Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI – EEG study

Tobias Larsen and John P. O’Doherty

1Division of Humanities and Social Sciences and Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA 91125, USA
2Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin 2, Ireland

While there is a growing body of functional magnetic resonance imaging (fMRI) evidence implicating a corpus of brain regions in value-based decision-making in humans, the limited temporal resolution of fMRI cannot address the relative temporal precedence of different brain regions in decision-making. To address this question, we adopted a computational model-based approach to electroencephalography (EEG) data acquired during a simple binary choice task. fMRI data were also acquired from the same participants for source localization. Post-decision value signals emerged 200 ms post-stimulus in a predominantly posterior source in the vicinity of the intraparietal sulcus and posterior temporal lobe cortex, alongside a weaker anterior locus. The signal then shifted to a predominantly anterior locus 850 ms following the trial onset, localized to the ventromedial prefrontal cortex and lateral prefrontal cortex. Comparison signals between unchosen and chosen options emerged late in the trial at 1050 ms in dorsomedial prefrontal cortex, suggesting that such comparison signals may not be directly associated with the decision itself but rather may play a role in post-decision action selection. Taken together, these results provide us new insights into the temporal dynamics of decision-making in the brain, suggesting that for a simple binary choice task, decisions may be encoded predominantly in posterior areas such as intraparietal sulcus, before shifting anteriorly.

1. Introduction

Considerable progress has been made in uncovering the brain systems involved in encoding predictions about future rewards and in using those predictions to guide behaviour [1–5]. Studies in both humans and other animals have identified contributions for a number of brain regions in valuation, learning and choice. Within the cortex, three regions that have received particular attention are the ventromedial prefrontal cortex (vmPFC) composed of medial orbital and adjacent medial prefrontal cortex, the lateral intraparietal sulcus (LIP) and the dorsomedial prefrontal cortex (dmPFC) extending from the anterior cingulate cortex dorsally along the medial wall. These regions have been found to encode value signals for decision options and actions [6–8], as well as signals corresponding to the difference in value between actions and/or options that are ultimately chosen versus those that are not [9–11]. However, the precise functions of these regions in the decision-making process remain controversial, particularly as regards where value signals from different options are ultimately compared in order to generate a choice. One possibility that has been proposed is that the vmPFC is involved in comparing stimulus values in order to generate a decision in at least certain types of choice processes [12]. Another viewpoint suggests that the comparison between the values of possible actions in order to yield a decision over which action to ultimately select is mediated within LIP...
participated in the EEG experiment. The experiments were conducted at Trinity College Institute of Neurosciences, Trinity College Dublin, Ireland. Of those EEG participants, 35 (16 males) also participated in an fMRI study using exactly the same task. Nineteen of the 41 participants took part in the EEG experiment first, whereas the other 22 took part in the fMRI experiment first. The gap between the EEG and fMRI experiments was 13–15 days for most participants; however, for six participants, it was 20–21 days, and for two participants, it was 5 weeks.

2. Methods

(a) Participants

Forty-one right-handed participants (19 males, average age 22) participated in the EEG experiment. The experiments were conducted in a quiet dimly lit room with participants sitting in a comfortable chair. They were instructed to make a decision in favor of one of two options, which were presented on a computer screen. Each decision option had a drifting probability of reward independent of the other, and participants were instructed to weigh the probabilities of the two options in order to obtain monetary rewards. To reduce the possibility that participants make a decision in advance of the options being presented on a particular trial, three decision options were used in total, two of which are then selected at random on a given trial. Thus, a participant does not know in advance which two decision options will be presented in advance of the onset of a particular decision trial.

On each particular trial, the decision options presented on that trial are randomly presented on either the left or the right of the screen, so as to enable value signals to the decision options per se to be disambiguated from value signals pertaining to particular actions. For the purpose of obtaining a neutral baseline, on 25% of the trials, a single white target was presented, followed by non-rewarding outcome.

Each decision option has a drifting probability of reward independent of the others. To ensure that the underlying reward distributions attached to each decision option are distinguishable enough for the participants to learn independently, three sets of three drifting probabilities (one set per block of trials) were created as noisy sine waves with a period between $\pi$ and $2\pi$, and with starting points distributed evenly (for an example set, see figure 2a). The participant indicates his/her choice by pressing a button with their left or right hand. Left choices were made with the left hand, right choices with the right hand.

(b) Task description

On each trial (figure 1), the participant is presented with two easily distinguishable decision options (coloured circles on black background), one on each side of a monitor, and is tasked with making a choice between them in order to obtain monetary rewards. To reduce the possibility that participants make a decision in advance of the options being presented on a particular trial, three decision options were used in total, two of which are then selected at random on a given trial. Thus, a participant does not know in advance which two decision options will be presented in advance of the onset of a particular decision trial.

On each particular trial, the decision options presented on that trial are randomly presented on either the left or the right of the screen, so as to enable value signals to the decision options per se to be disambiguated from value signals pertaining to particular actions. For the purpose of obtaining a neutral baseline, on 25% of the trials, a single white target was presented, followed by non-rewarding outcome.

Each decision option has a drifting probability of reward independent of the others. To ensure that the underlying reward distributions attached to each decision option are distinguishable enough for the participants to learn independently, three sets of three drifting probabilities (one set per block of trials) were created as noisy sine waves with a period between $\pi$ and $2\pi$, and with starting points distributed evenly (for an example set, see figure 2a). The participant indicates his/her choice by pressing a button with their left or right hand. Left choices were made with the left hand, right choices with the right hand.
We ran two versions of the experiment. In the one version, participants were presented with a fixed outcome of 0.25 euros, on each trial in which a reward was delivered. The other version of the experiment involved variable reward outcomes, centred on 0.25, but drawn from a uniform distribution (range: 0.1–0.4). Twenty participants took part in the first version of the experiment, whereas 21 participants took part in the second version. The two experimental versions were otherwise identical. In order to maximize statistical power, and because the differences in the outcome distributions are not pertinent for the hypotheses tested in the present manuscript, we combined the data across both groups when performing the analyses and presenting the results shown here.

(c) EEG and fMRI recordings
The 512 Hz EEG data were acquired at Trinity College Institute of Neuroscience using a Biosemi 128 + 2 channel cap system with eight flat-type active electrodes (six facial, two mastoids). At the start of each recording session, each connection was stable with offsets within a ±25 mV range. Data were recorded unreferenced and unfiltered with ActiView software.

The fMRI data were acquired using a Philips Achieva 3T scanner, also located at Trinity College Institute of Neuroscience. Scan parameters were optimized to obtain robust signals in vmPFC, but also to allow whole brain coverage: 45 slices recorded at a 30° angle, repetition time (TR) = 2.5 s, echo time (TE) = 28 ms, voxel size 3 × 3 × 3.35 mm, 440 volumes for each of the three experimental sessions. In addition, a 1 mm isotropic T1-weighted structural scan was acquired for each participant to enable localization of the activations.

(d) Data analysis
(i) EEG pre-processing
All EEG datasets have been pre-processed with the FASTER-package (http://sourceforge.net/projects/faster/) which is a fully automated, unsupervised method for processing of high density EEG data as well as human experimenters. The algorithm filters the data, removes overly noisy channels, performs independent component analysis and removes artefacts, epochs the data and removes epochs that are still noisy after artefact removal. After running each dataset through the FASTER algorithm, they were referenced and unfiltered with ActiView software.

(ii) Computational model
The subjects’ behavioural data were fitted to a simple SARSA-type reinforcement-learning (RL) model with discounted value indicating the amount won alongside a small picture of some euro coins. The ‘non-reward’ screen is visually similar, but has a red cross over the euro coins. For the neutral trials, a scrambled euro picture is shown at the time of outcome presentation.

(5) An ITI follows, with a duration randomized between trials with a uniform distribution between 2 and 9 s. The fixation cross is present.
where $t$ parameter probability of choosing stimulus softmax rule and the discounting rate. The fitted parameters were learning rate, temperature for the time, and prediction error for each reward delivery (or lack thereof). choice rule yields predictions of value for each stimulus at each is the reward received at time $t$ and $V_i$ is the reward value for stimulus $i$ at time $t$.

The fitted parameters were learning rate, temperature for the time, and prediction error for each reward delivery (or lack thereof).

The correlation with the difference and chosen value is then found from the fitted reinforcement-learning. To avoid the potentially strong correlation between difference value and chosen value, the design matrix for the regression consists of chosen and unchosen values, separated into left and right choices, giving four regressors. The correlation with the difference and chosen value is then found through averaging together the appropriate regressors.

This generated parameter estimates corresponding to the degree to which the computational signals were correlated with the EEG data at that specific time point in the trial across the experiment in that particular electrode. This was repeated across all time bins, all electrodes and all participants. To obtain between subjects effects, we computed the T-score for each time bin within each electrode across subjects. We could then produce maps in electrode space plotting the degree of correlation of different electrodes with the relevant computational signals at each different time interval within the trial.

In order to derive a statistical threshold for the T-scores, we needed to correct for multiple comparisons. As Bonferroni correction assuming independence of electrode position and time window was deemed too stringent given considerable intrinsic spatial and temporal smoothing, we used a permutation test to derive an appropriate statistical threshold. To carry out the permutation test, we randomly permuted the order of the model values and reran the regression with each EEG sequence for each electrode. This way we found the threshold at which random data with the same variance show no correlation with from our computational model. Next, we used the regress function in Matlab (Matlab. 2010b The MathWorks, Inc., Natick, MA) to regress each EEG time series against the relevant signals from our computational model, which are (i) the difference in value between the chosen and the unchosen decision options as predicted by the reinforcement-learning model and (ii) the value of the option that is ultimately chosen in the trial, also generated from the fitted reinforcement-learning. To avoid the potentially strong correlation between difference value and chosen value, the design matrix for the regression consists of chosen and unchosen values, separated into left and right choices, giving four regressors. The correlation with the difference and chosen value is then found through averaging together the appropriate regressors.

This generated parameter estimates corresponding to the degree to which the computational signals were correlated with the EEG data at that specific time point in the trial across the experiment in that particular electrode. This was repeated across all time bins, all electrodes and all participants. To obtain between subjects effects, we computed the T-score for each time bin within each electrode across subjects. We could then produce maps in electrode space plotting the degree of correlation of different electrodes with the relevant computational signals at each different time interval within the trial.

In order to derive a statistical threshold for the T-scores, we needed to correct for multiple comparisons. As Bonferroni correction assuming independence of electrode position and time window was deemed too stringent given considerable intrinsic spatial and temporal smoothing, we used a permutation test to derive an appropriate statistical threshold. To carry out the permutation test, we randomly permuted the order of the model values and reran the regression with each EEG sequence for each electrode. This way we found the threshold at which random data with the same variance show no correlation with

for unchosen alternatives and a softmax choice rule

\[
V_i(t + 1) = V_i(t) + \alpha(r - V_i(t)) \quad \text{if } i = \text{chosen alternative}
\]

\[
V_i(t + 1) = \gamma V_i(t) \quad \text{if } i \neq \text{chosen alternative}
\]

\[
P_i(t) = \frac{\exp(V_i(t)/\tau)}{\sum_{j=1}^{I} \exp(V_j(t)/\tau)},
\]

where $V_i(t)$ is the value of stimulus $i$ at time $t$, $\alpha$ is the learning rate, $r$ is the reward received at time $t$ and $\gamma$ is the discount rate. $P_i(t)$ is the probability of choosing stimulus $i$ at time $t$ given the temperature parameter $\tau$. This learning model combined with the softmax choice rule yields predictions of value for each stimulus at each time, and prediction error for each reward delivery (or lack thereof).

The fitted parameters were learning rate, temperature for the softmax rule and the discounting rate.

**Figure 3.** In order to validate our analytical approach, we tested for electrodes that correlate with the generation of a motor response, and source-localized the activity to motor cortex. (a) Electrodes correlating with responses in the time window −100 to 0 ms before button press. Electrodes shown show significance at a $p < 0.05$ corrected threshold determined through permutation test. (b) EEG activity for motor responses source-localized as comparisons of left versus right choices without an fMRI-based prior probability map. (Online version in colour.)

(iii) Computational model-based EEG analysis

To implement the computational model-based EEG analysis, for each electrode separately, we first extracted EEG signals corresponding to each trial. We extracted data from 500 ms prior to the stimulus presentation, and to 1500 ms after. Then we binned this high-frequency time series into 50 ms bins by averaging the EEG signal within each 50 ms timeframe. This binning process was conducted separately for each trial per electrode. Next, for each electrode separately, we concatenated the time bins covering similar time points in peristimulus time in each trial into a sequence corresponding to the average EEG signal within that time bin on each trial of the experiment. We repeated this for each time bin. Given that there are 128 electrodes, and 41 time bins, this gave us 5248 separate sequences. For each electrode therefore, we had sequences corresponding to each 50 ms interval within the trial which could subsequently be regressed on trial by trial value
capture prediction errors at outcome time. The results from the outcome time are not included in the present manuscript and will be presented elsewhere.

Finally, motion parameters were included as effects of no-interest to account for variance in the fMRI data induced by participant motion.

(v) fMRI-based localization of EEG data
To perform this source localization, we first divided each participants’ trials into six conditions according to their appropriate model values. Then, each condition was averaged, and a contrast \([-5 -3 -1 1 3 5]\) was set up to identify electrodes where the EEG amplitude in peristimulus time were linearly increasing with increasing model values. To maximize the likelihood of finding areas that are common between subjects, we used the group inversion method described in Litvak and Friston [18] as implemented in SPM8 (Wellcome Department of Imaging Neuroscience, UCL, UK) to localize the common brain areas that increased in activity with increasing model values. The group inversion method uses a canonical mesh in combination with a hierarchical Bayesian algorithm to find the most likely distribution of sources given the observed data. Because we also obtained fMRI data in the same group of participants (barring six participants who were scanned with EEG only), we could use the fMRI data to inform the source localization of the EEG data. For this, the statistical maps, coming from the fMRI results (see the electronic supplementary material, figure S1) for a particular contrast (thresholded at \(p < 0.005\) (uncorrected)), were entered into the source localization as a prior probability. After the result of each participant was inverted, the final result was obtained by summarizing the regionally specific correlates of value that were conserved over subjects. Other inversion parameters used were: GS model, no PST Hanning, 0 Hz high-pass, 48 Hz low-pass, no solution restriction, no time-frequency contrast.

4. Results

(a) Behavioural findings and model-fitting
In order to establish whether participants had learned the task, we examined the extent to which participants chose the best available option on a given trial, both relative to the objective reward probabilities and to the modelled subjective reward probabilities. We found that for the EEG study, participants selected the objectively best option on 63% of trials, and the subjectively best option on 68% of trials, both of which is significantly better than chance at \(p < 10^{-5}\), as estimated from a Monte Carlo simulation. Furthermore, for the fMRI study, participants selected the objectively best option on 68%, and subjectively best option on 76% of trials, both of which are significantly better than chance at \(p < 10^{-5}\).

We also found that the participants in both the EEG and the fMRI study increasingly chose the better option as the value difference between options increased, both for the objectively (figure 2b) and subjectively (figure 2c) best option. We compared the model fit with both a null model and with a model without the discount parameter, and the model we used outperformed the others according to the Akaike information criterion.

(b) EEG results

(i) Motor-action-related EEG responses
To validate our analysis approach, we first tested for EEG-related activity elicited by the motor response. When
using an $F$-test to compare motor responses elicited to right-handed choices versus those elicited to left-handed choices in native electrode space, we observed significant differences in event-related potentials at 100 ms prior to motor response with right- versus left-sided scalp distributions peaking centrally (figure 3a).

Next, we applied source localization for the EEG activity without using a prior probability map generated from the fMRI data (see Methods for details). When examining the localized results for the same time window (100 to 0 ms) prior to the motor action using a non-directional $F$-test, the EEG activity was found to be localized to primary motor cortex in the vicinity of the hand area (figure 3b; Montreal Neurological Institute coordinates, left: $[-54, -20, 28]$, $z = 4.34$; right: $[56, -22, 26]$, $z = 4.44$).

(ii) Chosen values

In order to find out where in the brain the decision first emerges, we tested for chosen value signals, a type of signal that by definition must emerge after but not before a decision has been rendered. Thus, detecting where and when this signal arises in the EEG data, should enable us to ascertain where and when in the brain the decision first emerges.

To characterize when and where this signal arises, we regressed our chosen value regressor (derived from the RL-model) with each of the electrodes in each of the 50 ms time bins as mentioned above. To ensure that the correlation was exclusively with chosen value and not with the value difference (chosen minus unchosen) or otherwise, we excluded any time bins/electrodes in which there was even a modest correlation ($p < 0.2$) with the value of the unchosen alternative. In a plot of electrodes showing significant chosen-value-related responses ($p < 0.05$ corrected) analysed separately for each 50 ms window after stimulus presentation (figure 4), we found evidence for an initial emergence of the chosen-value signal predominantly over a posterior scalp location in the range of 200–450 ms following the trial onset, as well as much more weakly (albeit still significantly) in an anterior scalp location. The initially posterior-located chosen-value signal then appears to propagate forward through the brain, emerging centrally 500–650 ms after stimulus presentation, and finally emerging anteriorly in the brain 850–1150 ms after stimulus presentation.

To further aid in localization of this signal, in figure 5a–c, we plot the correlating electrodes (thresholded at $p < 0.05$ corrected) on scalp maps for each time-period showing significant correlations. In addition to showing the posterior-to-anterior trend, the plots also indicate that the activity at 200–450 ms is strongest medially, the activity at 500–650 ms is more lateral and that the activity at 850–1150 ms after stimulus presentation has both a strong medial and a strong lateral component.

This posterior-to-anterior progression of chosen values is also seen by dividing the trials into high and low chosen value trials, and then plotting the scalp maps of the difference wave (see the electronic supplementary material, figure S2).

Next, we implemented source localization for the chosen-value signals using SPM8. For this, we used the fMRI data acquired from the same task as a prior probability map for the source localization. The fMRI results revealed significant chosen-value-related activity in areas reported in a number of previous fMRI studies, including lateral parietal cortex, lateral orbitofrontal and vmPFC (see electronic supplementary material, figure S1).

Source-localized results revealed activity in four principle locations for chosen-value: a region of intraparietal sulcus

Figure 5. Results of the analysis of the chosen value (a–c). Same data as in figure 4, but aggregated over time windows and plotted on scalp maps. Top row shows transverse view, bottom row coronal view. Bigger dots mean more significant time bins in the time window. (Online version in colour.)
In the present study, we used computational model-based EEG analysis in combination with model-based fMRI data acquired from an overlapping group of participants in order to ascertain the timing and localization of decision-related variables as estimated through a reinforcement-learning model.

In the present study, we used a very simple type of reinforcement-learning algorithm (SARSA) to estimate trial-by-trial value signals [19]. We note that this class of ‘model-free’ algorithm does have limitations, in particular when it comes to situations where a decision problem has higher-order structure, or where the value of an outcome to an agent changes across time [20,21]. A number of other types of algorithm have been proposed for this situation, including model-based reinforcement learning and Bayesian models [20–23]. However, in the present situation, the task was designed so that value signals could be adequately captured with even a simple reinforcement-learning model, because the reward distributions associated with each action were kept independent, and can thus be learned relatively efficiently by means of a ‘model-free’ reward prediction error. Furthermore, model-based and model-free RL algorithms will likely make very similar trial-by-trial predictions in the present case. Future extensions of this work could involve using more complex tasks in order to distinguish the temporal signatures of ‘model-based’ and ‘model-free’ value signals.

The results of this study provide several important new insights into how simple binary value-related decisions are made in the human brain. Specifically, we show that chosen value signals, which are by definition a consequence of the decision process [5,10], appear to initially emerge predominantly in a posterior location in the brain, with a weaker signal present also at an anterior locus. Using fMRI-informed source localization, we found that the posterior signal was localized to the vicinity of the lateral–intraparietal cortex and posterior lateral temporal lobe cortex. Such signals emerge as soon as 200 ms after the trial onset, suggesting that the decision itself may be computed very early on in the trial. The predominantly posteriorly localized chosen value signal then appears to propagate more anteriorly over the course of several

5. Discussion

In the present study, we used computational model-based EEG analysis in combination with model-based fMRI data acquired

**Figure 6.** fMRI-informed source localization of chosen value with threshold set at $p < 0.0005$ (unc). (Online version in colour.)

((−24, −46, 64), $z = 3.58$, figure 6 top), middle temporal gyrus ((64, −20, 2), $z = 3.90$), lateral prefrontal cortex ((48, 36, 6), $z = 4.23$) and vmPFC ((−6, 62, 4), $z = 5.18$, figure 6 bottom).

(iii) Value difference

Next, we tested for EEG signals correlating with the difference in values between the options that are chosen and not-chosen on each trial. In previous studies by our group, we have found neural correlates of the difference between the unchosen and chosen value in a region of dmPFC, which we suggested represented the outcome of the decision process as predicted by models of decision-making such as the drift–diffusion model [10,11]. Consequently, we aimed to test for the presence of such a signal in the EEG data across electrodes and time post-trial onset. We found that this signal appears to emerge predominantly 1050–1200 ms after trial onset in a central scalp location (figures 7 and 8). The source localization of the unchosen minus chosen value (figure 9) revealed significant effects of the value difference in dmPFC, most strongly in a more posterior part of the dmPFC ((8, 16, 62), $z = 5.93$), but also extending more anteriorly ((−24, 54, 22), $z = 4.12$).

There are two very important features of these value difference results. First, the timing; this value difference signal appears to emerge substantially later in time than does the value chosen signal (approx. 850 ms later). Given that the value chosen signal can only emerge as a consequence and not as a precursor of the decision process, the decision clearly must be made substantially earlier within the trial before the emergence of the value difference signal. As a consequence, it appears that this signal may not be critically related to the formation or immediate aftermath of the decision process itself, but instead must relate to some post-decision process. Second, the localization of the value difference most prominently in the dmPFC in the fMRI-constrained EEG data, is strongly consistent with the results of a number of previous fMRI studies that have localized this signal to the same region of dmPFC [10,11].
hundred milliseconds, emerging in a central location around 500 ms, and then finally in a more anterior location around 850 ms post-trial onset. In our fMRI-guided EEG source localization, these more anterior areas may correspond to the lateral prefrontal cortex and vmPFC, respectively.

These findings show that the consequence of the decision itself (for at least this type of simple reward-related choice) is represented very early on (only 200 ms after stimulus onset) most prominently in posterior parts of the brain such as LIP. A possible implication of these findings could be that posterior brain regions such as LIP are responsible for the formation of the initial decision, before this information is propagated anteriorly for the purposes of action selection. However, the fact that the chosen-value activity was also found in an anterior scalp location early on in the trial, albeit much more faintly than the posterior locus, suggests that such a conclusion may be premature on the basis of the current findings.

A more balanced interpretation would be that the decision process may emerge as a result of integrative interactions between anterior and posterior regions. It has been proposed that the decision process can be viewed in a hierarchical setting, with more anterior regions such as vmPFC playing a role in the selection or ranking of goals, and more posterior dorsal regions such as LIP playing a role in the selection of actions to obtain goals (see [5]). The degree to which these different regions contribute to the decision process for a given decision-making problem may depend on the extent to which the decision that needs to be made is weighted towards goal selection or action selection. For example, in a situation where the goal is the same over all the actions (i.e. to win money), we might expect less contribution from vmPFC to the decision process and more contribution of LIP and other dorsal cortical areas. Conversely, if the decision task involves selection over different types of goal outcomes (e.g. deciding whether to opt for a glass of soda or glass of milk), where each goal involves similar actions (i.e. a grasping movement), this type of decision might depend more on the vmPFC. Decisions that tax both goal selection and action selection might depend heavily on both brain areas. In the present case because the goal was the same for both actions (to win money), yet the actions differed and the probability of reward available on those actions differed, perhaps a greater initial dependence on posterior regions for computing the decision might pertain, as observed in our results.

The rapid emergence of chosen value signals in the vicinity of the lateral intraparietal cortex is particularly notable given prior literature implicating this region in decision-making in non-human primates [7,13,24–26]. In particular, neural signals have been found in these regions that are suggested to correspond to evidence integration processes in both perceptual and value-based decision-making [6,27–29]. It is also striking that the timing of emergence of decision signals in monkey LIP neurons is very compatible with the timing of the emergence of chosen value signals in the present study. In both the monkey studies and our study, around 200 ms post-stimulus, LIP appears to have a very robust representation of the option that ultimately is chosen.

Besides the lateral intraparietal cortex and vmPFC, another region that has previously been implicated in the decision process is the dmPFC [4,10,11,30]. This region has previously been found to encode the difference between unchosen and chosen values, which has been hypothesized to emerge as a consequence of the type of evidence integration process that may underpin decision-making at the neural level. When we tested for unchosen minus chosen value signals we found them to be localized predominantly to regions of dmPFC, as found in those previous studies. However, these signals appear to emerge 1050 ms into the trial, approximately 850 ms after the chosen values signals

![Figure 8](http://rstb.royalsocietypublishing.org/)  
**Figure 8.** Same data as in figure 7, but aggregated over the time window showing significant correlation and plotted on a scalp map. Top shows transverse view, bottom coronal view. Bigger dots mean more significant time bins in the time window. (Online version in colour.)

![Figure 9](http://rstb.royalsocietypublishing.org/)  
**Figure 9.** fMRI-informed source localization of value difference, with threshold set at $p < 0.0005$. (Online version in colour.)
are first observed elsewhere. A strong interpretation of these results is that value difference signals in the dmPFC are not involved in the implementation of the decision process itself, but instead emerge later on in the trial at the point when action selection is being implemented. One possibility, on which we speculated in the original fMRI study by Wunderlich et al. [10] that first reported this signal, is that the dmPFC may contribute to the implementation of the action selection by inhibiting selection of the action that is not taken on a particular trial.

In addition to the late emergence of value-difference signals, it is also notable that chosen value signals also occur mid to late in the trial (in the range of 850–1050 ms post-trial onset), localized predominantly to anterior brain regions such as vmPFC and lateral prefrontal cortex. What purpose could such late emerging signals serve? One possibility is that such chosen value signals are used in order to generate prediction errors, which correspond to the difference between actual and expected (chosen) outcomes [5]. Alternatively, chosen values might be used to motivate action selection, i.e. to increase response-vigour towards higher-valued actions relative to chosen actions that are less highly valued. Finally, chosen value signals could be used to modulate action-monitoring, so that actions leading to more highly valued outcomes are monitored more closely than actions resulting in less valuable outcomes.

One important caveat of our findings is that the fMRI constrained localization results report here does depend on the assumption that fMRI data and EEG data are generated by the same underlying neural sources, an assumption that may not always hold true [31]. Nevertheless, the brain regions identified in our source localization have been previously identified as having neural activity that is strongly implicated in value-related learning using direct neural recordings from those areas [7,8,32]. Thus, it is unlikely that the results we observe are an artefact of the assumption underlying fMRI-informed EEG localization.

In conclusion, the present findings provide us with valuable insights into the temporal and spatial processes underpinning human value-based decision-making for binary options. We show that value signals related to the consequence of the decision emerge within 200 ms after onset of the decision trial, most predominantly in posterior locations in the vicinity of the LIP as well as in posterior temporal lobe cortex, but also less prominently in an anterior cortical locus. These findings suggest an important role for posterior brain regions including the LIP in the generation of the decision itself, perhaps through cortical interactions with more anterior regions such as vmPFC. Furthermore, while dmPFC was found to report the difference in value between unchosen and chosen options as in a number of previous studies, this signal did not emerge until much later in the trial, approximately 1050 ms after the initial stimulus presentation, and 850 ms after the first emergence of the value signals reflecting the consequence of the decision (chosen value signals). These findings suggest that dmPFC may not play a consequential role in the initial formation of the decision itself, but rather may contribute at a later stage during the process of implementation of action selection.

Acknowledgements. The authors thank Simon Dunne and Teresa Furey for their help running the experiment, and Dr Paul Dockree for his advice and suggestions.

Funding statement. This work was supported by grants from Science Foundation Ireland and the Gordon and Betty Moore Foundation to J.O.D.

References


