The human prefrontal cortex mediates integration of potential causes behind observed outcomes.

Klaus Wunderlich[1,3*], Ulrik R Beierholm[4], Peter Bossaerts[1,2,5] and John P O'Doherty[1,2,6]

[1] *Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA*

[2] *Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA*

[3] *Wellcome Trust Centre for Neuroimaging, UCL, London, UK*

[4] *Gatsby Computational Neuroscience Unit, UCL, London, UK*

[5] *Ecole Polytechnique Federale de Lausanne, Swiss Confederation*

[6] *Trinity College Institute of Neuroscience and School of Psychology, Trinity College, Dublin, Ireland*

[*] To whom correspondence should be addressed at:

Wellcome Trust Centre for Neuroimaging
12 Queen Square
London, WC1N 3BG
United Kingdom
Email: kwunder@fil.ion.ucl.ac.uk

Abstract

Prefrontal cortex has long been implicated in tasks involving higher-order inference in which decisions must be rendered not only about which stimulus is currently rewarded but also which stimulus dimensions are currently relevant. However, the precise computational mechanisms used to solve such tasks have remained unclear. We scanned human participants with fMRI while they performed a hierarchical intra-dimensional / extra-dimensional shift task to investigate what strategy subjects use while solving higher order decision problems. By using a computational model-based analysis we found behavioral and neural evidence that humans solve such problems not by occasionally shifting focus from one to the other dimension, but by considering multiple explanations simultaneously. Activity in human prefrontal cortex was better accounted for by a model that integrates over all available evidences than by a model in which attention is selectively gated. Importantly, our model provides an explanation for how the brain determines integration weights, according to which it could distribute its attention. Our results demonstrate that at the point of choice, the human brain and the prefrontal cortex in particular is capable of a weighted integration of information across multiple evidences.

Introduction

The human prefrontal cortex has long been implicated in the ability to solve decision problems that contain a hierarchical dimensional structure, in which inferences must be rendered over different levels of a stimulus hierarchy (Koechlin et al. 2003). This type of decision problem is typified in the Wisconsin Card Sorting Task (Grant and Berg 1948) and its modern antecedents such as the intra-extra dimensional set shifting task (Downes et al. 1989). In such tasks, a participant is presented with stimuli that contain multiple dimensions such as color and shape, and at any one moment in time, one of these dimensions is causally linked to reward. However, within the "relevant" dimension, only one out of a set of possible within-dimension features is currently rewarded (e.g. if color is relevant, then either "red" or "green" might at a given point in time be rewarded). The decision problem faced by the participant is therefore two-fold: to identify which stimulus dimension is currently relevant, and within a given dimension to identify which feature is currently rewarded. This task can become especially challenging if the relevant dimensions and features change over time, as the participant must constantly update his/her inferences at each level of the decision hierarchy.

It is well established that lesions to the prefrontal cortex impair the capacity to perform such hierarchical decision problems (Dias et al. 1996; Drewe 1974; Milner 1963; Robinson et al. 1980). However, the specific computations underlying the capacity of prefrontal cortex to solve such decision problems are much less well understood. This kind of problem could be solved by one of two strategies: an attention-gated strategy, in which the agent first identifies which dimension is most likely to be relevant at a given point in time. The agent then focuses exclusively on what is considered to be the relevant dimension, deciding separately which feature is likely to be rewarding. The advantage of such an attention-gating approach is that by focusing only on the dimension for which there is the strongest evidence, there is no need to incorporate evidence for the features in the less compelling dimension in the decision process, thereby imposing less demand on limited cognitive resources. An alternative integration strategy is to simultaneously evaluate evidence from both levels of the hierarchy weighted by how much this dimension is judged to be relevant. In other words, even if color is deemed very likely to be relevant, information about which feature could be correct if shape were the relevant dimension is also taken into account to some extent, and used in the decision over which compound

stimulus ultimately to select. This type of "integration" strategy typifies Bayesian analysis (Berger 1980).

Here we aimed to address whether during the performance of a hierarchical reversal learning task, the human brain follows an attention-gated approach to problem solving or rather the integrative procedure whereby less likely explanations for observed phenomena still influence decisions. Our task is a probabilistic variant of a Wisconsin card-sorting task in which each stimulus possessed attributes along two dimensions: color and motion. Within each dimension there were two features, so that color was red or green, while motion (indicated by moving dots) was either leftward or rightward. At any one point in time, one dimension e.g. color was "relevant" while within this dimension one of the two features was correct (e.g. green), in that choice of a stimulus possessing that feature yielded monetary reward with a high probability and monetary loss with a low probability, while choice of the other stimulus yielded only a low probability of reward and loss with a high probability. We explicitly imposed a hierarchical structure between dimension and feature on this task by reversing the correct feature at a higher rate than the relevant dimension.

Human subjects were scanned with fMRI while they participated in this probabilistic hierarchical decision task. We then used a variety of computational models in order to account for subjects' behavioral performance on the task. After confirming that subjects used a hierarchical strategy over simply model-free learning of either stimulus or feature values, we compared two possible models: one that implemented an attention-gating strategy, and another that implemented an integration strategy. We then applied these models to the fMRI data in order to uncover brain regions in which activity at the time of decision-making reflected one or other of these computational strategies. In order to discriminate which of these models was contributing to the decision process, we used the value-signals estimated by each of the model and tested for regions of the prefrontal cortex correlating with such signals. We focused in particular on the ventromedial prefrontal cortex as this region has previously been found to encode value signals at the time of decision making (Boorman et al. 2009; Daw et al. 2006; FitzGerald et al. 2009; Hampton et al. 2006; Hare et al. 2008; Knutson et al. 2005; Wunderlich et al. 2009). We compared and contrasted the capacity of the integration and attention-gating models to capture value-related computations in this region of prefrontal cortex. We hypothesized that activity in

ventromedial prefrontal cortex would be best captured by the value-signal derived from the model that best explained participant's behavioral choice patterns.

Materials and Methods

Subjects

16 healthy subjects (drawn from the Caltech student population; 6 female; 18–28 years old) with no history of neurological or psychiatric illness participated in the study. The study was approved by the Institutional Review Board of the California Institute of Technology and all subjects provided informed consent before their participation.

Task description

The task is a variant of a hierarchical intra/extra-dimensional (ID/ED) shift task. In any given trial, subjects were presented with a choice between two stimuli, each of which consists of a feature in the motion dimension and one feature in the color dimension. The two features belonging to the same stimulus are shown horizontally next to each other, grouped together by a bone-shaped structure in the background (Fig. 1A). The motion dimension (always presented on the left side of the stimulus) consisted of a moving dot sequence where the dots moved either to the left or the right; the color dimension could take on either a red or green color filled rectangle. The features for the upper stimulus were assigned pseudo randomly in each trial and converse features were assigned to the lower stimulus, i.e. if green and right motion were assigned to the upper stimulus then the lower stimulus contained red and left motion. We implemented a constraint that identical color/motion pairings within the same stimulus did not occur more than two times in a row to avoid strings of trials in which subjects cannot associate the outcome unambiguously to a chosen motion or color feature. Subjects chose the upper or lower stimulus by pressing one of two distinct buttons on a button box with their right thumb.

At any given time, only one stimulus dimension (color or motion) was causally linked to reinforcement and subjects were rewarded at a higher rate if they selected the stimulus that contained the correct feature (either red or green for color and leftward or rightward for motion)

based on the currently relevant dimension, i.e. one dimension was "relevant", and within that dimension a particular feature was correct. For example, if "color" was the relevant dimension then within color, "green" may have been correct. Choice of the stimulus that had the correct feature yielded monetary rewards on a probabilistic basis with 80% probability, whereas selection of the other stimulus yielded reward with only 20% probability. Rewarded trials yielded a prize of 25 cents, while the other trials resulted in a loss of 25 cents (there were no neutral outcomes). This trial outcome, i.e. whether subjects gained or lost money on this trial, was the only feedback provided. Note that due to the probabilistic contingencies subjects may occasionally see a loss outcome (-25 cents) after they chose correctly, or see a rewarding outcome (+25 cents) after they chose incorrectly. After subjects chose the stimulus with the correct feature three times in a row (indicating that they learned the relevant dimension and correct feature), there was a 50% probability in each further trial that the correct feature would switch. Furthermore, after a variable number of such within-category switches (1-4; uniformly distributed), the relevant dimension also switched. The adaptive switching rule used here and previously (Hampton et al. 2006) served two purposes: it created a reasonable degree of unpredictability for the next switch and prevented feature switches from occurring before the subject detected which feature is correct. Such a situation is to be avoided, as it would no longer provide us with the desired dynamic range in certainty and value estimates in the imaging analysis. The total number of trials in the fMRI experiment varied across subjects due to these probabilistic switching rules and differences in how fast individual subjects learned the correct response. However, every subject experienced the same number of intra and extradimensional switches. One experiment always contained 9 dimensional switches (5 motion & 5 color blocks) and 30 feature switches (uniformly distributed and pseudorandomly arranged). At the end of the experiment subjects were paid a flat amount of 25$ plus their accumulated earnings (in total 26 – 41$, sd = 4.3).

This design imposes a hierarchical structure on the task with feature reversals occurring on a faster timescale than the dimensional switches. As in many real-life problems, participants in our experiment have more than one theory ("which is the relevant dimension?") to go by in any trial, and corresponding characteristics ("which is the correct feature?") on which to base their decisions between the upper or lower stimulus. Subjects were instructed that at any given time only one dimension was relevant for determining reward and within that dimension one feature

was correct. We told subjects that choice of the stimulus containing the correct feature would yield reward with a high probability and choice of the other stimulus with a much lower probability. Subjects were also provided with the information that the correct feature would change after a number of trials and that the relevant dimension would change at a slower timescale than the correct feature, i.e. the problem was specifically posed as a probabilistic hierarchical reversal task. However, subjects were not given any further explicit information about the exact reward probabilities or about the mechanism of how changes in correct features and dimension were realized. It is also important to note that participants never got direct feedback about whether their theory was currently correct, but only whether their choice proved successful or not. Immediately preceding the fMRI scan subjects underwent a series of practice blocks in which we familiarized them with the task by gradually increasing its complexity. This training consisted of three blocks: first, subjects worked on a simple reversal task with only one modality. Next, we introduced the second dimension but indicated every dimensional switch by an auditory signal. Finally, subjects practiced the task without the help of the auditory signal. Subjects experienced four dimensional switches during this stage, which happened according to the same rules as during the following fMRI experiment. As at all times reversals only progressed after subjects chose consistently the correct stimulus, this rule imposed a performance criterion before subjects progressed to the fMRI stage.

The task was presented via back projection on a translucent screen, viewable through a headcoil mounted mirror.

Algorithms

We compared a number of computational strategies for how subjects would solve this task and will focus our investigation on the comparison of two hierarchical strategies, a selection versus an integration strategy.

Attention-gating model

The first model in this class, which we will call an attention-gating strategy, computes a decision in a two step procedure by first evaluating which dimension is currently relevant, and then

allocating attention to only that dimension, before subsequently working out which feature within the dimensional category is currently reinforced (**Fig. 1B**). Dimension relevance is evaluated using a reinforcement learning mechanism that integrates over the previous rewards obtained when selecting that dimension (Sutton and Barto 1998).

In decision theoretic terms, this type of strategy is called maximizing, because probabilistic information about which feature is currently correct is used to guide choice only over the dimension deemed currently relevant (and hence gated by selective attention). For instance, if the model decides color is relevant then choice is computed by taking into account the probability that within color, red is correct or that green is correct, but probabilistic information over the unselected dimension of motion is ignored. The selection strategy relates to complexity reducing heuristics wherein choices are proposed to be taken using the best available evidence (Gigerenzer and Goldstein 1996).

Integration model

While the attention-gating model maintains only information deemed relevant, an alternative approach is to distribute attention and fully integrate over all of the probabilistic information available to the individual at that moment in time, and use this complete information to guide choice (**Fig. 1C**). In other words, even if color is deemed highly likely to be the relevant dimension (i.e. color has a high probability), the model not only takes into account the probability that red or green is correct, but also uses the information it has from the less likely motion dimension about which movement direction might be correct weighted by the low probability of motion being relevant. In decision theoretic terms such an integration strategy is formally called marginalization, and is one important step into the direction of the full Bayesian approach (see the supplementary materials). According to Bayesian theory (Berger 1980), an optimal estimate results from combining information across dimensions with evidence from past trials. The available information is integrated based on weights that are assigned to the dimension and correspond to the likelihood that the dimension is relevant (hereby no theory, even the least likely, is ever excluded). This issue of whether to commit to a theory is frequently encountered in the machine learning domain where it is well known that integrating over

parameters is optimal (MacKay 1999). The difference between this integration model and a full Bayesian approach is that the latter requires a change in the updating scheme from prediction error based to likelihood ratio based (Yang and Shadlen 2007). We also implemented a full Bayesian model (supplementary methods), which, in a test against behavior, did not result in any significant improvement of model likelihoods compared to the integration model.

Hierarchical model implementations

For modeling the data we assumed that subjects learned the relevant values assigned to the different stimuli on the basis of trial by trial experience using prediction error based updating (Sutton and Barto 1998). If stimulus *s* is selected on trial *t*, the value of its two features *a* (one from each color and motion dimension) are updated via a prediction error, $\delta(t)$, as follows: $V_a(t+1) = V_a(t) + \alpha^i \delta(t)$ where α^i is a learning rate between 0 and 1. The prediction error $\delta(t)$ is calculated by comparing the actual reward received, *r(t)*, with the reward that the subject expected to receive from that action in that trial; that is, $\delta(t) = r(t) - V_a(t)$. Specifically for this task, three variables had to be learned and updated in each trial: V_{green}, V_{right} and W_{color}, keeping track respectively of the value of the green stimulus (versus red stimulus), rightwards motion stimulus (versus left stimulus) and the weight of the color versus motion modality. We assume that V_{red} = -V_{green}, V_{left} = -V_{right} and W_{motion} = -W_{color}. Hence, the updating can be done with the inverse prediction error $-\delta(t)$ if the complementary action is chosen.

In each trial, a subject chooses either UP or DOWN, and thus selects a combination of one color feature (red or green) and one motion feature (left or right). The values of these 'intermodality' choices are updated in each trial using the following RL scheme:

$\delta_i(t) = r(t) - V_i(t)$, where i=red (green) if red (green) was chosen

$\delta_j(t) = r(t) - V_j(t)$, where j=right (left) if right (left) was chosen.

The dimensional weight is updated according to:

$W_c(t+1) = W_c(t) + \alpha^e \delta_c(t)$ with $\delta_c(t) = r_t \left(V_i(t) - V_j(t) \right) - W_c(t)$ where α^e is the learning rate for the dimensional weight. The dimensional weight is increased if the difference in expected value for each of the two modalities was larger than the expected dimensional weight. In terms of the

typical RL model the dimensional weight is tracking the ability of one dimension to predict reward, relative to the other.

Using these learned values we compared two ways to generate choices:

1. *Selection*: Determine which dimension D has the highest expected value, W_C or $W_M =-W_C$. The decision whether the top or bottom stimulus is chosen is then based on that modality only by choosing the stimulus which contains the feature with the higher value within dimension D. Given choice between $UP=\{V_{i1},V_{j1}\}$ and $DOWN=\{V_{i2},V_{j2}\}$ calculate $V_{UP}=V_{i1}*H(W_C)+V_{j1}*H(-W_C)$ where H is the Heaviside operator, $H(+)=1$, $H(-)=0$.

2. *Integration*: The value for choosing the top or bottom stimulus is calculated as a weighted sum of both feature values in each stimulus. The integration weights are determined by the current dimensional weight, i.e. weight the two features by how well their dimension has performed. The value for UP is then $V_{UP}=V_{i1}*w_{color}+V_{j1}*(1-w_{color})$ where $w_{color}=exp(\kappa*W_{color})/(exp(\kappa*W_{color})+exp(\kappa*W_{motion}))$. Using the definition of sigma from below this last equation simplifies to $w_{color} = \sigma(2*\kappa*W_{color})$.

We used a softmax procedure to generate choices where in every trial the probability (P) of choosing action $a \in \{UP,DOWN\}$ over b is given by: $P_{a,t}=\sigma(\beta(V_a(t)-V_b(t)))$, where $\sigma(z)=1/(1+e^{-z})$ is the Luce choice rule or logistic sigmoid, and β determines the degree of stochasticity involved in making decisions. We fit the parameters for each model and each subject (learning rates α^i, α^e, softmax β, as well as the extra parameter κ for the integration model) such that the model best explained subjects' choices (maximum likelihood, with likelihood $l=\prod_t P_{x_t,t}$ where x_t is the actual choice in trial t).

Non-hierarchical-RL models

In order to account for the possibility that subjects solved the task without considering the hierarchical structure at all, we considered various different non-hierarchical RL models.:

(1) A compound stimulus-based RL model. This model learns a value per stimulus *s* (the actual composite stimuli, not individual features). That is by separately learning the values of the four composite stimuli using a Rescorla-Wagner learning rule (Rescorla and Wagner 1972) for $V_{RightRed}$, $V_{RightGreen}$, $V_{LeftRed}$, $V_{LeftGreen}$. In each trial, the values of the two presented stimuli are considered for choice and only the value of the chosen stimulus is updated via a prediction error. In this model not only the information on hierarchical structure but also information about dimensional grouping are neglected.

(2) A feature-based RL model in which each of the 4 features (red, green, left, right) are considered to be separate items for the RL model to learn about (i.e. V_{red} and V_{green} are not assumed anti-correlated) and furthermore V_{Color} is also kept at 0 with the choice therefore based on a sum of the two features encountered within each stimulus (e.g. for UP={left,green}, $V_{UP}=(V_{left}+V_{green})/2)$. Thus after the choice (DOWN={red, left}, no updating would be done for V_{green} or V_{right}.

(3) A 1 layer version of the hierarchical model above, where V_{Color} is always kept at 0 and hence there is never any information with regard to which modality is more likely to be correct. This is equivalent to the special case of the Integration model with κ=0. Such a learner mixes the dimensions equally (e.g. for UP={left,green}, $V_{UP}=(V_{left}+V_{green})/2$ and since $V_{green}=-V_{red}$ and $V_{right}=-V_{left}$, for our simple setup effectively just chooses option with the larger value of the four stimuli {V_{red}, $V_{green}=-V_{red}$, V_{left}, $V_{right}=-V_{left}$ }.

All other implementation details of models 2 and 3 were otherwise identical to the hierarchical models above. We used a softmax rule to generate choices and fitted the 2 parameters (learning rate α^i, softmax scaling β) by maximizing the likelihood of the subject choices.

Behavioral model comparison

To compare the behavioral fits for the different models while accounting for differences in model complexity we report the Bayesian Information Criterion (BIC) (Burnham and Anderson 2002) in Table S1, which corrects for the number of parameters, *k*, in a model based on the number of data points *n*: $BIC=-2*\log(l)+k*\log(n)$. The model with the lower *BIC* explains the subjects' behavior better. For the model comparison between integration and attention-gating we can also calculate an approximation to the log Bayes Factor (Kass and Raftery 1995) as log BF=-$0.5*(BIC_{int} - BIC_{select})$. The Bayes Factor specifies the ratio of marginal model likelihoods and a probability that the attention-gating model is more likely is thus given as $P_{attgated} = exp$(-log BF). This allows us to assign a probability to our null hypothesis that the attention-gating model is more likely and, in analogy to a p-value in statistical testing, reject the attention-gating model if $P_{select}<0.05$. This criterion is met if the log BF > ~3 (i.e. BF > 20). To allow inference on the population we used a pairwise Bayesian model comparison (Stephan et al. 2009) between all our tested models using the BIC corrected likelihoods as model evidences.

Note that we base our model comparison on log evidences of the models instead of the fraction of correctly predicted choices. This method is superior to comparing the number of correctly predicted choices (which are calculated from binary data) because such data introduce noise into the system when the model predictions are passed through softmax and fails to take into account parametric variations in the model predictions (Daw 2011).

FMRI data acquisition

Data were acquired with a 3T scanner (Trio, Siemens, Erlangen, Germany) using an eight-channel phased array head coil. Functional images were taken with a gradient echo T2*-

weighted echo-planar sequence (TR = 2.65 s, flip angle = 90°, TE = 30 ms, 64 × 64 matrix). Whole brain coverage was achieved by taking 45 slices (3 mm thickness, no gap, in-plane resolution 3 × 3 mm), tilted in an oblique orientation at -30deg to the AC-PC line to minimize signal dropout in OFC. Subjects' head was restrained with foam pads to limit head movement during acquisition. A high resolution T1-weighted anatomical scan of the whole brain (MPRAGE sequence, 1x1x1 mm resolution) was also acquired for each subject.

FMRI data analysis

Imaging analysis was performed using SPM5 (Wellcome Trust Centre for Neuroimaging, Institute of Neurology, London, U.K.). Images were first slice time corrected to TR/2, realigned to the first volume to correct for subject motion, spatially normalized to a standard T2* template with a voxel size of 3mm, and spatially smoothed with an isotropic Gaussian kernel of 8mm FWHM to account for anatomical differences between subjects and to allow for valid statistical inference at the group level. Intensity normalization and high pass temporal filtering (using a filter width of 128s) were also applied to the data.

First, we estimated for each individual subject a GLM for the attention-gating model and separately another GLM for the integration model, differing only in the model predicted parametric modulator values. Two events were modeled in each trial: the time of the stimulus presentation, parametrically modulated by the variables described below, and the time of the presentation of the outcome, modulated by the binary outcome (+1/-1).

Parametric modulators

The regressor at the time of choice was parametrically modulated by the following decision variables (in the two separate GLMs, the parametric modulators contained values from the respective model, selection or integration):

(1) Stimulus value of the better choice (max(V_{UP}, V_{DOWN}))

We reasoned that if subjects make a value based decision between the upper or lower stimulus, we would necessarily find a representation of a decision variable in the brain that encodes the

stimulus value. This value differs between our two models: in the case of the integration model this value is a weighted combination of the two feature values within the stimulus, and in the attention-gating model it represents the feature value of only the dimension that subjects deem more likely (as indicated by the dimensional weight in our model).

We also considered alternative representations of the stimulus values: one possibility are chosen values, which are stimulus values V_{UP}, V_{DOWN} weighted by subjects' actual choices. We tested for chosen values in a separate GLM in which we replaced the max stimulus value by a chosen value modulator. In this test we find similar results as for the max stimulus value (although slightly weaker effect size) but as our study was not designed to distinguish between these highly correlated value signals this difference was not further explored. Another possibility is that the brain might encode the value difference between the top and bottom stimuli instead of the best stimulus. However, in our present task the values of the two stimuli are mutually dependent on each other, i.e. $V_{down} = 1-V_{up}$. A value difference between the higher and lower valued stimuli ($|V_{up} - V_{down}|$) is therefore in our case identical to a scaled representation of max(V_{up}, V_{down}). Another alternative would be a representation of the choice probability $P_{up}=\sigma(\beta(V_{up}-V_{down}))$, i.e. the value passed through the sigmoidal softmax function. Applying such a sigmoid essentially asks to what degree subjects use the available information: two subjects could have the exact same information but utilize it differently due to differences in the steepness of the softmax. Thus, the sigmoid transformed variables encode the usage of the available information rather than the available value. As the inverse temperatures (β, κ) for most of our subjects are close to (or smaller than) 1, the sigmoidal transformation leads to an almost linear transformation over a large part of the input space and therefore value and the sigmoid transformed variable will be highly correlated anyways.

(2) Within-dimensional certainty for color (max(V_{red}, V_{green})), and (3) motion (max(V_{right}, V_{left}))

We defined within-dimensional certainty as confidence about one feature being correct. A high value for one feature indicates that choice of a stimulus containing this feature was repeatedly rewarded. On the other hand, a feature value close to 0 means that in the recent past neither feature within this dimension was rewarded more often than the other and we expect that in this situation subjects have high uncertainty about which feature within this dimension is currently

correct. We can calculate certainty as argmax over the feature values because V_{red} and V_{right} fluctuate within the range [-1, 1] and $V_{green}=1-V_{red}$ ($V_{left}=1-V_{right}$).

We also tested for neural signals related to uncertainty, which we defined in our case as 1-certainty. Uncertainty is at maximum when subjects assign equal value to the two features (e.g. V_{red} and V_{green}), whereas the lowest uncertainty occurs when one value is much larger than the other (e.g. large V_{red}). We also considered alternative definitions for uncertainty, specifically the entropy $V_{red}*V_{green} = V_{red}*(1-V_{red})$. In practice this signal is highly correlated with 1-certainty and when we run an additional GLM with the entropy based uncertainty signals we found very similar activations to the ones reported here.

(4) Across-dimensional certainty ($\max(W_{MOTION}, W_{COLOR})$)

The across dimension certainty was similarly defined as the relative ability to predict reward based on the dimension.

In a separate GLM we tested alternatively for the signed values V_{red}, V_{right} (instead of the within-dimensional certainties) and W_{color} (instead of the across-dimensional certainty). We did not find any activity at a corrected significance level in these contrasts. Note however that these contrasts are on a scale from red to green and left to right. As there is no common reference point for encoding the dimensions it is unlikely that the subject group encoded the features on the same scale (i.e., red to green or green to red).

We entered all regressors and modulators described above under 1-4 independently (without serial orthogonalization) into the design matrix. Thereby only the additional variance that cannot be explained by any other regressor is assigned to the effect, preventing spurious confounds between regressors (Andrade et al. 1999; Draper and Smith 1998).

The regressors were convolved with the canonical HRF, and low frequency drifts were excluded with a high-pass filter (128-s cutoff). Short term temporal autocorrelations were modeled using an AR(1) process. Motion correction regressors estimated from the realignment procedure were

entered as covariates of no interest. Statistical significance was assessed using linear compounds of the regressors in the GLM, generating statistical parametric maps (SPM) of t values across the brain for each subject and contrast of interest.

The within-dimension certainty contrast shown in Figure 3A and 3B is an equally weighted linear combination of within-dimension certainty regressors (2) and (3). We also looked at correlations with (2) and (3) separately and found that areas activated by (2) and (3) overlap and are both exclusively located at the same region as the combined contrast shown in Figure 3.

These contrast images were then entered into a second level random effects analysis using a one sample t-test against zero. The structural T1 images were coregistered to the mean functional EPI images for each subject and normalized using the parameters derived from the EPI images. Anatomical localization was carried out by overlaying the t-maps on a normalized structural image averaged across subjects, and with reference to an anatomical atlas (Duvernoy 1999).

We used a Bayesian model comparison (Stephan et al. 2009) to determine which model (GLM_attention-gating or GLM_integration) better explained the neural activity in vmPFC. First, we computed log-likelihoods for the two models, averaged within a 12mm sphere (1.5x smoothing kernel size) in vmPFC. Since the attention-gating model activated cluster was completely contained by the integration model activated area, we centered the sphere on the group peak of the weaker attention-gating model (any selection bias towards the stronger correlating model would then be working against us). Next, we calculated posterior model probabilities in this region for every subject and the group of subjects. In brief, the procedure by Stephan et al. rests on treating the model as a random variable and estimating the parameters of a Dirichlet distribution, which describes the probabilities for all models considered. These probabilities then define a multinominal distribution over model space, allowing one to compute how likely it is that a model generated the subjects' data. To decide which model is more likely, we use the conditional model probabilities to quantify an exceedance probability, i.e. a belief that a particular model is more likely than the other model, given the group data. For a detailed description of the technique also see the methods paper by Stephan, Friston and colleagues (Stephan et al. 2009).

All reported activations survive p<0.05 FWE corrections for multiple comparisons across the whole brain at the cluster level (p<0.001 height threshold for each voxel). Cluster level correction is more sensitive for activations with a larger spatial extent than for activations with very small clusters.. As our functional neuroanatomical hypotheses are all about cortical areas as opposed to sub-cortical areas (in which activations with smaller spatial extent are more typically detected), the use of cluster correction is appropriate in this case.

Results

Behavioral analysis and model comparison

On average across subjects, participants performed 294 trials (sd=62), chose the correct stimulus 60% of the time (sd=4) and got rewarded in 56% of the trials (sd=3). We tested if subjects' performance improved over time by splitting the trials in half and comparing the fraction of correct choices in the first half with the correct choices in the second half. There was no difference in the number of correct choices ($T_{df=15}$=0.26, p=0.80). To test if subjects became better at anticipating intra-and extra dimensional shifts during the task we further compared the number of trials required for the first half of switches (n_1 = 146, sd = 31) with the second half of switches (n_2 = 147, sd = 36) and did not observe any difference (t-test p>0.93).

Equally, response times between correct (rt=1.00 sec, sd=0.16) and incorrect (rt=1.02 sec, sd=0.17) trials were not different (paired t-test p>0.15). Subjects responded slightly faster when color was the relevant feature than when it was motion: color (rt=.97 sec, motion rt = 1.04 sec; paired t-test p<0.001).

While all subjects experienced the same number of switches one would nevertheless predict that in relative terms (compared to the total number of trials), fast learners experience a higher proportion of switching trials. To investigate this hypothesis we correlated the learning rate against the proportion of switches / number of total trials and found a significant relation (r = 0.83, p = 0.0002).

 We found behavioral evidence that when solving our task, subjects tend to consider evidences from all dimensions concurrently and rely more on the integration strategy rather than on a

hierarchical strategy that only occasionally shifts attention between dimensions (Table S1). To statistically evaluate the fit of the selection and the integration model we next calculated Bayes Factors for the comparison of the integration with the selection strategy. Bayes Factors indicate the evidence that one model is more likely than the other, with a log Bayes Factor of 3 or greater corresponding to an exceedance probability of 95% (strong evidence) (MacKay 2003). Our results indicate that the integration model was more likely to explain subjects' behavior than the attention-gating model in all subjects, and 13 out of 16 subjects showed strong evidence for the integration model (**Fig. 1D; Table S2**). Alternatively, as the models are nested, we performed a statistical likelihood ratio test which provides support for the integrating model based on all subject choices ($X^2_{df=1}$=500.5, p<<0.00001; or on the individual subject level for every subject p<0.01). Strong support in favor of the integration model is also provided if we compare both models using a Bayesian Model Comparison (Stephan et al. 2009) (alpha = 17.0 for integration versus 1 for attention-gating; posterior probability = 0.94 versus 0.06, exceedance probability $\approx$ 1.0 versus 0). Together, this suggests that subjects utilized information from the characteristics of both dimensions, rather than the alternative of concentrating only on the dimension that is believed more likely correct. **Figure 2** shows profiles of hidden targets, subject's choices and rewards, and the integration model predicted values of a representative subject.

In addition to the attention-gating and integration models we also tested other model variants such as a fully Bayesian model or simpler none hierarchical variants of the RL model but found that none of these could explain subject behavior as well as our integrating RL model, even after correcting for the different levels of complexity of the models (Table 1). The attention-gating model can be seen as a special case of the integration model (i.e. a nested model). However, this only happens when the transformation from value to weighting *w* approximates a Heaviside function (when κ->∞), as implemented in the attention-gating model. The fitted parameter κ is shown in Figure S3 for all subjects. Notice that the function is far from being a Heaviside for most subjects. Although the extra parameter gives more flexibility in fitting for the integration model, the BIC correction takes this into account.

Adhering to the model predicted choices is directly task relevant. Those subjects, whose choices could be relatively well explained by the integration strategy, also earned more money, as indicated by a significant correlation between the model likelihood of the integration model with

the average earned money / trial (r=0.50, p<0.05). Furthermore, the experienced ratio of switches in the task correlates both with the log likelihood of the integration model (r=0.86, p<0.001), and the average across-dimensional certainty (r=0.55, p=0.03). Subjects adhering to the model (good model fit) experience relatively high switch ratios as the task progresses fast. A high average certainty suggests that subjects have a better grasp of what dimension is correct, again this is related to how fast the task progresses. To analyze whether the success of one strategy over another depends on the frequency of the switches we correlated across subjects the fraction of experienced switches / total number of trials against the subjects' Bayes Factors from the model comparison (indicating the relative model fit). This relationship is not significant (r = 0.35, p > 0.17).

Neural correlates of decision variables

In order to identify neural correlates of the valuation process we separately regressed neural activity onto trial by trial value signals of the selection and integration models. In particular, we were interested in the maximum stimulus value signal on each trial as this is the key output variable of the decision making process. In the case of the attention-gating model, this value corresponds to the RL value of selecting the better stimulus within the dimension that the model predicts is currently relevant. In the integration model, this value is a linear combination of RL values for the color and motion features, weighted by the model predicted across dimension likelihood.

Based on previous findings (Hare et al. 2008; Padoa-Schioppa and Assad 2006; Plassmann et al. 2007; Wunderlich et al. 2010) we specifically predicted to find stimulus value signals in ventromedial prefrontal cortex (vmPFC). We found that the stimulus value signal correlated most strongly with activity in vmPFC, extending dorsally and medially along PFC (**Fig. 3A**). We also found chosen value signals in vmPFC (**Fig. S1**), though the effect size of the stimulus value was higher than that of the chosen value signal at the peak of the activation.

We found analogous activation patterns correlating with the attention-gating model's decision variables, albeit with a much weaker effect size and smaller extent (**Fig. S2**). In particular, neural activity in medial prefrontal cortex correlated both with the stimulus value of the gated and

integration model. The area correlating with the attention-gating model predicted variables was thereby entirely contained within the larger area activated by the integration model (**Fig. 3B**).

Neurometric model comparison between attention-gating and integration models

To discriminate between our two hypotheses about potential strategies, we tested if one model's prediction correlated better with BOLD activation than the other model's prediction. In order to perform a quantitative comparison of the value signals from both models within vmPFC, we estimated the two models in separate GLMs and determined their relative probability to explain the measured neural signal in every subject by means of a Bayesian model comparison approach (Stephan et al. 2009). The advantage of this method is that it circumvents any collinearity issues occurring from correlated regressors in a single GLM. Consistent with our behavioral results, we found that across subjects the integration model was more likely to be the underlying cause of the neural variability in vmPFC (Dirichlet alpha=13.34 for integration vs. alpha=4.66 for attention-gating posterior probability = 0.74, exceedance probability = 0.98; **Fig. 3C**). The decisive measure in this comparison is the exceedance probability, which in our case indicates that the integration model explains neural data better than the attention-gating model with 98% confidence.

Certainty and uncertainty related signals

To further investigate the neural substrates of this hierarchical decision task, we next looked more closely at decision related variables of the integration model. Specifically, we tested for correlations between certainty signals and the fMRI data. The across-dimension certainty reflects how likely it is that one of the complementary dimensions is the correct one. The within-dimensions certainty pertains to subjects getting the features for the two dimensions right, which in our model reflects a metric of subjects' choices being reinforced. We also hypothesized that we would see increased activity relating to within-dimension uncertainty in areas related to the allocation of cognitive resources and control on the basis of task difficulty. Areas most prominently implicated in this function on the basis of past literature are the anterior cingulate

cortex and dorsolateral prefrontal cortex (dlPFC) (Kerns et al. 2004; MacDonald et al. 2000; Matsumoto and Tanaka 2004)

We found that within-dimension certainty correlated with activity located in ventral medial PFC (**Fig. 4A**). The peak of this certainty related activity was located slightly anterior to the peak of the value related activity described above (Table 2).

We found within-dimension uncertainty correlating with activity in dorsomedial PFC adjunct to the ACC, bilaterally in dorsolateral PFC, the frontal poles, and parietal cortex (**Fig. 4B**).

No significant correlation was found between the BOLD signal and either our measure of certainty or uncertainty across dimensions even at $p < 0.001$ uncorrected. A list of all activated regions is shown in Table 2.

Discussion

We investigated how humans deal with complex decision tasks with multiple dimensions, and evaluated the computations subserved by the prefrontal cortex in this function. We specifically compared two strategies: do humans first determine the most likely dimension that explains how rewards are tied to stimuli and then choose on the basis of that theory (attention gated), or choose by integrating over multiple evidences and weighting each dimension by the likelihood that it is correct (integration strategy). The integration model had a significantly higher likelihood for explaining behavior than the attention-gating model, and neural activity in vmPFC correlated more strongly with trial by trial valuations according to the integration model than the attention-gating model.

Although there is much prior evidence indicating that prefrontal cortex plays a critical role in hierarchical decision making (Dias et al. 1996; Downes et al. 1989; Drewe 1974; Lawrence et al. 1996; Milner 1963; Owen et al. 1993; Owen et al. 1991; Robinson et al. 1980), up to now little was known about the specific computational processes underlying such a contribution. Our results begin to shed light on this question, by indicating that at the level of inference – when working out what choice to take next – the prefrontal cortex uses probability information in an integrated fashion. In this approach, even the least likely explanation of observed phenomena is

taken into account in order to generate a decision and its weight is commensurate with its likelihood. Such integration by likelihood has been formalized in Bayesian analysis. Bayesian like processes have been suggested to occur during cross-modal integration in terms of how information from perceptual cues in different sensory modalities is integrated together during perceptual decision making (Battaglia et al. 2003; Beers et al. 1999; Ernst and Banks 2002; Knill and Pouget 2004; Rao et al. 2002), to account for human cognition in perception (Jacobs 1999; Körding et al. 2007), action selection (Kording and Wolpert 2004; Trommershauser et al. 2003), and human causal inference (Gopnik et al. 2004; Griffiths and Tenenbaum 2005; Körding et al. 2007; Tenenbaum 1999). Note that optimality in our task alludes to the integration of possible evidences rather than optimal updating of evidences. Indeed, our decision models do not differ at the learning level: both use the same reinforcement learning principles to update values, ensuring that gated and integration approaches to decision making were compared in a clean fashion. When we fit a fully Bayesian model (with Bayesian learning of likelihoods, see supplementary methods) to behavior, results were slightly inferior, although the fully Bayesian model still performed better than the attention-gating model. We further examined posthoc whether the integration weights would affect learning by fitting a model that allows for a variable learning rate based on the weights of a dimension. Such a model gave very similar fits and produced very similar regressors to those produced by the integration model used in the fMRI analysis, indicating that we have no evidence on the basis of the present results that the integration weights are being incorporated into the learning rate. Instead these weights appear only to exert influence at the time of choice. Taken together, these findings suggest that while we have evidence that subjects integrated evidence at the point of choice, we cannot rule out the possibility that during learning, the updating of values may proceed through a variant of reinforcement learning as opposed to full Bayesian updating.

Our findings add significantly to current understanding about the role of the ventromedial prefrontal cortex in computing value during choice. The present results show that at the point of choice, the computations appear to involve dynamic integration of relevant information from the environment in order to compute value, as opposed to selectively representing the value of the most relevant features. These findings build on previous evidence that value signals in vmPFC dynamically reflect knowledge of the underlying structure in a task as opposed to solely reflecting the predictions of a reinforcement learning model (Hampton et al. 2006), as well as

findings that value signals in this area reflecting strategic computations during social decision making (Hampton et al. 2008). It is important to note that the present findings do not rule out a role for attention in the computation of value (Krajbich et al. 2010). Rather, our findings can be interpreted as suggesting that during the choice process attention might get distributed between the features of relevant stimuli during value computations.

The most prominent performance deficit reported on such hierarchical tasks following lesions of prefrontal cortex is perseveration, a tendency to persist in responding on the previously reinforced category following a change in stimulus contingencies (Milner 1963; Nelson 1976). These deficits have led to the proposal that prefrontal cortex, and particularly its lateral aspect, contributes to facilitating the switching of attentional focus between relevant cues (Dias et al. 1996; Rogers et al. 2000), which is predicated on the notion that attention is allocated selectively to only one dimension at a time. A failure to switch behavior following a change in contingencies would therefore, according to this suggestion, occur due to an inability to disengage attention from the previously reinforced category and relocate it toward the new category. Our results suggest that rather than facilitating the allocation of attention to only one dimension selectively at a time, prefrontal cortex instead may contribute toward the distribution of attention across multiple probable causes. Thus, prefrontal cortex keeps track of multiple contingencies, even if some of them are temporarily unlikely to be correct. Therefore, the pattern of impairment following lesions in prefrontal cortex could result from an inability to integrate across contingencies, and hence, an inability to distribute attention (or indeed readjust the distribution of attention) optimally across multiple causes, rather than exclusively controlling the attentional focus toward the most likely cause. In the evidence integration approach, trust in one's decision is best measured by our within-dimension certainty measure, which combines the likelihood that one has identified the right feature for both color and motion dimensions. Accordingly, we discovered positive correlations between the within-dimension certainty measure and activation in anterior medial PFC.

We found a correlation with within-dimensional uncertainty in dorsomedial prefrontal cortex, dorsolateral prefrontal cortex, the frontal poles, and parietal cortex. These latter areas are more active in those trials in which subjects are unsure about the correct choice. Our finding of

uncertainty related signals in dorsomedial PFC resonates with a recent implication of this region in tracking the volatility in the environment to adjust the rate at which learning occurs (Behrens et al. 2007). Furthermore, ACC and dlPFC have long been suggested to interact together for the purposes of cognitive control (Kerns et al. 2004; Koechlin et al. 2003; MacDonald et al. 2000; Matsumoto and Tanaka 2004); while frontopolar cortex has previously been associated with uncertainty (Yoshida and Ishii 2006) and exploring multiple behavioral alternatives in order to maximize reward (Koechlin et al. 1999; Koechlin and Hyafil 2007). It is not efficient to maintain a high level of control all the time because cognitive control requires effort and the brain therefore regulates control according to demand. The best strategy for this regulation is, in our task, a variation according to the uncertainty about the correct choice, which is captured by our within-dimension uncertainty signal. Note that in a two-step attention gated approach, for the effective allocation of attention between dimensions, cognitive control would have to be recruited in relation to uncertainty about what dimension is relevant. High uncertainty across dimensions would then render the task more difficult for subjects because they first have to settle on one dimension before they decide on the feature. In contrast, the continuously integrating model utilizes a one-step procedure and thus task difficulty is only reflected in the uncertainty of the final output. Indeed our findings indicate that areas involved in recruiting cognitive control correlated only with the general uncertainty about the correct features (within-dimension uncertainty) but not the across dimension uncertainty, suggesting that in our task cognitive control is recruited as a result of general uncertainty about which stimulus is correct rather than which dimension.

A number of previous studies have attempted to differentiate the functional contributions of different regions of prefrontal cortex in hierarchical decision making (Badre et al. 2010; Bechara et al. 2000; Goldman-Rakic 1996; Hampton et al. 2006; Koechlin and Summerfield 2007; Owen 1997; Wallis et al. 2001). Using a selective lesion approach it was recently shown that damage to a number of different areas of prefrontal cortex can cause impairment on hierarchical decision making in monkeys, including the orbitofrontal cortex, anterior cingulate and ventral and dorsal prefrontal cortices, and moreover that these regions may make distinct contributions underlying performance (Buckley et al. 2009). Our finding of distinct task relevant computational signals within prefrontal cortex is broadly consistent with these findings. While ventromedial prefrontal cortex is involved in computing the value of particular choices on the basis of prior

reinforcement history in order to guide choice, lateral prefrontal cortex, frontopolar cortex and dorsomedial prefrontal cortex may contribute to the allocation of attention across the stimuli according to the degree of uncertainty (or indeed conflict) about which feature is currently correct.

It is also important to consider the relationship between the present results and previous trial based analyses (Cools et al. 2002; Ghahremani et al. 2010; Hampton et al. 2006; O'Doherty et al. 2001) and neurochemical investigations (e.g. (Clarke et al. 2005; Robbins and Roberts 2007; Roberts et al. 1994)) of reversal learning. Most importantly, these prior studies have reported activity in lateral orbital and adjacent prefrontal cortex following a change in reinforcement contingencies leading to a switch in behavior (Cools et al. 2002). Those previous findings have implicated these areas in either the evaluation of negative reinforcement per se, or in the facilitation of behavioral switches through response inhibition and/or response selection, or the detection of contingency changes. The present results are not incompatible with those previous findings on reversal, but rather address a separate issue. Here we address the issue about how value computations are generated at the point of choice when information about value is available through multiple dimensions, while those previous studies pertain to the mechanism by which choice is alternated between different stimulus/action pairs. In order to assess the extent to which these previous results could be replicated in the present data, we ran an analysis in which we looked for activity immediately preceding a switch in stimulus choice and replicated at least some of the findings from this previous literature (see Figure S4). However, it should be noted that because in the present study we used a probabilistic structure at both levels of the hierarchy, the present task design is not optimized for analyzing the data in such a trial based manner. This is because we cannot unambiguously ascertain when subjects subjectively understood a switch in contingencies had occurred for a particular stimulus dimension. Here, each stimulus contained both dimensions and subjects' choices were therefore not unique in the decomposition of whether their choice was based on color or motion, or in case of the integration model a weighted combination of both. Instead, the design of the task was optimized for a computational model based analysis as opposed to the trial based "switch" analysis used in those previous studies. Specifically, switches in our task were more rapid than in previous studies, increasing sensitivity for differentiating between integrating and selection: in a stable environment, the weight for the irrelevant dimension in the integration model would become very small over time

and the integration model would eventually predict identical outcomes as the attention-gating model.

One possible concern with the present approach is that it is difficult to determine whether attention is distributed at the same time to both dimensions or switched stochastically from one to the other on a trial by trial basis. In this scenario, gated-attention is noisy and the dimensions are considered one at a time, but stochastically. However, this can produce similar fits to the integration model only if alternative dimensions are not considered purely randomly, but with a frequency that mimics integrating behavior. Importantly, such a strategy still relies on quantifying the relevance of the two dimensions, which is captured by the weights in our model. Furthermore, stochastic switching would still require a two-stage decision with a choice about dimension in every trial. In order to perform such a behavior the subject would require knowledge about the certainty of the most relevant dimension (across dimensional certainty), a signal that is conspicuously absent from the imaging data, even though other signals such as the within-dimensional certainty were present at robust statistical significance levels. Although we cannot rule out the presence of this signal on the basis of a null result, the most parsimonious explanation for our imaging data seems to be an integrative account rather than a stochastic switching account. In general it remains also possible, that if humans are faced with a hierarchical problem of sufficient complexity (for example requiring integration over many more than two dimensions), keeping track of all theories simultaneously becomes both cognitively too challenging and normatively ill-advised (Bröder and Schiffer 2003; Diaconis and Freedman 1986). In those instances, subjects might switch to a simpler strategy like attention gating, which employs fewer resources.

Our results indicate that the human brain and the prefrontal cortex in particular, is capable of integrating information in an optimally weighted manner. These findings could be used as a basis to promote better decision making strategies in everyday life and in policy making, whereby decision makers are encouraged to take into account probabilistic information about multiple causes simultaneously rather than being trained to generate decisions based on sequential assumptions that may ultimately prove false.

Acknowledgements

References

Andrade A, Paradis AL, Rouquette S, and Poline JB. Ambiguous results in functional neuroimaging data analysis due to covariate correlation. *Neuroimage* 10: 483-486, 1999.

Badre D, Kayser AS, and D'Esposito M. Frontal cortex and the discovery of abstract action rules. *Neuron* 66: 315-326, 2010.

Battaglia PW, Jacobs RA, and Aslin RN. Bayesian integration of visual and auditory signals for spatial localization. *J Opt Soc Am A Opt Image Sci Vis* 20: 1391-1397, 2003.

Bechara A, Tranel D, and Damasio H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 123 (Pt 11): 2189-2202, 2000.

Beers RJv, Sittig AC, and Gon JJ. Integration of proprioceptive and visual position-information: An experimentally supported model. *J Neurophysiol* 81: 1355-1364, 1999.

Behrens TE, Woolrich MW, Walton ME, and Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci* 10: 1214-1221, 2007.

Berger J. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer, 1980.

Boorman ED, Behrens TE, Woolrich MW, and Rushworth MF. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62: 733-743, 2009.

Bröder A, and Schiffer S. Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making* 213: 193-213, 2003.

Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PG, Kwok SC, Phillips A, and Tanaka K. Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science* 325: 52-58, 2009.

Burnham KP, and Anderson D. *Model Selection and Multi-Model Inference*. Springer, 2002.

Clarke HF, Walker SC, Crofts HS, Dalley JW, Robbins TW, and Roberts AC. Prefrontal Serotonin Depletion Affects Reversal Learning But Not Attentional Set Shifting. *Journal of Neuroscience* 25: 532-538, 2005.

Cools R, Clark L, Owen AM, and Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 22: 4563-4567, 2002.

Daw ND. Trial-by-trial data analysis using computational models. In: *Affect, Learning and Decision Making, Attention and Performance*, edited by E.A. P, T.W. R, and M D. Oxford: Oxford University Press, 2011.

Daw ND, O'Doherty JP, Dayan P, Seymour B, and Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature* 441: 876-879, 2006.

Diaconis P, and Freedman D. On the Consistency of Bayes Estimates. *The Annals of Statistics* 14: 1-26, 1986.

Dias R, Robbins TW, and Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380: 69-72, 1996.

Downes JJ, Roberts AC, Sahakian BJ, L. EJ, Morris RG, and Robbins TW. Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: Evidence for a specific attentional dysfunction. *Neuropsychologia* 27: 1329-1343, 1989.

Draper NR, and Smith H. Extra Sums of Squares and Tests for Several Parameters being Zero. In: *Applied Regression Analysis* Wiley, 1998.

Drewe EA. The effect of type and area of brain lesion on Wisconsin card sorting test performance. *Cortex* 10: 159-170, 1974.

Duvernoy HM. *The Human Brain. Surface, Blood Supply and Three-Dimensional Section Anatomy*. NewYork: Springer, 1999.

Ernst MO, and Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415: 429-433, 2002.

FitzGerald TH, Seymour B, and Dolan RJ. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci* 29: 8388-8395, 2009.

Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, and Poldrack RA. Neural Components Underlying Behavioral Flexibility in Human Reversal Learning. *Cerebral Cortex* 20: 1843-1852, 2010.

Gigerenzer G, and Goldstein DG. Reasoning the fast and frugal way: models of bounded rationality. *Psychol Rev* 103: 650-669, 1996.

Goldman-Rakic PS. The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philos Trans R Soc Lond B Biol Sci* 351: 1445-1453, 1996.

Gopnik A, Glymour C, Sobel D, Schulz L, Kushnir T, and Danks D. A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review* 111: 1-31, 2004.

Grant DA, and Berg EA. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *J Exp Psychol* 38: 404-411, 1948.

Griffiths TL, and Tenenbaum JB. Structure and strength in causal induction. *Cognit Psychol* 51: 334-384, 2005.

Hampton AN, Bossaerts P, and O'Doherty JP. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc Natl Acad Sci U S A* 105: 6741-6746, 2008.

Hampton AN, Bossaerts P, and O'Doherty JP. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 26: 8360-8367, 2006.

Hare TA, O'Doherty J, Camerer CF, Schultz W, and Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28: 5623-5630, 2008.

Jacobs RA. Optimal integration of texture and motion cues to depth. *Vision Res* 39: 3621-3629, 1999.

Kass RE, and Raftery AE. Bayes Factors. *Journal of the American Statistical Association* 90: 773-795, 1995.

Kerns JG, Cohen JD, MacDonald AW, III, Cho RY, Stenger VA, and Carter CS. Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science* 303: 1023-1026, 2004.

Knill DC, and Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci* 27: 712-719, 2004.

Knutson B, Taylor J, Kaufman M, Peterson R, and Glover G. Distributed Neural Representation of Expected Value. *J Neurosci* 25: 4806-4812, 2005.

Koechlin E, Basso G, Pietrini P, Panzer S, and Grafman J. The role of the anterior prefrontal cortex in human cognition. *Nature* 399: 148-151, 1999.

Koechlin E, and Hyafil A. Anterior prefrontal function and the limits of human decision-making. *Science* 318: 594-598, 2007.

Koechlin E, Ody C, and Kouneiher F. The architecture of cognitive control in the human prefrontal cortex. *Science* 302: 1181-1185, 2003.

Koechlin E, and Summerfield C. An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences* 11: 229-235, 2007.

Körding KP, Beierholm U, Ma WJ, Quartz SR, Tenenbaum JB, and Shams L. Causal Inference in Multisensory Perception. *PLoS ONE* 2: e943, 2007.

Kording KP, and Wolpert DM. Bayesian integration in sensorimotor learning. *Nature* 427: 244-247, 2004.

Krajbich I, Armel C, and Rangel A. Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci* 13: 1292-1298, 2010.

Lawrence BJ, Sahakian BJ, Hodges AE, Rosser KW, Lange TW, and Robbins TW. Executive and mnemonic functions in early Huntington's disease. *Brain* 119: 1633-1645, 1996.

MacDonald AW, Cohen JD, Stenger VA, and Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288: 1835-1838, 2000.

MacKay DJC. Comparison of Approximate Methods for Handling Hyperparameters. *Neural Computation* 11: 1035-1068, 1999.

MacKay DJC. *Information theory, inference and learning algorithms*. Cambridge: Cambridge University Press, 2003.

Matsumoto K, and Tanaka K. Neuroscience. Conflict and cognitive control. *Science* 303: 969-970, 2004.

Milner B. Effects of different brain lesions on card sorting. *Archives of Neurology* 9: 100-110, 1963.

Nelson HE. A modified card sorting test sensitive to frontal lobe defects. *Cortex* 12: 313-324, 1976.

O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, and Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4: 95-102, 2001.

Owen AM. Cognitive planning in humans: neuropsychological, neuroanatomical and neuropharmacological perspectives. *Prog Neurobiol* 53: 431-450, 1997.

Owen AM, Roberts AC, Hodges JR, Summers BA, Polkey CE, and Robbins TW. Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain* 116 (Pt 5): 1159-1175, 1993.

Owen AM, Roberts AC, Polkey CE, Sahakian BJ, and Robbins TW. Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia* 29: 993-1006, 1991.

Padoa-Schioppa C, and Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441: 223-226, 2006.

Plassmann H, O'Doherty J, and Rangel A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* 27: 9984-9988, 2007.

Rao R, Olshausen B, and Lewicki M. *Probabilistic Models of the Brain*. Cambridge, Massachusetts: MIT Press, 2002.

Rescorla RA, and Wagner AR. *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, Classical Conditioning II*. New York: Appleton-Century-Crofts, 1972.

Robbins TW, and Roberts AC. Differential Regulation of Fronto-Executive Function by the Monoamines and Acetylcholine. *Cerebral Cortex* 17: i151-i160, 2007.

Roberts AC, De Salvia MA, Wilkinson LS, Collins P, Muir JL, Everitt BJ, and Robbins TW. 6-Hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin Card Sort Test: possible interactions with subcortical dopamine. *Journal of Neuroscience* 14: 2531-2544, 1994.

Robinson AL, Heaton RK, Lehman RA, and Stilson DW. The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *J Consult Clin Psychol* 48: 605-614, 1980.

Rogers RD, Andrews TC, Grasby PM, Brooks DJ, and Robbins TW. Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *J Cogn Neurosci* 12: 142-162, 2000.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, and Friston KJ. Bayesian model selection for group studies. *Neuroimage* 46: 1004-1017, 2009.

Sutton RS, and Barto AG. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

Tenenbaum JB. Bayesian modeling of human concept learning. In: *Advances in Neural information Processing Systems 11*, edited by Kearns MS, Solla SA, and Cohn DA. Cambridge, MA: MIT Press, 1999, p. 59-65.

Trommershauser J, Maloney LT, and Landy MS. Statistical decision theory and the selection of rapid, goal-directed movements. *J Opt Soc Am A Opt Image Sci Vis* 20: 1419-1433, 2003.

Wallis JD, Anderson KC, and Miller EK. Single neurons in prefrontal cortex encode abstract rules. *Nature* 411: 953-956, 2001.

Wunderlich K, Rangel A, and O'Doherty JP. Economic choices can be made using only stimulus values. *Proc Natl Acad Sci U S A* 107: 15005-15010, 2010.

Wunderlich K, Rangel A, and O'Doherty JP. Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences* 106: 17199-17204, 2009.

Yang T, and Shadlen MN. Probabilistic reasoning by neurons. *Nature* 447: 1075-1080, 2007.

Yoshida W, and Ishii S. Resolution of uncertainty in prefrontal cortex. *Neuron* 50: 781-789, 2006.

Figure Captions

Figure 1: **Task and Models**

(A)　Subjects chose one of two items, of which each had a color (red or green) and a motion (left or rightwards moving dots) attribute. The features were randomly assigned to both items. Once the subject selected an item, a box was placed around the target and remained on the screen until 2s after stimulus onset. After a 3s delay they either received a 25 cent reward or a subtraction of 25 cents from their payout. One feature was designated the correct feature, and the choice of the item carrying that feature led to a reward on 80% of the occasions and a loss 20% of the time. Consequently, by choosing this correct item subjects accumulated monetary gain. The other item was incorrect, and choosing it led to a reward 20% of the time and a loss 80% of the time, leading to a cumulative monetary loss. After subjects chose the correct item on three consecutive occasions, the contingencies reversed with a probability of 50% in every consecutive trial. After two to four of such within-dimension reversals the relevant dimension changed (extra-dimensional switch). The inter-trial-interval was variable.

(B)　Hierarchical decision model based on attentional-shifts. Stimulus-outcome associations are learned for color (C) and motion direction (M). According to this model, subjects choose in a two-step process: they first form a hypothesis about which dimension is relevant (either C or M) and then base their reward expectation and choice exclusively on the information learnt about that dimension.

(C)　In the integration model, the available information from both dimensions (C and M) is integrated as a weighted sum and the decision is based on a linear combination of evidence from both dimensions. Subjects form a hypothesis about the likelihoods that each dimension is relevant, corresponding to weights in the linear combination. Weights are updated on every trial.

(D)　Bayes Factors of the comparison of the integration versus attention-gating model. The integration model fits better to subjects' behavior in every single subject and for 13 out of 16 subjects the log Bayes Factor indicated strong evidence in favor of the integration model.

Figure 2: Integration model predictions and behavior

(A)   Top: model predicted value for tracking the color feature. Bottom: model predicted value for tracking the motion feature. Middle: Learned weight for the dimension. X-axis is time, unit is trial. The relevant dimension is indicated by a gray background in either the color or motion plot. Feature reversals within this block are shown as black lines and the icon directly above this line denotes the new correct feature. The green/red boxes below the timecourses show whether subject's choice on the trial was correct/incorrect and a blue box indicates that the trial was rewarded. Data is shown for a representative subject over the time of the entire experiment.

(B)   Model certainty for the two features (red = color; blue = motion) and the relevant dimension.

(C)   Model predicted choice versus actual choice for the subject shown in A-B. The proportion of choices of the upper stimulus (dark) increases with higher model predicted value for the upper stimulus. The proportion of bottom stimulus choices (light) follow the opposite course. The model predicted value for the bottom stimulus is 1-(top stimulus value).

Figure 3: Neural correlates of decision value and model comparison

(A)   BOLD responses in medial PFC correlate significantly with the stimulus value signal from the integration model.

(B)   Neural activity in vmPFC correlates with the trial-by-trial stimulus value signal. Shown is the area that is commonly activated by both the integration and the attention-gating model at $p < 0.001$ corrected. The comparison between the two models in (c) is based on this commonly activated region.

(C)   We used a Bayesian Model comparison to identify the model that can better explain neural activity in this area. Overall, the integration model (I) explains activity in this vmPFC area better than the attention-gating model (A). The exceedance probability for the integration model is 0.98. The exceedance probability is the probability that one model is more likely than the other one, i.e that the posterior probability for the integration model is larger than 0.5.

Figure 4: Certainty and uncertainty of the integration model

(A)   BOLD responses in a sub-cluster of medial PFC correlate significantly with the within-dimension certainty (averaged across color and motion) from the integration model.

 (B)   Dorsomedial & dorsolateral PFC and the frontal poles show negative correlations with the within-dimension certainty (thus indicating a positive correlation with uncertainty).

The color-coding is identical to Figure 3.

Table Captions

Table 1:

Pairwise model comparison using Bayesian Model Comparison. Numbers indicate exceedance probabilities of the column model versus the row model.

Table 2:

Locations of significant correlation with value signals of the integration model (threshold $p<0.001$ uncorrected and 20 voxels extent; * $p<0.05$ FWE corrected for multiple comparisons at the cluster level). MNI coordinates denote the group peak voxel of each cluster.