions of interest (ROI) with post-hoc t-tests. We first verified that the three main associations were significant in each task (Fig. 3C): it was indeed the case for Val in vmPFC activity (rating: $\beta_{Val} = 0.701 \pm 0.126$, $p = 2{\cdot}10^{-6}$; 1D-2O: $\beta_{Val} = 0.293 \pm 0.126$, $p = 0.025$; 1O-2D: $\beta_{Val} = 0.700 \pm 0.180$, $p = 4{\cdot}10^{-4}$), for Conf in mPFC activity (rating: $\beta_{Conf} = 0.768 \pm 0.111$, $p = 3{\cdot}10^{-8}$; 1D-2O: $\beta_{Conf} = 0.335 \pm 0.114$, $p = 0.006$; 1O-2D: $\beta_{Conf} = 0.313 \pm 0.133$, $p = 0.024$) and for DT in dmPFC activity (rating: $\beta_{DT} = 0.410 \pm 0.106$, $p = 4{\cdot}10^{-4}$; 1D-2O: $\beta_{DT} = 0.875 \pm 0.123$, $p = 2{\cdot}10^{-8}$; 1O-2D: $\beta_{DT} = 0.676 \pm 0.125$, $p = 4{\cdot}10^{-6}$). We then verified that the three dissociations were significant (Fig. 3D): Val was better reflected in vmPFC activity ($\beta_{Val/vmPFC} > \beta_{Val/mPFC}$: $p = 9{\cdot}10^{-9}$; $\beta_{Val/vmPFC} > \beta_{Val/dmPFC}$: $p = 5{\cdot}10^{-8}$), Conf in mPFC activity ($\beta_{Conf/mPFC} > \beta_{Conf/vmPFC}$: $p = 0.0035$; $\beta_{Conf/mPFC} > \beta_{Conf/dmPFC}$: $p = 2{\cdot}10^{-7}$) and DT in dmPFC activity ($\beta_{DT/dmPFC} > \beta_{DTvmPFC}$: $p = 0.021$; $\beta_{DT/dmPFC} > \beta_{DTmPFC}$: $p = 2{\cdot}10^{-4}$).

Thus, the triple dissociation observed in the maps was robust across tasks and was supported by significant differences between ROI. However, the triple dissociation does not imply that the three variables of interest were solely represented in one single brain region. In particular, Conf and DT were also significantly related to vmPFC activity ($\beta_{Conf} = 0.256 \pm 0.098$, $p = 0.013$; $\beta_{DT} = 0.377 \pm 0.108$, $p = 0.001$), even if these activities were dominated by Val-related activity. In a pilot study, we compared the standard EPI acquisition sequence used in the main experiment to sequences using multiband acceleration and multi-echo acquisition.

Although the number of participants (n=15) and the number of task sessions (one out of three) was too low for sound statistical inference, we nonetheless examined whether the triple dissociation would hold with these alternative sequences. This dataset was independent from the main dataset based on which regions of interest were defined. We observed the same trends (see Fig. S1), with regression estimates higher for Val in vmPFC, Conf in mPFC and DT in dmPFC. However, only Val in vmPFC and DT in dmPFC survived statistical thresholds for significance.

We developed variants of our GLM to further assess the robustness of these findings. Regarding Val, we examined whether other value-related variables employed in previous studies, such as decision value (e.g., chosen minus unchosen option value) could better account for vmPFC activity. When replacing our proxy for stimulus value by these other variables, we did not find any stronger correlation with activity in the vmPFC, even when defined anatomically or from meta-analyses of fMRI data. Regarding Conf, we assessed whether the dorsal location of the cluster (mPFC instead of vmPFC) could be related to the orthogonalization with Val, by simply removing the Val regressor, but results were unchanged. Regarding DT, we tested whether the association with dmPFC activity could arise from the delay itself, and not from a prolonged deliberation. When replacing the delta function modeling stimulus onset by a boxcar function (whose duration varies with DT), the association between DT and dmPFC activity was still significant, suggesting a modulation in amplitude and not just a prolongation of the signal.

We looked for further generalization of the valuation signal, not solely across tasks but also across stimuli. We focused on the rating task, in which the link with neural activity is easier to assess, as there is only one stimulus to valuate (Fig. 3D). FMRI time series were regressed against a GLM that separated stimulus categories (Rti, Rt and Et) into different onset regressors, each modulated by corresponding ratings. Results show that vmPFC activity was positively related to the value (likeability rating) of reward items, whether or not they are presented with an image, and negatively correlated to the cost (dislikeability rating) of effort items (Rti: $\beta_{Val} = 0.630 \pm 0.224$, $p = 0.008$; Rt: $\beta_{Val} = 0.803 \pm 0.177$, $p = 6{\cdot}10^{-5}$; Et: $\beta_{Val} = -0.670 \pm 0.205$, $p = 0.002$). Thus, the association between Val and vmPFC activity was independent of the presentation mode, and integrated costs as well as benefits. Importantly, the association with reward value or effort cost was not observed in putative opponent brain regions such as the dmPFC, whose activity even tended to decrease with dislikeability rating of effort items (Et: $\beta Val = -0.225 \pm 0.116$, $p = 0.060$).

Thus, it appeared that dmPFC activity did not reflect the effort cost attached to the option on valuation but the effort cost of the meta-decision (selecting a response). 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. As DT is a very indirect proxy for the effort invested in solving the task, we investigated the relationship with 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 all 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 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). In particular, there was no consistent association between vmPFC and pupil size, 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 triple dissociation 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 meta-decision 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 disprove opponent systems theories that would predict separate representations, notably those assuming that the vmPFC is involved in stimulus valuation, while action costs would implicate the dmPFC (Padoa-Schioppa, 2011; Rangel and Hare, 2010; Schneider and Koenigs, 2017). 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 previous demonstrations that the vmPFC integrates costs such as potential loss or delay in reward delivery (Hare et al., 2009, 2011; Kable and Glimcher, 2007; Talmi et al., 2009; Tom et al., 2007). 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 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 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; Lebreton et al., 2015). 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. An alternative interpretation 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 previously described (De Martino et al., 2017). On the contrary, dmPFC activity tended to decrease with confidence, but this association 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 both with 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 (Kiani et al., 2014), 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 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 (Botvinick et al., 2001; Shenhav et al., 2013). 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 (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 decision 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 meta-decisional 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, 2019). This model would predict that increasing dmPFC activity is meant to increase mPFC activity, which we could not test here because our correlational approach precludes any inference about causality. Yet it remains possible that dmPFC activation is not about demand for control but about estimating the amount of control being invested, 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 meta-decisional regulation.

Finally, we have shown 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 and general 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. Participants were recruited through the online RISC (Relais d'Information en Sciences de la Cognition) online platform (https://www.risc.cnrs.fr/). All participants were screened for the use of psychotropic medications, alcohol and drug use, and history of psychiatric disorders, cognitive/neurological 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 excessive movement inside the scanner (>3mm within-session per direction). Seven to nine additional participants were excluded from pupil size analysis, depending on the task, due to poor signal detection.

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 first protocol (n=15) aimed at comparing scanning sequences. Each task was subdivided into 3 sessions. Each session was scanned through a different scanning protocol using regular EPI, EPI with multiband acceleration, EPI with multiband + multi-echo acquisition. The main analysis only includes fMRI data recorded during the first session using regular EPI acquisition. The other participants were scanned with regular EPI during the nine sessions, which were all included in the analysis.

Behavioral tasks

All tasks were programmed using Psychtoolbox (http://psychtoolbox.org/) in Matlab 2012 (The MathWorks, Inc., USA). Participants were given a 4-button box (fORP 932, Current Designs Inc, Philadelphia, USA) placed under their right to provide their responses. For further data analyses, stimulus luminance was calculated using standard function of red-green-blue composition 0.299·red + 0.587·green + 0.114·blue (http://www.w3.org/TR/AERT#color-contrast), which was estimated through the Screen('GetImage') built-in psychtoolbox command. 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 session (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). Each of the three fMRI sessions included three blocks of 24 R_{ti} , 24 R_t and 24 E_t trials, the order of blocks being counterbalanced across participants. The items were presented one by one, and participants rated them by moving a cursor along an 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 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. 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 items (food vs. food, goods vs. goods, mental vs mental, physical vs physical). Each item was presented twice, following two inter-mixed 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 three sessions of the 1D-2O choice task included each 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 E_t). Each item was presented twice, following two inter-mixed 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 (this varying choice difficulty). The mean net value was also balanced across 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 sessions of 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 with intercept and decision value weighted by a free parameter.

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 rating between left and right options.

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

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

where V_R and V_E are the ratings provided for the reward and effort items. Thus, the decision value (or net value) here is a weighted sum reward likeability and effort dislikeability, with an additional parameter scales the two dimensions.

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 regressions models, with 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 minus effort dislikeability) 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	$[V\text{-mean}(V)]^2$	$[P_{left} - \text{mean}(P_{left})]^2$	$[P_{accept} - \text{mean}(P_{accept})]^2$

In each case, P is probability generated with the logistic regression and V is either reward likeability or effort dislikeability, provided by z-scored individual ratings.

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 = 25ms, flip angle = 78°, number of slices = 37, slice thickness = 2.5mm, field of view = 200mm. 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 9 sessions (3 per task) using this standard EPI sequence, but a pilot subgroup (n=15) performed only 3 sessions (1 per task). In this subgroup, 3 sessions were scanned using a sequence with a multi-band acceleration factor, and 3 sessions using multi-band + multi-echo acquisition. Functional data collected with these two other sequences have been analyzed to select the best acquisition sequence (standard EPI) for the main experiment (see results in supplementary information).

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 described above. The regressors of interest 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 across tasks of corresponding regression estimates against zero. Note that dislikeability ratings obtained for effort items were negatively weighted in all regressors (meaning that they can only decrease stimulus value).

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. Parameter estimates and t-values were extracted from each voxel within these clusters and then averaged across voxels. Finally, a last GLM was built with one event per trial, modeled with a stick function, at the time of stimulus onset. It was used to extract trial-by-trial activity levels in the clusters of interest, which then served as regressors to explain pupil size data (see below).

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 in paper mentioning the 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, after calibration before 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 1Hz) and zscored 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 behavioral data analysis) 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/).

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.

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.

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.

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.

Boureau, Y.-L., and Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology *36*, 74–97.

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.

Cisek, P., and Kalaska, J.F. (2010). Neural mechanisms for interacting with a world full of action choices. Annu. Rev. Neurosci. *33*, 269–298.

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.

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

Hare, T.A., Camerer, C.F., and Rangel, A. (2009). Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. *324*, 4.

Hare, T.A., Malmaud, J., and Rangel, A. (2011). Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice. Journal of Neuroscience *31*, 11077–11087.

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.

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.

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

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.

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. (2019). Trading Mental Effort for Confidence: The Metacognitive Control of Value-Based Decision-Making. BioRxiv 837054.

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.

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.

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 Delgado, M.R. (2015). The good, the bad and the brain: Neural correlates of appetitive and aversive values underlying decision making. Curr Opin Behav Sci *5*, 78–84.

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

Schneider, B., and Koenigs, M. (2017). Human lesion studies of ventromedial prefrontal cortex. Neuropsychologia *107*, 84–93.

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.

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.

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.

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

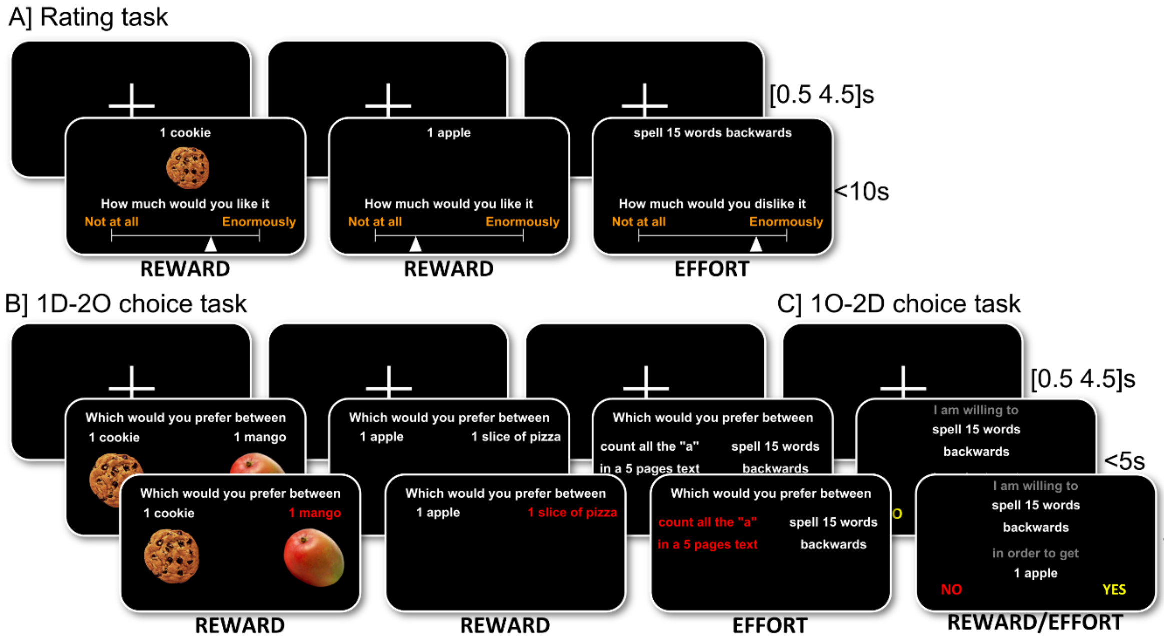

Figure 1. Behavioral tasks.

Example trials are illustrated as a succession of screenshots from top to bottom, with durations in seconds. Only the 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 the 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.

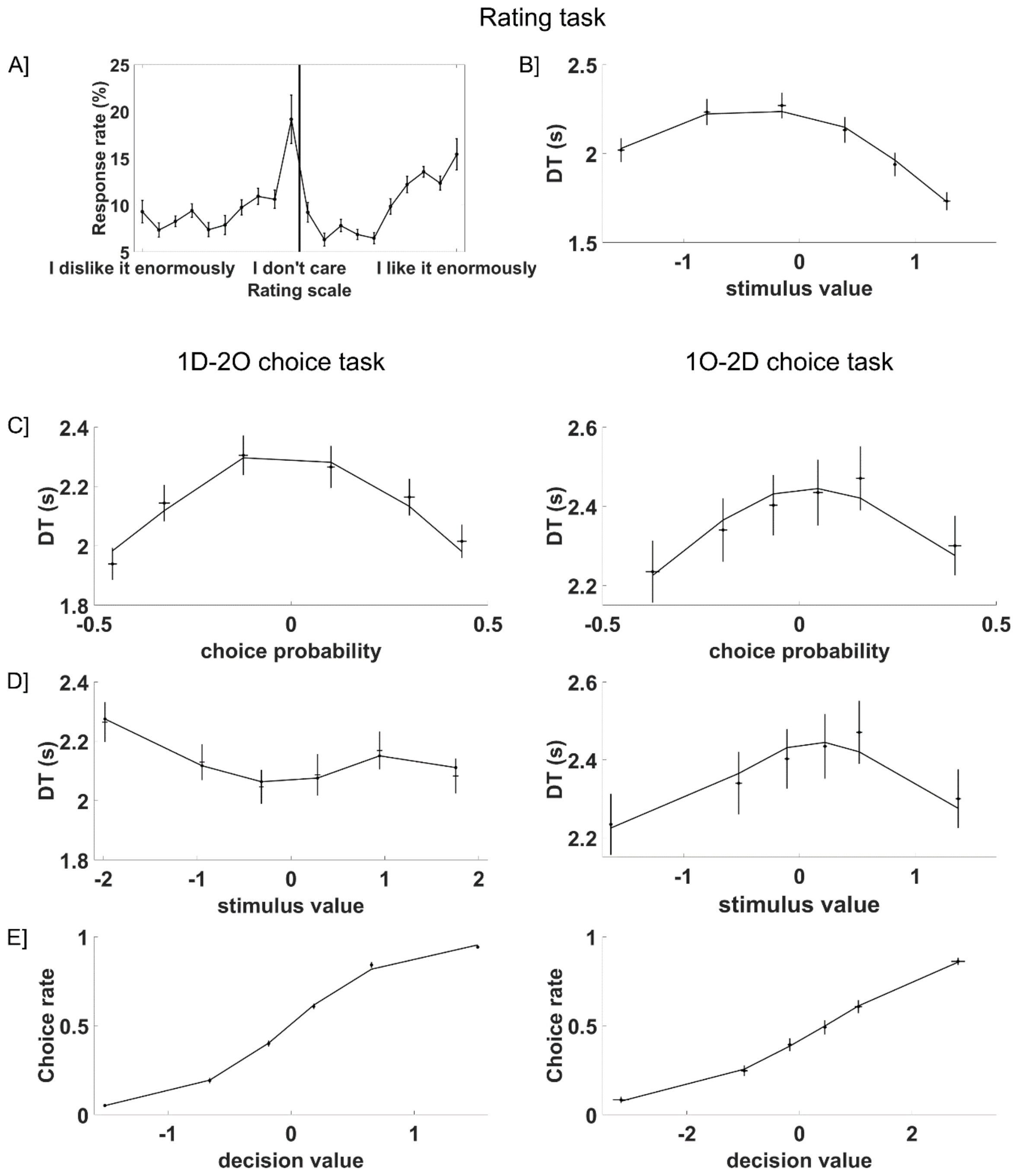

Figure 2: Behavioral results.

A] Distribution of ratings. Bars show the average response rate for each bin of ratings. 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').

B] Deliberation time in the rating task, as a function of the centered item value (likeability rating). Positive value means higher likeability (for rewards) or lower dislikeability (for efforts), 0 is the mean rating across trials.

C] Deliberation time as a function of stimulus value. In the 1D-2O choice task (left graph), stimulus value is the sum of likeability ratings for left and right options ($V_{left} + V_{right}$). In the 1O-2D choice task (right graph), stimulus value is the likeability of reward plus the dislikeability of effort. Thus, in this yes/no choice task, stimulus value is equivalent to decision value ($\beta_R \cdot V_R + \beta_E \cdot V_E$). In any case, stimulus value was centered, such that 0 is the mean across trials.

D] Deliberation time as a function of choice probability. In the 1D-2O choice task (left graph), choice probability is the output of the softmax function for the left option, centered such that 0 is the mean across trials. In the 1D-2O choice task (right graph), choice probability is the output of the softmax function for the yes option, centered such that 0 is the mean across trials.

E] Choice rate as a function of decision value. Decision value is simply the difference between left and right option ratings in the 1D-2O choice task, and the weighted sum of reward and effort ratings in the 1O-2D choice task. The set of decision values across trials were distributed over 6 bins. Dots are mean choice rate ('left' response in the 1D-2O and 'yes' response in the 1O-2D task). Lines show binned logistic regression fits.

Dots represent mean across participants, error bars are inter-participant standard errors.

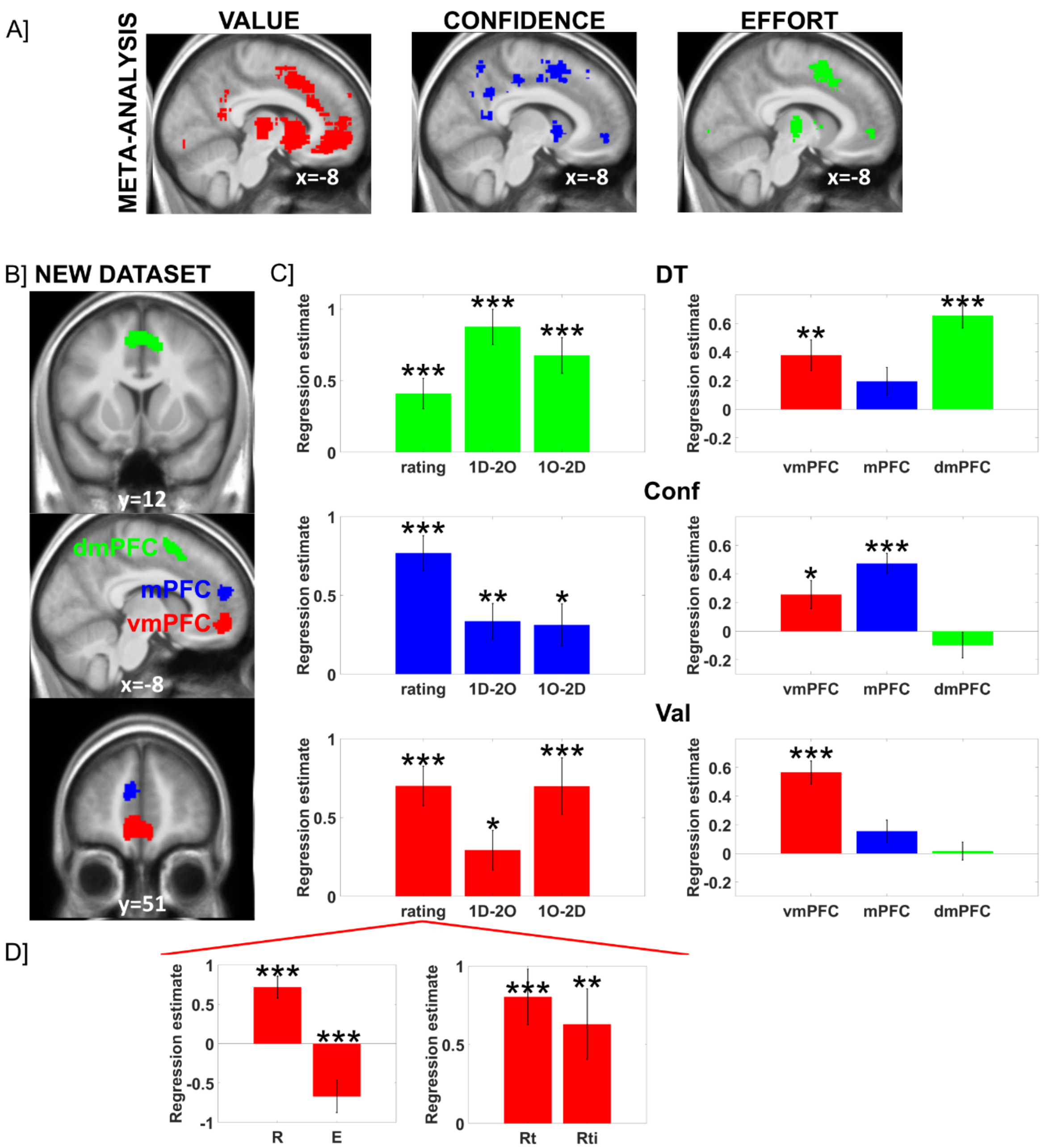

Figure 3: Neural results.

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.

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] Decomposition of regression estimates obtained for each variable of interest, per task (rating, 1D-2O and 1O-2D choice) on the left, and per ROI (vmPFC, mPFC, dmPFC) on the right.

D] Decomposition of regression estimates, obtained for Val in the vmPFC during rating, per stimulus category (reward R versus effort E, and reward presented as text + image R_{ti} versus text only R_t).

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

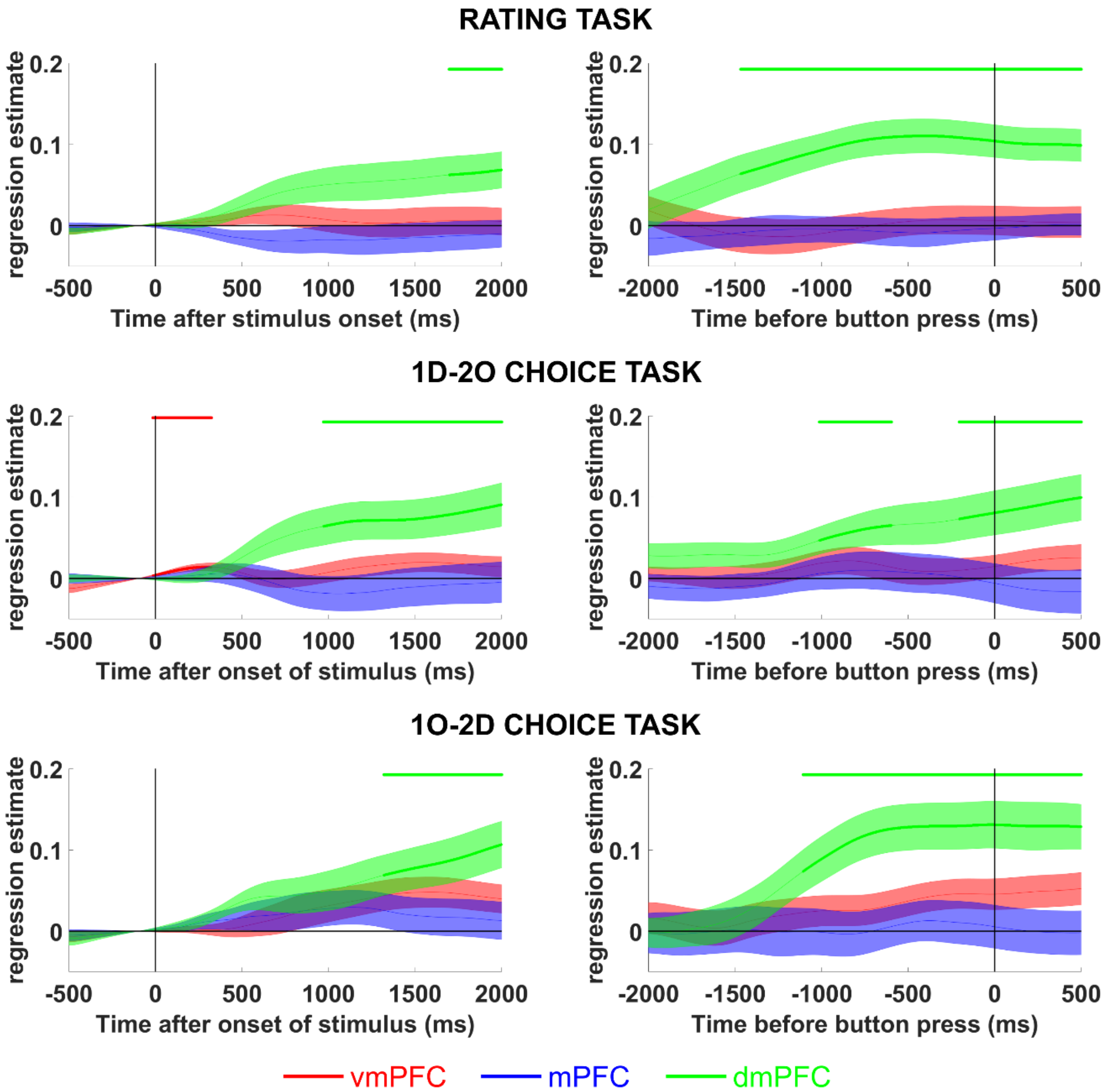

Figure 4: Pupillometric results.

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, 1D-2O and 1O-2D choice tasks). Left and right columns show time courses aligned on 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.