Behavioral/Systems/Cognitive

Prior Expectation Modulates the Interaction between Sensory and Prefrontal Regions in the Human Brain

Dobromir Rahnev,[1,2] **Hakwan Lau,**[1,2] **and Floris P. de Lange**[2]

[1]Columbia University, Department of Psychology, New York, New York 10027, and [2]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behavior, 6525 EN Nijmegen, The Netherlands

How do expectations about the identity of a forthcoming visual stimulus influence the neural mechanisms of perceptual decision making in the human brain? Previous investigations into this issue have mostly involved changing the subjects' attentional focus or the behavioral relevance of certain targets but rarely manipulated subjects' prior expectation about the likely identity of the stimulus. Also, because perceptual decisions were often paired with specific motor responses, it has been difficult to dissociate neural activity that reflects perceptual decisions from motor preparatory activity. Here we designed a task in which we induced prior expectations about the direction of a moving-dot pattern and withheld the stimulus–response mapping until the subjects were prompted to respond. In line with current models of perceptual decision making, we found that subjects' performance was influenced by their expectation about upcoming motion direction. The integration of such information into the decision process was reflected by heightened activity in the dorsolateral prefrontal cortex. Activity in this area reflected the degree to which subjects adjusted their decisions based on the prior expectation cue. Furthermore, there was increased effective connectivity between sensory regions (motion-sensitive medial temporal area MT+) and dorsolateral prefrontal cortex when subjects had a prior expectation about the upcoming motion direction. Dynamic causal modeling suggested that stimulus expectation modulated both the feedforward and feedback connectivity between MT+ and prefrontal cortex. These results provide a mechanism of how prior expectations may affect perceptual decision making, namely by changing neural activity in, and sensory drive to, prefrontal areas.

Introduction

Perceptual decision making has recently received great attention by researchers (Heekeren et al., 2008; Ratcliff and McKoon, 2008; Tosoni et al., 2008; Donner et al., 2009; Ferrera et al., 2009; Egner et al., 2010; Noppeney et al., 2010). Perceptual decisions are almost always informed and heavily biased by our prior expectations. For example, we identify objects much more rapidly in a familiar context that sets up priors for object recognition (Bar, 2004; Enns and Lleras, 2008).

Most studies of perceptual decision-making have looked at the influence of expectations on perceptual inference by selectively biasing participants to process one type of information over another (Summerfield et al., 2006; Summerfield and Koechlin, 2008; Esterman and Yantis, 2010; Preuschhof et al., 2010). Although this may selectively bias processing for one stimulus type, the prior probability of occurrence for a stimulus (i.e., its likeli-

hood) is not manipulated in these studies. Rather, they manipulated the amount of attention devoted to the stimulus (i.e., its relevance) (Summerfield and Egner, 2009). One recent human imaging study that manipulated prior information (rather than relevance) in perceptual decision making (Forstmann et al., 2010) observed that expectation for leftward or rightward motion selectively activated a contralateral corticostriatal circuit that was linked to the associated behavioral response. However, as in most primate neurophysiological studies (Gold and Shadlen, 2007), in the study by Forstmann et al. perceptual priors (i.e., "expecting leftward motion") covaried with motor priors (i.e., "expecting to make a leftward response"). Therefore, it is not clear whether subjects integrated the prior information into the perceptual decision-making process or whether the cue helped them to prepare the most likely response.

In the current study, we manipulated prior probability of sensory evidence during a perceptual decision-making task, while avoiding a fixed stimulus–response mapping. This allowed us to assess how perceptual prior expectations influence neural activity in sensory and prefrontal areas and uncover the mechanisms of decision making independently from motor preparation (Bennur and Gold, 2011). Behavioral results showed that stimulus expectation changed subjects' perceptual choice, rendering subjects more biased (shifted criterion) and slightly less sensitive (decreased d'). Neuroimaging results indicated that, when subjects had a prior expectation about the sensory evidence, this resulted in both increased activity in dorsolateral prefrontal cortex and increased recurrent con-
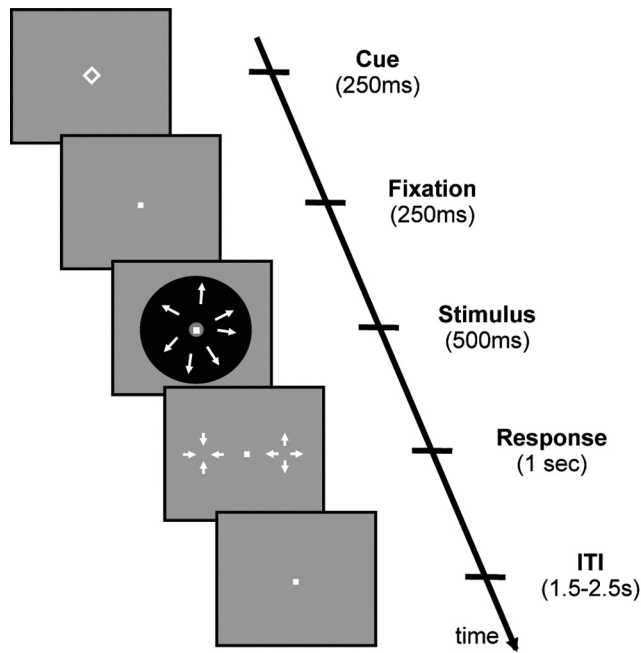
Figure 1. Task design. In each trial, subjects were asked to press keys to indicate whether a patch of dots was moving in an expanding or contracting direction. A cue in the form of a simple geometric shape indicated the likely direction of the motion. In half of the trials, the cue was predictive of the motion direction (75% valid, 25% invalid), whereas in the other half of the trials, a neutral cue was presented, which gave no information about the likely direction of the upcoming motion (i.e., a non-predictive cue). To perform optimally when the cue was predictive, subjects needed to combine the information from the cue and stimulus. The response mapping was only shown after the offset of the stimulus, so that subjects could not prepare for a motor response before the end of the motion presentation. Each trial lasted between 3.5 and 4.5 s (chosen from a uniform distribution), and subjects had up to 1 s to give an answer after the offset of the stimulus. ITI, Intertrial interval.

nectivity between this area and motion-sensitive medial temporal sensory area MT+.

Materials and Methods

Subjects. Twenty-three healthy right-handed subjects took part in the experiment. All participants had normal or corrected-to-normal vision. Two subjects performed at chance in discriminating the direction of motion during the experiment and were therefore excluded from additional analysis. All the analyses were done with the remaining 21 subjects (15 females; mean ± SD age, 21.6 ± 2.6 years). A written informed consent was obtained from all subjects. The research was approved by the local ethics committee in which the experiment was performed (CMO region Arnhem-Nijmegen, The Netherlands).

Stimuli and task. Subjects were required to indicate the overall direction of motion of white dots (density, 2.4 dots/degree²; speed, 6°/s) presented inside a black annulus (outer circle radius, 10°; inner circle radius, 1°). The motion direction was either contracting or expanding. A separate set of dots was chosen to carry the coherent motion each trial. Incoherent dots moved randomly with the same speed as coherent dots. Each dot had a lifetime between three and five frames (projector refresh rate, 60 Hz), after which it was replaced by another dot at a random location within the annulus. A small fixation square was presented for the duration of the trial, and subjects were required to maintain fixation on it. The stimuli were presented on gray background (Fig. 1) and were generated using Psychophysics Toolbox (Brainard, 1997) in MATLAB (MathWorks).

Each trial began with the presentation of a cue that was a simple geometric shape (Fig. 1). Four different shapes were used: square, diamond, triangle pointing up, and triangle pointing down. For half of the subjects, the first two cues were predictive of the forthcoming stimulus, whereas the last two were not. For the remaining half of the subjects, this

relationship was inverted. To further avoid any confounds from low-level physical characteristics of the cues, each shape from either pair of shapes was counterbalanced across subjects to indicate either a contracting or expanding stimulus (this was only relevant when the shape was predictive). The predictive cue indicated the forthcoming stimulus correctly on 75% of the trials. We refer to the trials in which the subsequent stimulus was correctly predicted by the cue as "valid." In the remaining 25% of the time, the predictive cues indicated the wrong direction of motion. We refer to these trials as "invalid." The non-predictive cues were followed by expanding and contracting motion equally often. Thus, the non-predictive cues are also referred to as "neutral" throughout. Subjects were fully informed about the above contingencies and were encouraged to take the cue into account when making their perceptual decisions.

Subjects gave their response with the index fingers of their left and right hands. On each trial, they were instructed which hand should be used for which response. Thus, the answer "expanding" might be mapped to the right-hand button press on one trial and to a left-hand button press on another trial. Importantly, the instruction only came after the offset of the stimulus, thus ensuring that any accumulation of evidence during stimulus presentation was not conflated with motor preparation.

The cue was presented for 250 ms, followed by a fixation dot for the same duration. Then, the motion stimulus appeared for 500 ms, and finally the trial-specific response mapping appeared. Subjects were not allowed to respond before the appearance of the response mapping and had 1 s to respond once it appeared. Reaction times (RTs) were measured from the onset of the response mapping. During the experiment, no feedback was provided. The total duration of a single trial was between 3.5 and 4.5 s, with the intervals sampled from a uniform distribution.

Before the fMRI experiment, each subject took part in a 1 h training session administered on a separate day, in which the subject practiced 882 trials. During the first half of the training session, subjects were given trial-by-trial feedback, whereas there was no feedback during the second part of the training to prepare subjects for the experiment in the scanner environment. Just before the fMRI experiment, subjects practiced for additional 5 min (64 trials) to remind them of the task requirements.

Based on the data from the behavioral session, three motion coherence levels were chosen for each subject. The intermediate coherence level was individually chosen to produce ~75% correct responses in the absence of a cue (mean ± SD coherence, 4.4 ± 0.7%). The low motion coherence was fixed at 50% of the intermediate coherence, whereas the high motion coherence was 150% of the intermediate coherence. The presentation of motion coherence was pseudorandomized for each subject such that each coherence level appeared equally often with each combination of cue identity and motion direction.

The experiment consisted of two sessions of 336 trials each. Trials were grouped in 42 blocks of eight trials. Each block consisted of trials with either predictive or non-predictive cues. The predictive and non-predictive blocks alternated, with the first block counterbalanced across subjects to be either predictive or non-predictive. Each block began with a 1 s presentation of the two geometric shapes that served as cues during the subsequent eight trials, followed by a 1 s interval of fixation. This was done to remind subjects of the type of cue during the following block. Blocks were separated by an additional 4–8 s (sampled from a uniform distribution), resulting in an average block length of 40 s.

We recorded a localizer for the MT+ region in a separate session after the main experiment. Forty-two blocks of moving dots (block duration of 16 s) were alternated with 14 blocks of stationary dots (block duration of 16 s), resulting in ~15 min scan duration.

Behavioral analyses. We computed the signal detection theoretic (SDT) measures d' and c (Macmillan and Creelman, 2005). The variable d' is a measure of a subject's stimulus discrimination sensitivity, whereas c is a measure of a subject's bias. These measures were calculated on the basis of hit and false-alarm rates. By only considering the trials in which subjects were cued that the stimulus was likely to be expanding, we computed d'_{expand} and c_{expand}. Similarly, by only considering the trials in which the cue indicating that the stimulus was likely to be contracting, we computed $d'_{contract}$ and $c_{contract}$. Because we were not interested in gen-

eral performance differences between contracting and expanding motion, we then averaged across these motion types to obtain $d'_{predictive}$, which reflects the discriminability of the motion stimuli in the context of a predictive cue. Conversely, c_{expand} and $c_{contract}$ are the biases for answering "expand" when an expanding or contracting cue was presented, respectively. Thus, we expected c_{expand} to be negative (corresponding to answering "expand" >50% of the time), and $c_{contract}$ to be positive (corresponding to answering "expand" <50% of the time). Criterion shift was defined as $c_{contract} - c_{expand}$, i.e., the difference of how much the cues were able to move subjects' criterion for picking one response option over the other. The bigger this shift was, the more the subject adjusted their behavior based on the predictive cue.

fMRI acquisition. Images were acquired on a 3 tesla Trio MRI system (Siemens). Functional images were acquired using a 32-channel coil, with a single-shot gradient echo-planar imaging sequence (repetition time, 1950 ms; echo time, 30 ms; 31 ascending slices; voxel size, 3 × 3 × 3 mm; flip angle, 80°; field of view, 192 mm). A high-resolution anatomical image was acquired using a T1-weighted MPRAGE sequence (repetition time, 2300 ms; echo time, 3.03 ms; voxel size, 1 × 1 × 1 mm).

fMRI data analysis. Analysis was performed using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK). The first six volumes of each scan were discarded to allow for scanner equilibration. Preprocessing consisted of realignment through rigid-body registration to correct for head motion, slice timing correction to the onset of the first slice, coregistration of the functional and anatomical images, segmentation of the anatomical image, normalization to Montreal Neurological Institute (MNI) space using the gray matter image obtained from the segmentation, interpolation of functional images to 2 × 2 × 2 mm, and smoothing with a Gaussian kernel with a full-width at half-maximum of 8 mm. A high-pass filter (cutoff, 128 s) was applied to remove low-frequency signals, such as scanner drift.

Regressors for the first-level analysis were obtained by convolving the unit impulse time series for each condition with the canonical hemodynamic response function. We modeled the three different cue–stimulus types (valid, invalid, neutral) separately for all trials in which subjects responded correctly. Trials with incorrect or no responses were modeled as a regressor of no interest. The onsets of the trials were defined at the onset of the cue presentation, and the duration was set to 1 s, the period from the onset of the cue to the offset of the motion stimulus. We included 12 nuisance regressors related to head motion: three regressors related to translation and three regressors related to rotation of the head, as well as their derivatives (Lund et al., 2005).

Unless stated otherwise, the reported activations are at $p < 0.05$ corrected at the cluster level for multiple comparisons using an auxiliary (uncorrected) voxel threshold of $p < 0.001$. This auxiliary threshold defines the extent of activated clusters that formed the basis of our statistical inference (Friston et al., 1996).

Psychophysiological interaction analysis. We performed psychophysiological interaction (PPI) analyses to test whether connectivity between MT+ and regions in the parietal and prefrontal cortex depended on the identity of the cues. PPI represents a measure of context-dependent connectivity, explaining regionally specific responses in one brain area in terms of the interaction between responses in another brain region and a cognitive or sensory process (Friston et al., 1997). We defined the left and right MT+ for each subject using the data from the localizer by applying the contrast "moving dots > stationary dots." We also defined the dorsolateral prefrontal cortex (DLPFC) and the intraparietal sulcus (IPS) based on the contrast "valid + invalid (i.e., predictive) > neutral (i.e., non-predictive)" in the main experiment. For each subject, BOLD response time series values were extracted from these regions. The analysis was constructed to test for differences in the regression slope of MT+ activity on IPS and DLPFC, depending on whether subjects were engaged in tasks with either predictive or non-predictive cues.

To construct the PPI regressors, we multiplied the time courses extracted from left and right MT+ ("physiological factor") with the time course of the experimental manipulation (predictive trials − non-predictive trials, "psychological factor") (Gitelman et al., 2003). This PPI regressor, alongside the task regressors and the time course of the seed region, were included in the statistical model for each subject. Separate

models were created for left and right MT+. To visualize the regression slopes separately for predictive and non-predictive trials, we multiplied the time course of neural activity in MT+ with task regressors relating to predictive and non-predictive trials, respectively. However, it should be noted that the difference in slope values between conditions, rather than their absolute values, are of interest. The reason is that the slope for predictive trials is implicitly biased by the non-predictive trials and vice versa.

Average parameter estimates for each region of interest were calculated for each subject separately. These estimates were then subjected to a paired-samples t test to test the prediction that connectivity between MT+ and left DLPFC and/or IPS is different for predictive compared with non-predictive blocks.

Dynamic causal modeling analysis. Dynamic causal modeling (DCM) aims to model directed (causal) influences between regions (Friston et al., 2003). In the PPI analysis, we observed stronger connectivity between MT+ and the left DLPFC. DCM allows us to adjudicate between stronger feedforward, feedback, or recurrent interactions between MT+ and DLPFC. This analysis was performed using SPM8 (Wellcome Department of Imaging Neuroscience). To perform the DCM analysis, we estimated a first-level analysis with the following regressors: stimulus (modeling the main effect of stimulus, i.e., stimulus > baseline) and expectation (modeling the main effect of expectation, i.e., predictive trials > non-predictive trials). To extract the time courses for each of the regions of interest, we thresholded the subject-specific statistical F maps for the effects of no interest from the first-level model described above at $p < 0.01$ and masked the resulting activations with the subject-specific MT+ regions obtained from the localizer or with the left DLPFC region obtained from the group analysis above.

The main goal of the DCM analysis was to investigate whether the modulation of the connection between MT+ and left DLPFC by expectation was primarily feedforward, feedback, or recurrent. Thus, we specified three different models that differed in whether expectation influenced feedforward, feedback, or both connections between MT+ and DLPFC. All models included intrinsic bilateral connections from left and right MT+ to left DLPFC (see Fig. 6A). Furthermore, we considered three different possibilities for how the stimulus itself influenced activity in and connectivity between the nodes of the network. We constructed models in which the stimulus only modulated MT+, models in which the stimulus modulated both MT+ and DLPFC, and models in which the stimulus modulated MT+, DLPFC, and the connectivity between MT+ and DLPFC. Together, this resulted in nine models (feedforward, feedback, and recurrent models, each with three possible implementations of how stimulus affected activity). We fitted each of these nine models for each subject separately. Then, using a hierarchical Bayesian approach (Stephan et al., 2009), we compared the feedforward, feedback, and recurrent families of models by computing the exceedance probability of each class of models. The exceedance probability in this analysis reflects the likelihood that a class of models explains the data better than the other class of models in the comparison. We also used Bayesian model comparison to compare all individual models across subjects at the second level.

Results

Behavioral results

The identity of the cues (valid, invalid, or neutral) significantly influenced the proportion of correct responses ($F_{(2,40)} = 12.63$, $p < 0.001$). In particular, the proportion of correct responses was higher when the cue was valid rather than neutral ($t_{(20)} = 3.09$, $p = 0.006$, average difference of 5%) and lower when the cue was invalid rather than neutral ($t_{(20)} = -3.30$, $p = 0.004$, average difference of 14%). Cue identity also significantly influenced RTs ($F_{(2,40)} = 12.57$, $p < 0.001$), although the pattern was more complex (Fig. 2A). The level of motion coherence influenced both the proportion of correct responses ($F_{(2,40)} = 95.94$, $p < 0.001$) and RTs ($F_{(2,40)} = 10.24$, $p < 0.001$), with higher motion coherence leading to faster and more accurate responses. There was also an interaction between motion coherence and cue identity for
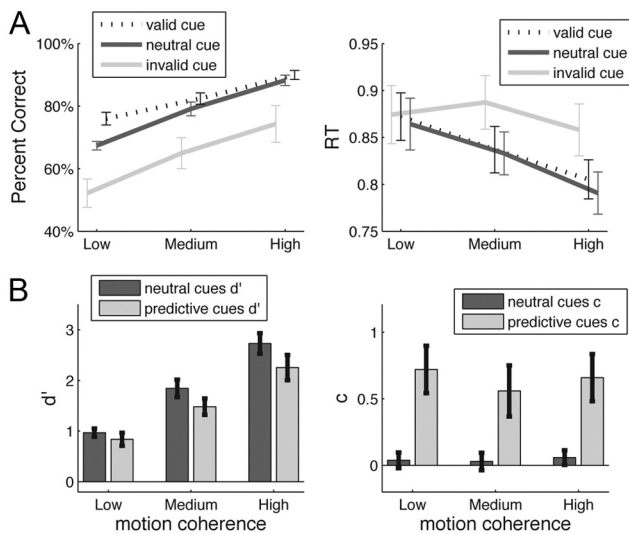
Figure 2. Behavioral results. **A**, Accuracy and reaction times are plotted as a function of motion coherence (low, medium, high) and cue type (valid, neutral, invalid). Overall, higher motion coherence led to higher accuracy and lower reaction times. Similarly, invalid cues decreased performance, whereas valid cues improved it. Error bars show the SEM. **B**, The SDT measures d' and c were computed independently for predictive (i.e., valid and invalid) and non-predictive (i.e., neutral) cues to gauge subjects' ability to do the task with/without the expectation provided by the cues. Surprisingly, d' was lower for predictive than for neutral cues. Subjects were unbiased ($c \sim 0$) when doing the task with neutral cues but shifted their criterion toward the expected percept when predictive cues were presented. Error bars show the SEM.

both accuracy ($F_{(4,80)} = 3.93$, $p = 0.013$) and RT ($F_{(4,80)} = 5.23$, $p = 0.005$).

We investigated the effect of the predictive cues on discrimination sensitivity, as measured with the signal detection measure d' (Macmillan and Creelman, 2005). The idea behind this analysis was to estimate subjects' capacity to perform the motion discrimination task, either with or without an expectation about upcoming motion direction. We computed $d'_{predictive}$ using all the valid and invalid trials and $d'_{neutral}$ using all the neutral trials. The key point to note about the procedure is that, to compute $d'_{predictive}$, we separated the predictive (valid and invalid) trials into two categories: one in which the cue indicated that the forthcoming stimulus was likely to be expanding (d'_{expand}) and one in which the cue indicated a likely contracting stimulus ($d'_{contract}$). Then, $d'_{predictive}$ was simply the average of d'_{expand} and $d'_{contract}$ (for details, see Materials and Methods). Similarly, we computed the signal detection measure of bias c for predictive ($c_{predictive} = c_{contract} - c_{expand}$) and neutral ($c_{neutral}$) trials.

Figure 2B shows that, not surprisingly, d' increased with motion coherence ($F_{(2,40)} = 88.63$, $p < 0.001$). More interestingly, the cues also influenced d' ($F_{(1,20)} = 8.66$, $p = 0.008$) such that predictive (valid and invalid) cues decreased d'. This suggests that, although subjects benefited from the predictive cue to improve overall accuracy, the underlying stimulus processing sensitivity (independent of the aid provided by the cue) became worse after the presentation of a predictive cue. The decrease in d' for predictive cues became bigger with increasing motion coherence, with the medium and high motion coherence showing a significant d' difference (medium coherence: d' difference = 0.36, $t_{(20)} = 2.69$, $p = 0.014$; high coherence: d' difference = 0.47, $t_{(20)} = 2.59$, $p = 0.018$), unlike the lowest motion coherence, which did not show a significant d' decrease (d' difference = 0.13, $p = 0.23$).

Finally, the predictive cues significantly shifted the criterion (c: $F_{(1,20)} = 11.89$, $p = 0.003$), demonstrating that subjects did

incorporate the information of the cues. Conversely, motion coherence had no effect on c ($F_{(2,40)} = 1.21$, $p = 0.31$).

Neural activity
We first identified regions in which activity during the perceptual decision was modulated by expectation. More specifically, we investigated whether there was larger activity during trials with predictive cues (valid and invalid) compared with trials with non-predictive cues (neutral). This comparison revealed larger activity for predictive trials bilaterally in the IPS and DLPFC. The anterior part of the left IPS region extended into the gyral surface posterior to the postcentral sulcus, whereas its most posterior part reached just anterior and dorsal of the transverse occipital sulcus. The right IPS activation was smaller and did not extend as laterally and anteriorly as the activation in left IPS. In view of their anatomical locations, these activations are most properly referred to as "anterior IPS" (Shulman et al., 1999). The anatomical location of DLPFC falls close to cytoarchitectonically defined Brodmann area 46 (Rajkowska and Goldman-Rakic, 1995). These regions are shown in Figure 3A, and details are provided in Table 1. We next tested for differences in these regions as a function of cue type (valid vs invalid), as well as belief updating (as indexed by criterion shift induced by the predictive cue).

Figure 3B shows the percentage signal change associated with the three cue types (invalid, valid, neutral), for each of the four regions. Interestingly, there was larger activity during invalidly cued trials than validly cued trials in the left ($t_{(20)} = 2.49$, $p = 0.01$) and right DLPFC ($t_{(20)} = 2.54$, $p = 0.01$). This activity difference was not present in either left or right IPS ($p > 0.8$). In Figure 3C, we have plotted the time courses for each trial type in each of the regions of interest.

Next, we tested whether interindividual differences in the updating of one's belief on the basis of the predictive cue during the perceptual decision was related to activity levels in the expectation-related regions of interests (IPS, DLPFC), as well in the sensory region MT+. Specifically, we investigated whether between-subject variations in brain activity differences between predictive and non-predictive trials were correlated with the extent to which subjects incorporated the predictive cues (criterion shift). Criterion shift was formalized by each individual subject's propensity to adjust his or her decision criterion based on the predictive cues (i.e., $c_{contract} - c_{expand}$; see Materials and Methods). We computed correlation coefficients between these measures for bilateral DLPFC, IPS, and MT+ (Fig. 4). Because three subjects used the cue particularly extensively and could potentially drive some of the correlations, we used a nonparametric correlation test (Spearman's rank correlation), which is unaffected by outliers. There was a positive correlation between criterion shift and brain activity in left DLPFC (Spearman's $\rho = 0.50$, $p = 0.01$) and left IPS (Spearman's $\rho = 0.61$, $p = 0.002$). In other words, subjects who used the cues more extensively also showed higher activity in parietal and prefrontal regions in the left hemisphere. Conversely, criterion shift showed a negative correlation with right MT+ (Spearman's $\rho = -0.38$, $p = 0.04$).

Effective connectivity
The previous analysis suggests that expectation modulated neural activity in dorsolateral prefrontal cortex as a function of the bias introduced by expectation. We next tested whether the interaction between sensory (MT+) and dorsolateral prefrontal cortex was also affected by expectation. In particular, we investigated whether activity changes in MT+ was associated with larger activity changes in IPS and DLPFC during trials in which subjects
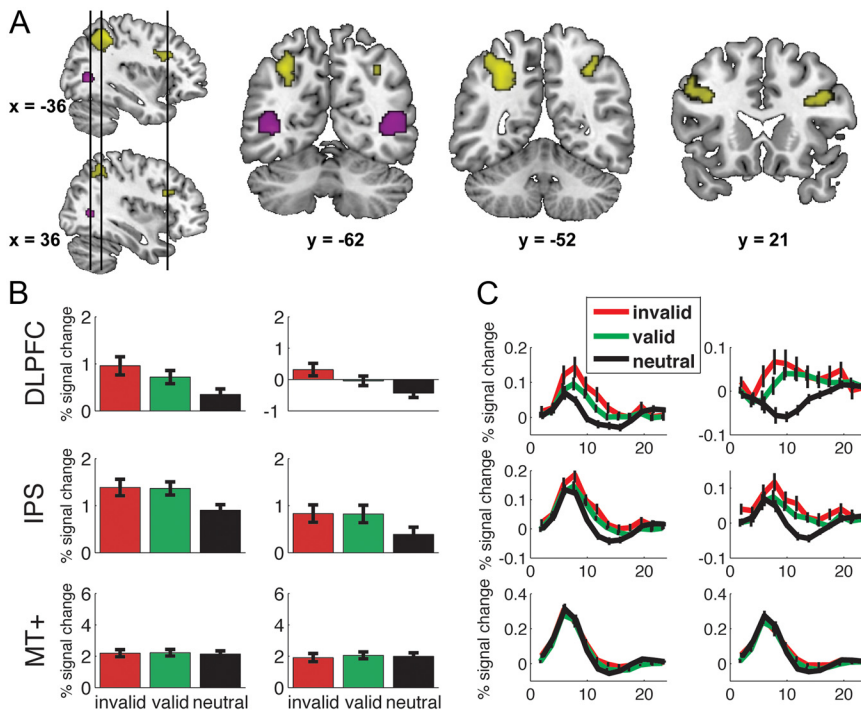
Figure 3. Neural activation differences induced by expectation. *A*, Larger activation was found bilaterally in both IPS and bilateral DLPFC for trials in which subjects had an expectation than for trials in which subjects had no expectation (shown in yellow). Bilateral MT+ (shown in purple) was functionally localized using an independent localizer for each subject. *B*, Percentage signal change is plotted for each of the three cue types (invalid, valid, and neutral) for left DLPFC, IPS, and MT+ (left column) and its right hemisphere counterpart (right column). DLPFC showed larger activity for invalidly cued trials compared with validly cued trials (both *p* values <0.02). No such difference was found for IPS (both *p* values >0.8). There were no differences in MT+ for the differently cued trials (all *p* values >0.2). *C*, Time courses for each of the six regions of interest are plotted for each trial type. Error bars represent the SEM.

Table 1. Localization of activation difference for the contrast predictive (valid + invalid) > nonpredictive (neutral)

Anatomical region	*t* value	Cluster size	Corrected *p* value	Coordinates (*x, y, z*)
Left DLPFC	4.49	257	0.001	−50, 22, 36
Right DLPFC	4.27	162	0.01	40, 20, 30
Left IPS	5.68	1016	<0.001	−40, −44, 50
Right IPS	4.90	149	0.014	32, −56, 44

Coordinates correspond to the standard MNI brain. All results are cluster-level corrected for multiple comparisons.

had a prior expectation about the forthcoming stimulus than during trials in which subjects did not have a prior expectation about the stimulus. Indeed, we observed that, when subjects had an expectation, there was a stronger effective connectivity between MT+ and left DLPFC (left MT+: $t_{(20)} = 2.19$, $p = 0.02$; right MT+: $t_{(20)} = 2.79$, $p = 0.006$) than when subjects had no expectation about the upcoming motion direction (Fig. 5). These effects were observed in neither the right DLPFC ($p > 0.07$) nor the left or right IPS ($p > 0.14$).

The effective connectivity uncovered by the above analysis could be the result of connectivity that is primarily feedforward, primarily feedback, or recurrent (both feedforward and feedback). Therefore, we performed DCM analysis to differentiate between these possibilities. We constructed families of models in which expectation modulated the feedforward, the feedback, or both the feedforward and feedback connections between MT+ and DLPFC (Fig. 6A). Each family consisted of individual models in which the stimulus was allowed to modulate (1) only MT+, (2) MT+ and DLPFC, and (3) MT+, DLPFC, and the recurrent

connections between MT+ and DLPFC (see Materials and Methods). This produced a total of nine models.

Using a hierarchical Bayesian approach (Stephan et al., 2009), we compared the feedforward, feedback, and recurrent families of models by computing the exceedance probability of each class of models. The exceedance probability in this analysis reflects the likelihood that a class of models explains the data better than the other class of models in the comparison. The results showed that the recurrent family of models explained the data best (Fig. 6B). This family had exceedance probability of 80%. The feedforward-only family of models had exceedance probability of 18%, whereas the feedback-only class of models had exceedance probability of 2%. Thus, it appears that expectation modulated both the feedforward and the feedback connections.

We also compared all nine individual models separately by examining their average posterior probabilities across subjects. The feedforward, feedback, and recurrent models in which the stimulus only influenced MT+ were generally preferred, with average posterior probabilities of 12, 10, and 44%, respectively. The models in which the stimulus directly influenced activity in both MT+ and DLPFC had average posterior probabilities of 5, 2, and 2% for the feedforward, feedback, and recurrent models, respectively. Last, models in which the stimulus modulated MT+, DLPFC, and the connectivity between MT+ and DLPFC had average posterior probabilities of 9, 5, and 10% for the feedforward, feedback, and recurrent models, respectively.

Finally, we looked at the strength of the connections between regions in the winning model (the recurrent model in which stimulus only modulates MT+) (Fig. 6C). We were interested how expectation modulated the connections between bilateral MT and left DLPFC. Expectation significantly increased the strength of the forward connection from left MT to left DLPFC ($d = 0.098$, $t_{(20)} = 3.43$, $p = 0.003$) and from right MT to left DLPFC ($d = 0.075$, $t_{(20)} = 3.07$, $p = 0.006$). Conversely, the feedback connections were negatively modulated by expectation, although the effect was not significant ($p > 0.10$).

Discussion

The likelihood of a perceptual event has a marked effect on our ability to perceive it (Bar, 2004). Here we examined how perceptual decisions about motion are influenced by inducing a prior expectation regarding the likelihood of motion direction. Behavioral results suggest that the prior cue changed subjects' expectation about upcoming motion direction, such that decisions were more accurate when subjects were provided with a valid cue (Fig. 2). Moreover, DLPFC cortex showed higher activity when subjects had a prior expectation about motion direction (Fig. 3) and a modulation of activity as a function of whether the expectation was violated or not. Furthermore, between-subject activity differences in the left DLPFC reflected individual differences in how much the expectation cue biased subjects' decisions (Fig. 4). Ad-
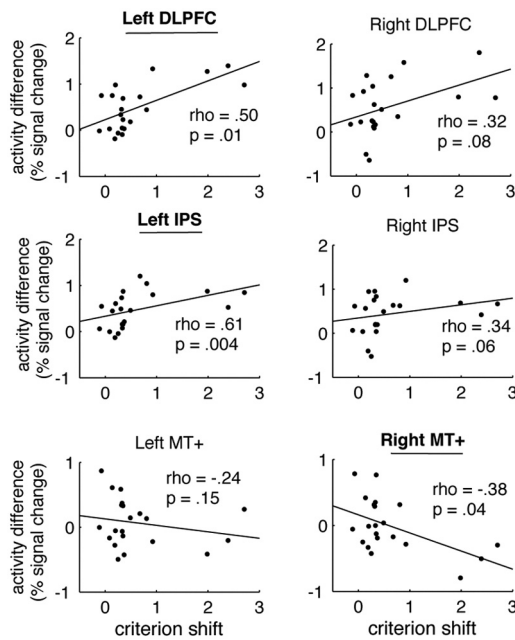
Figure 4. Brain-behavior correlation of expectation-induced bias. We observed a significant positive across-subjects correlation between the behavioral criterion shift induced by the prediction cue and the neural activity increase for predictive trials in left DLPFC (Spearman's $\rho =$ 0.50, $p = 0.01$) and left IPS (Spearman's $\rho = 0.61$, $p = .002$). Only trends were found in right DLPFC and right IPS. Conversely, this correlation was negative in left and right MT+, with the effect being significant in right MT+ (Spearman's $\rho = -0.38$, $p = 0.04$). The activity increase was computed from the contrast predictive (valid and invalid) > non-predictive (neutral) trials. Criterion shift is a measure of the degree to which subjects adjusted their decision bias (for details, see Materials and Methods). We used Spearman's rank correlation, a nonparametric test that is insensitive to extreme values in the variables. All significant correlations remain significant if Pearson's product-moment correlation was used.
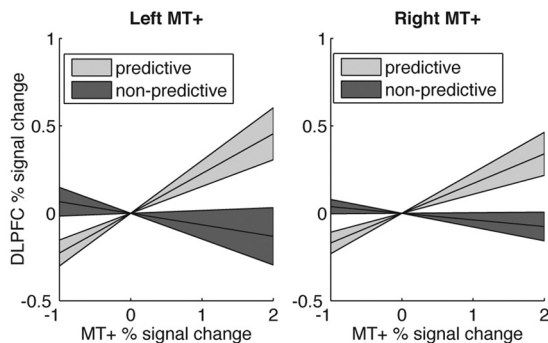

Figure 5. Effective connectivity between bilateral MT+ and DLPFC. We used a PPI analysis to look at the effective connectivity between left DLPFC and left (left panel) and right (right panel) MT+. Left DLPFC was chosen because it was sensitive to the presence and validity of the cue (Fig. 3) and correlated with the extent to which subjects shifted their criterion based on the predictive cues (Fig. 4). We tested whether the connectivity between MT+ and left DLPFC depended on the cue identity [computed from the contrast predictive (valid + invalid) > non-predictive (neutral)]. Connectivity between MT+ and left DLPFC was higher when subjects had a prior expectation about upcoming motion direction, as revealed by the difference in the slope of regression between fMRI activity. The regression lines reflect averages across subjects. The shaded regions correspond to one SEM.

ditionally, the left DLPFC showed enhanced recurrent connectivity with the motion-sensitive area MT+ (which is likely representing the sensory evidence in this experiment) when subjects had an expectation about the direction of upcoming motion (Fig. 5). Below we will interpret and discuss these results within the context of current models of perceptual decision making.

Prior expectation increases activity in dorsolateral prefrontal cortex

The prefrontal and parietal cortices are thought to play an important role in perceptual decision making (Gold and Shadlen, 2007; Heekeren et al., 2008; Kayser et al., 2010a,b; Scheibe et al., 2010). Specifically, electrophysiological studies suggest that neural activity in these regions may reflect the accumulation of evidence during the decision process (Shadlen and Newsome, 1996; Schall, 2003). When we compared trials in which subjects had a stimulus expectation with trials in which subjects had no stimulus expectation, we found heightened activity in both IPS and DLPFC. One interpretation of this finding is that the incorporation of prior expectation into the decision process changes the computations in these areas related to evidence accumulation.

However, it could be argued that the expectation cue simply induced a larger cognitive demand, because subjects needed to attend to both the cue and the stimulus for these type of trials. Although this interpretation can explain the fact that bilateral DLPFC and IPS are more active in the presence of expectation, it does not seem compatible with our exact pattern of results. In particular, the activation of DLPFC was additionally modulated by the validity of the cue: there was larger activity for invalidly cued than validly cued trials. This is consistent with a role of DLPFC in evidence accumulation, because subjects will need to accumulate more evidence for the correct motion direction when they have an invalid prior expectation. An alternative explanation for the activity difference in DLPFC between valid and invalid expectations could be that additional activity in the DLPFC arose during the invalidly cued trials because of cognitive conflict (Ridderinkhof et al., 2004). However, this alternative explanation would not predict larger effective connectivity between sensory area MT+ and DLPFC when subjects had a prior expectation. This connectivity modulation is more in line with a specific role of the DLPFC in sending its predictions to sensory area MT+, as well as integrating the prior expectation and the sensory information provided by MT+. Also, the fact that individual differences in criterion shift were correlated with the activity difference between predictive (valid and invalid) and non-predictive (neutral) cues in DLPFC is not in line with cognitive conflict. The activity in DLPFC is rather in line with previous studies that observed activity modulations in DLPFC consistent with evidence accumulation (Kim and Shadlen, 1999; Heekeren et al., 2004, 2006; Ho et al., 2009), and therefore this region may play an important role in incorporating prior expectation into perceptual decisions.

Perhaps somewhat surprisingly, we did not find an overall increase in sensory area MT+ as a function of perceptual expectation, and the activity in MT+ only showed a weak negative correlation with the extent to which each subject used the predictive cues. Top-down modulations on sensory representations are well known (Gilbert and Sigman, 2007), and indeed previous studies have shown that expectation about motion direction can boost activity in direction-selective MT+ cells (Treue and Martínez Trujillo, 1999). In line with this, a previous neuroimaging study did find a modulation of MT+ activity when subjects had an expectation about upcoming motion (Shulman et al., 1999). One important difference between the study by Shulman et al. and the current experiment relates to detection versus discrimination. Shulman et al. asked subjects to detect motion, and subjects were given a cue that indicated the direction in which coherent motion could occur (up, down, left, or right). In that task, it is plausible that the activity of relevant MT+ neurons may be upregulated without suppressing other MT+ neurons. In our task, subjects were asked to discriminate between two motion
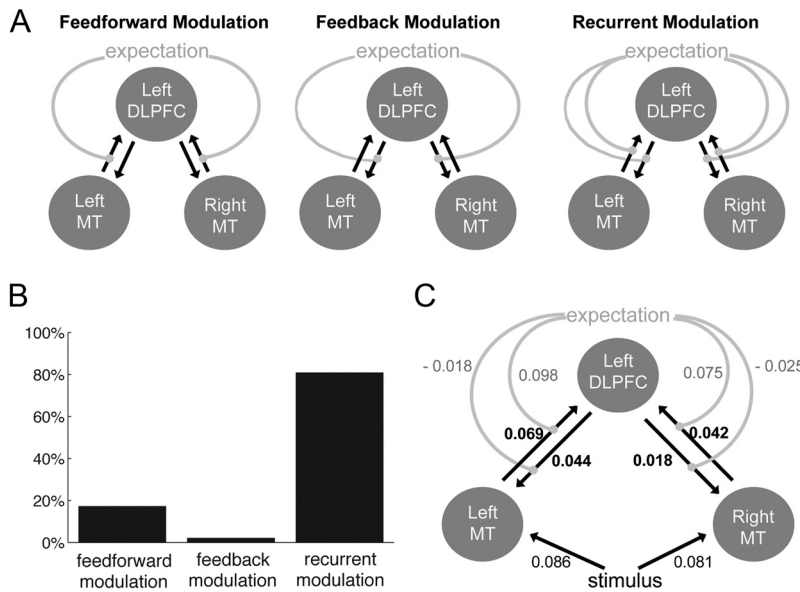
Figure 6. Dynamic causal modeling of connectivity between bilateral MT+ and left DLPFC. **A**, Three families of models tested whether expectation modulated only the feedforward, the feedback, or both connections between MT+ and DLPFC. In all models, there were recurrent intrinsic connections between bilateral MT+ and left DLPFC. The contribution of the stimulus varied between models and is not depicted in this figure (for details, see Materials and Methods). **B**, Bayesian model comparison was used to compute the exceedance probability for each of the three families of models. The exceedance probability was largest for the recurrent family of models, suggesting that expectation modulated both the feedforward and feedback connections between MT+ and DLPFC. **C**, All connections and their values are shown for the best-fitting model (the recurrent model in which the stimulus modulated only MT+). Across subjects, expectation significantly modulated the feedforward connection from left MT+ to left DLPFC ($p = 0.003$) and from right MT+ to left DLPFC ($p = 0.006$).

directions (expanding or contracting), thereby the expectation cues may have led to simultaneous enhancement and suppression of two competing neuronal assemblies, precluding an overall activity increase. This interpretation could be directly tested in future research.

Perceptual decisions and motor preparation
In many previous studies, perceptual decisions are linked to specific motor responses throughout the experiment (Tosoni et al., 2008). This means that the accumulated evidence for the perceptual decision itself does not need to be held "online" but can be directly translated into the preparation of specific motor plans. This may explain why previous studies have reported results of perceptual decision-related activity in the motor cortex (Donner et al., 2009) or in the corticostriatal network, which is thought to be related to action selection (Forstmann et al., 2010). However, studies that dissociated the perceptual decision from the response modality have observed neural activity in parietal (Bennur and Gold, 2011) and prefrontal (Heekeren et al., 2006; Ferrera et al., 2009) cortices, regions that may be more related to the perceptual decision itself rather than motor preparation. In our experiment, we have carefully minimized the possibility that our results would be contaminated by motor preparation, by informing subjects of the response mapping only after the presentation of the stimulus. Under these circumstances, prior expectation modified the activity level in the DLPFC and connectivity between MT+ and DLPFC in a manner that is in line with a role for this region in evidence accumulation. Although we did not find the same relationship in the parietal cortex, we did observe a general modulation of the presence versus absence of expectation on neural activity in the anterior intraparietal sulcus. This is generally compatible with a role of this region in the accumulation of evidence

(Bennur and Gold, 2011), in view of its response to both prior expectation and sensory evidence.

Connectivity between MT+ and DLPFC
Our PPI analysis demonstrated that expectation increased the effective connectivity between MT+ and DLPFC. Furthermore, the DCM analysis suggested that this increased connectivity is likely driven by modulations of both the feedforward and the feedback connections. Increased feedback connectivity from DLPFC to MT+ could embody a change of gain setting in sensory regions by expectation (Shulman et al., 1999) and thus provide a mechanism of how prior information could change decision making, namely by frontal regions changing the responsiveness of relevant neurons in sensory cortex. Complementing this, increased feedforward connectivity could reflect the result of this gain setting, which leads to a stronger impact of the expected sensory evidence on prefrontal cortex. Interestingly, this change of feedforward connectivity is consistent with recent neurophysiological findings (Law and Gold, 2009).

Relationship with other biasing effects on perceptual decision making
Two ways that have been traditionally used to bias decision making is by either altering base rate of certain stimulus categories or altering reward contingencies associated with certain stimulus categories. One may speculate that these top-down effects are hard to distinguish, because they have superficially similar effects on behavior (making a particular response more likely and its reaction time faster). Also, reward can change activity in both sensory (Serences, 2008) and decision-related (Rorie et al., 2010; Summerfield and Koechlin, 2010) areas. Conceptually, one can imagine that changing reward can influence behavior without changing subjects' belief as to what is actually the stimulus (e.g., "I do not really think the stimulus contains expanding motion, but since it will more likely give me high reward, I am going to answer as such anyway"). Changing prior expectation, conversely, can plausibly change subject's beliefs and potentially the content of perception (Sterzer et al., 2008). Therefore, although both manipulations change behavior, the reward structure may dictate how behaviorally relevant the stimulus is, whereas the prior expectation pertains to its likelihood (Summerfield and Egner, 2009). These differences may explain why changing prior expectation may change the interaction between sensory and decision-making processes, whereas changing reward may selectively change the decision stage.

Conclusion
We provide behavioral and neural evidence for how prior expectation biases perceptual decisions. Behaviorally, prior expectations lead to both more biased and less sensitive perception. Neuroimaging data show that this bias is implemented by in-

creasing activity in, as well as sensory drive to, the dorsolateral prefrontal cortex.

References

Bar M (2004) Visual objects in context. Nat Rev Neurosci 5:617–629.

Bennur S, Gold JI (2011) Distinct representations of a perceptual decision and the associated oculomotor plan in the monkey lateral intraparietal area. J Neurosci 31:913–921.

Brainard DH (1997) The psychophysics toolbox. Spatial Vision 10:433–436.

Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Curr Biol 19:1581–1585.

Egner T, Monti JM, Summerfield C (2010) Neural integration of top-down spatial and feature-based information in visual search. J Neurosci 30:16601–16608.

Enns JT, Lleras A (2008) What's next? New evidence for prediction in human vision. Trends Cogn Sci 12:327–333.

Esterman M, Yantis S (2010) Perceptual expectation evokes category-selective cortical activity. Cereb Cortex 20:1245–1253.

Ferrera VP, Yanike M, Cassanello C (2009) Frontal eye field neurons signal changes in decision criteria. Nat Neurosci 12:1458–1462.

Forstmann BU, Brown S, Dutilh G, Neumann J, Wagenmakers EJ (2010) The neural substrate of prior information in perceptual decision making: a model-based analysis. Front Hum Neurosci 4:40.

Friston KJ, Holmes A, Poline JB, Price CJ, Frith CD (1996) Detecting activations in PET and fMRI: levels of inference and power. Neuroimage 4:223–235.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229.

Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19:1273–1302.

Gilbert CD, Sigman M (2007) Brain states: top-down influences in sensory processing. Neuron 54:677–696.

Gitelman DR, Penny WD, Ashburner J, Friston KJ (2003) Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. Neuroimage 19:200–207.

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535–574.

Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. Nature 431:859–862.

Heekeren HR, Marrett S, Ruff DA, Bandettini PA, Ungerleider LG (2006) Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. Proc Natl Acad Sci U S A 103:10023–10028.

Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. Nat Rev Neurosci 9:467–479.

Ho TC, Brown S, Serences JT (2009) Domain general mechanisms of perceptual decision making in human cortex. J Neurosci 29:8675–8687.

Kayser AS, Buchsbaum BR, Erickson DT, D'Esposito M (2010a) The functional anatomy of a perceptual decision in the human brain. J Neurophysiol 103:1179–1194.

Kayser AS, Erickson DT, Buchsbaum BR, D'Esposito M (2010b) Neural representations of relevant and irrelevant features in perceptual decision making. J Neurosci 30:15778–15789.

Kim JN, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat Neurosci 2:176–185.

Law CT, Gold JI (2009) Reinforcement learning can account for associative and perceptual learning on a visual-decision task. Nat Neurosci 12:655–663.

Lund TE, Nørgaard MD, Rostrup E, Rowe JB, Paulson OB (2005) Motion or activity: their role in intra- and inter-subject variation in fMRI. Neuroimage 26:960–964.

Macmillan N, Creelman C (2005) Detection theory: a user's guide. New York: Cambridge UP.

Noppeney U, Ostwald D, Werner S (2010) Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. J Neurosci 30:7434–7446.

Preuschhof C, Schubert T, Villringer A, Heekeren HR (2010) Prior information biases stimulus representations during vibrotactile decision making. J Cogn Neurosci 22:875–887.

Rajkowska G, Goldman-Rakic PS (1995) Cytoarchitectonic definition of prefrontal areas in the normal human cortex. II. Variability in locations of areas 9 and 46 and relationship to the Talairach coordinate system. Cereb Cortex 5:323–337.

Ratcliff R, McKoon G (2008) The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput 20:873–922.

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004) The role of the medial frontal cortex in cognitive control. Science 306:443–447.

Rorie AE, Gao J, McClelland JL, Newsome WT (2010) Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. PLoS One 5:e9308.

Schall JD (2003) Neural correlates of decision processes: neural and mental chronometry. Current Opin Neurobiol 13:182–186.

Scheibe C, Ullsperger M, Sommer W, Heekeren HR (2010) Effects of parametrical and trial-to-trial variation in prior probability processing revealed by simultaneous electroencephalogram/functional magnetic resonance imaging. J Neurosci 30:16709–16717.

Serences JT (2008) Value-based modulations in human visual cortex. Neuron 60:1169–1181.

Shadlen MN, Newsome WT (1996) Motion perception: seeing and deciding. Proc Natl Acad Sci U S A 93:628–633.

Shulman GL, Ollinger JM, Akbudak E, Conturo TE, Snyder AZ, Petersen SE, Corbetta M (1999) Areas involved in encoding and applying directional expectations to moving objects. J Neurosci 19:9480–9496.

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

Sterzer P, Frith C, Petrovic P (2008) Believing is seeing: expectations alter visual awareness. Curr Biol 18:R697–R698.

Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. Trends Cogn Sci 13:403–409.

Summerfield C, Koechlin E (2008) A neural representation of prior information during perceptual inference. Neuron 59:336–347.

Summerfield C, Koechlin E (2010) Economic value biases uncertain perceptual choices in the parietal and prefrontal cortices. Front Hum Neurosci 4:208.

Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for forthcoming perception in the frontal cortex. Science 314:1311–1314.

Tosoni A, Galati G, Romani GL, Corbetta M (2008) Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. Nat Neurosci 11:1446–1453.

Treue S, Martínez Trujillo JC (1999) Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399:575–579.