Neuropsychologia 50 (2012) 522–529

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Pre-stimulus pattern of activity in the fusiform face area predicts face percepts during binocular rivalry

P.-J. Hsieh[a,b,*], J.T. Colas[b], N.G. Kanwisher[b]

[a] Program in Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School Singapore, Singapore
[b] Department of Brain and Cognitive Sciences, McGovern Institute, Massachusetts Institute of Technology, United States

ARTICLE INFO

Article history:
Received 5 April 2011
Received in revised form
14 September 2011
Accepted 14 September 2011
Available online 21 September 2011

Keywords:
Binocular rivalry
Pre-stimulus activity
fMRI pattern activity
Fusiform face area

ABSTRACT

Visual input is ambiguous, yet conscious experience is unambiguous. In binocular rivalry the two eyes receive conflicting images, but only one of them is consciously perceived at a time. Here we search for the neural sites of the competitive interactions underlying this phenomenon by testing whether neural pattern activity occurring before stimulus presentation can predict the initial dominant percept in binocular rivalry and, if so, where in the brain such predictive activity is found. Subjects were scanned while viewing an image of a face in one eye and an image of a house in the other eye with anaglyph glasses. The rivalrous stimulus was presented briefly for each trial, and the subject indicated which of the two images he or she preferentially perceived. Our results show that BOLD fMRI multivariate pattern activity in the fusiform face area (FFA) *before the stimulus is presented* predicts which of the two images will be dominant, suggesting that higher extrastriate areas, such as the FFA, are not only correlated with, but may also be involved in determining the initial dominant percept in binocular rivalry. Furthermore, by examining pattern activity before and after trial onset, we found that pre-trial activity in the FFA for the rivalrous face trials is no more similar to the post-trial activity for the non-rivalrous face trials than to that for the non-rivalrous house trials, indicating a dissociation between neural pattern information, which predicts a given state of awareness, and mean responses, which reflect the state of awareness ultimately achieved.

1. Introduction

All perceptual stimuli are inherently ambiguous. How then do our brains manage to produce conscious perceptual experiences of the world in which this ambiguity has been resolved? Although several imaging studies have examined the neural correlates of conscious experience when viewing bistable or ambiguous stimuli (Donner, Sagi, Bonneh, & Heeger, 2008; Hsieh & Tse, 2009, 2010; Hsieh, Caplovitz, & Tse, 2006; Hsieh, Vul, & Kanwisher, 2010; Lumer, Friston, & Rees, 1998; Polonsky, Blake, Braun, & Heeger, 2000; Schoth, Waberski, Krings, Gobbele, & Buchner, 2007; Tong & Engel, 2001; Tong, Nakayama, Vaughan, & Kanwisher, 1998), the causal relationship between neural activity and conscious experience in these studies remains unclear. Here we attempt to go beyond merely analyzing the neural correlates of consciousness by testing whether and how neural pattern activity occurring before stimulus onset can predict the initial percept in binocular rivalry.

Using functional magnetic resonance imaging (fMRI) and multivariate pattern analysis, we investigated whether and how mean responses and pattern activity in the candidate neural sites before stimulus presentation is correlated with the initial dominant percept. Subjects were scanned while viewing an image of a face presented to one eye and an image of a house to the other eye with anaglyph glasses. A rivalrous stimulus was presented briefly for each trial, and subjects were required to indicate which of the two images they perceived (Fig. 1). We sought to determine what, if any, pre-trial pattern of neural responses is predictive of a particular subsequent percept. According to one intuitive hypothesis, states of awareness are determined in part by a simple amplification of pre-trial sensory biases, and hence pre-trial activity that *resembles* the neural signature of percept A more than that of percept B would bias the subsequent percept toward percept A. For example, an above-average mean response in the fusiform face area (FFA) before trial onset might predict a greater likelihood of a face percept in an upcoming rivalrous trial. Note, however, that this hypothesis need not be true; percepts could be determined by distinct pre-trial activity in the same regions or could be predicted by activity in brain regions other than those associated with a characteristic percept. In this case the relevant predictive pre-trial activity would not resemble that of the subsequent percept

* Corresponding author.
E-mail addresses: pojang.hsieh@duke-nus.edu.sg, hsieh.pj@gmail.com (P.-J. Hsieh).

0028-3932/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2011.09.019

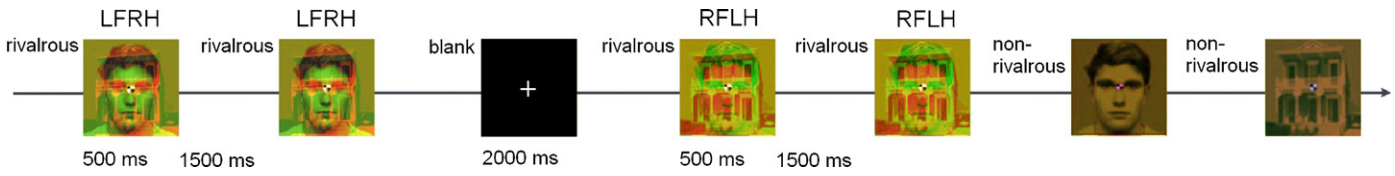

Fig. 1. For each trial subjects viewed one of four stimuli through red- and green-filtered anaglyph glasses. There were two kinds of rivalrous stimuli—one consisting of superimposed images of a green face and a red house, and the other consisting of superimposed images of a red face and a green house. The two non-rivalrous stimuli included individual yellow images of the face and the house. Stimuli appeared for 500 ms and were followed by a white fixation cross for at least 1500 ms. All subjects completed a minimum of twelve runs, each with a duration of 240 s. The order of the trials was optimized within each run.

at all. Furthermore, we asked whether any predictive neural activity takes the form of changes in the mean responses of ROIs or in pattern information in those ROIs. A number of prior studies have shown dissociations between mean responses and pattern information, including several cases in which pattern information can discriminate between conditions that cannot be discriminated on the basis of mean responses (Haxby et al., 2001; Kamitani & Tong, 2005) and even cases in which pattern information can discriminate between conditions when no net mean response is observed (Harrison & Tong, 2009; William et al., 2008).

Here we investigate this possibility and show that pre-trial fMRI pattern activity in the FFA predicts which of two rivaling percepts will dominate. Furthermore, we found that pre-trial pattern activity in the FFA for the rivalrous face trials is no more similar to the post-trial activity for the non-rivalrous face trials than to that for the non-rivalrous house trials, suggesting that the spatial pattern information reflects endogenous neural activity, whereas the ultimate perceptual decision is neurally manifested as mean activation after stimulus onset. These findings show that spatiotemporal information in multivariate patterns of fMRI activity may constrain theories of human information processing by (1) going beyond merely analyzing neural correlates to approach neural causes of perceptual awareness and (2) revealing how neural representations, captured with fMRI as multivariate patterns, evolve before and after stimulus onset.

2. Materials and methods

2.1. Participants

Twelve subjects between 18 and 30 years old participated in the fMRI study and were paid 60 dollars per session. Seven other young adult volunteers participated in the first psychophysical experiment outside of the scanner, which tested whether a stimulus duration of 500 ms sufficiently induces rivalry, and were paid 5 dollars. Another seven adults participated in the second psychophysical experiment, which tested the feasibility of a slow event-related design, and were paid 15 dollars. All subjects were healthy and right-handed and had normal or corrected-to-normal visual acuity. All subjects gave written consent within a protocol passed by the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects.

2.2. Experimental procedures

Scanning was performed in the McGovern Institute at the Massachusetts Institute of Technology in Cambridge, MA, with the Athinoula A. Martinos Imaging Center's 3T Siemens Trio scanner. Functional MRI runs were acquired using a gradient-echo, echo-planar sequence (TR = 2 s, TE = 40 ms, $3.1 \times 3.1 \times 3.1$ mm + 10% spacing). Note that for the time point of "0 s," for example, fMRI data were actually acquired between 0 and 2 s. 32 slices were collected with a 12-channel head coil. Slices were oriented roughly perpendicular to the calcarine sulcus and covered the whole brain.

For each trial subjects viewed one of four stimuli through red- and green-filtered anaglyph glasses. There were two kinds of rivalrous stimuli—one consisting of superimposed images of a green face and a red house, and the other consisting of superimposed images of a red face and a green house. The two non-rivalrous stimuli included individual yellow images of the face and the house, which were generated by combining the original red and green images. Each of the four stimuli appeared in 25% of the trials. All stimuli were presented against a black background, centered on the fixation point, and subtended $1.75^\circ \times 1.75^\circ$ of visual angle. Stimuli were presented for 500 ms, a duration that induces a behavioral state previously classified as predominantly rivalrous (Fox & Check, 1972; Williams, Morris, McGlone, Abbott, & Mattingley, 2004; Wolfe, 1983) but is brief enough to prevent within-trial perceptual switching from occurring (see Fig. 5 for psychophysical results). Stimuli appeared at the beginning of a 2-s scanning repetition (TR) and were followed (for at least 1500 ms) by a white fixation cross subtending $0.35^\circ \times 0.35^\circ$. Note that our rivalrous stimuli were presented only for 500 ms in each trial, as opposed to being continuously on the screen as in Tong et al. (1998). As a result, the rivalrous percept is naturally shorter and less stable than what had previously been reported.

While being scanned, all subjects completed a minimum of twelve runs, each with a duration of 240 s. Prior to the functional scanning, subjects completed dynamic psychophysical testing with a staircase procedure to determine the luminance values of the red and green channels for which the rivalrous stimulus could be perceived as a face or a house with equal frequency (i.e. approximately half of the trials perceived as a face). The order of the trials was optimized within each run using the optimal sequencing program *Optseq2* (NMR Center, Massachusetts General Hospital, MA, USA). One third of the total scanning time consisted of null events with variable duration (2–6 s) that were randomly inserted between trials. For each stimulus presentation, subjects were required to press one of two buttons on a response box (two-alternative forced choice; 2AFC) to indicate which of the two images (face or house) he or she perceived preferentially.

2.3. ROI identification

Functional localization of two of the regions of interest (ROIs) was based on three independent runs of 20-s blocks with grayscale images of faces, scenes, common objects and scrambled objects (four blocks per category per run). The fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997) was defined as the region of the fusiform gyrus that responded more strongly to images of faces than to images of intact objects ($p < 10^{-4}$). The parahippocampal place area (PPA; Epstein & Kanwisher, 1998) was defined as the region of the parahippocampal gyrus that responded more strongly to images of scenes than to images of intact objects ($p < 10^{-4}$).

The foveal confluence (FC; Dougherty et al., 2003) was identified as an ROI from two runs of a retinotopic localizer scan. It was defined as the small region at the posterior end of the calcarine sulcus that responded more strongly to flickering checkerboards presented in the center of the visual field than to those presented in the periphery of the visual field ($p < 10^{-4}$).

2.4. Data analysis

Data analysis was conducted using the fMRI software package *FreeSurfer* (http://surfer.nmr.mgh.harvard.edu/) and MATLAB (MathWorks). The processing steps for both the localizer and experimental runs included motion correction and intensity normalization. Processing for the localizer runs also included spatial smoothing with a 6 mm kernel. A gamma function with *delta* = 2.25 and *tau* = 1.25 was used to estimate the hemodynamic response for the localizer runs. For the experimental runs, the time courses were obtained using a finite impulse response (FIR) model without assuming a particular hemodynamic response function, which is optimal for identifying uncontaminated pre-trial signals. Trials in excess of the minimum number of trials between the two conditions ("face" and "house" when grouping by category or "left" and "right" when grouping by eye) were excluded from analysis to ensure that each condition contained an equal amount of data.

To avoid contamination of the pre-trial blood oxygen level-dependent (BOLD) signal by the response to a previous trial (Leopold, Wilke, & Maier, Logothetis, 2002; Pearson & Brascamp, 2008), we eliminated trials as needed to balance the trial histories for each condition post hoc, going back one trial in each case. For example, when we examined whether pre-trial activity differs between rivalrous face and house trials, we grouped the rivalrous stimuli by perceived category such that the two compared conditions were equally preceded by all six possible percepts/stimuli—namely, rivalrous face to the left eye, rivalrous face to the right eye, rivalrous house to the left eye, rivalrous house to the right eye, non-rivalrous face, and non-rivalrous house. If a trial was preceded by fixation, it would instead be categorized according to the percept/stimuli of the trial preceding prior fixation. Note that such equating ensures that the activity of the current trial is not dependent upon the activity of the previous trial.

To test whether the percept/stimulus of the second trial back (T_{-2}) can influence the percept of the current trial (T_0), we compared the following two likelihoods: (1)

the likelihood of state X inducing state *Z* though state Y (X → Y → Z) and (2) the likelihood of any state other than *X* inducing state *Z* though state Y (∼X → Y → Z). For example, given $T_{-2} = E_L C_F$ (eye dominance: left, category perceived: face) and $T_{-1} = E_L C_F$, we evaluated the probability of finding $T_0 = E_L C_F$, as well as the probability of finding $T_0 = E_L C_F$ given $T_{-2} = \sim E_L C_F$ and $T_{-1} = E_L C_F$. A paired *t*-test was then performed across subjects to determine whether there is a difference between the two probabilities. Observing a significant difference would indicate that the percept/stimulus of the second trial back does, in fact, have an impact on the percept of the current trial. A total of 168 pairs of comparisons were performed, permuting across 6 possibilities for X ($E_L C_F$, $E_R C_F$, $E_L C_H$, $E_R C_H$, non-rivalrous face, non-rivalrous house), 7 possibilities for Y ($E_L C_F$, $E_R C_F$, $E_L C_H$, $E_R C_H$, non-rivalrous face, non-rivalrous house, fixation), and 4 possibilities for Z ($E_L C_F$, $E_R C_F$, $E_L C_H$, $E_R C_H$). Only comparisons for some sequences in which Y is fixation showed a significant difference ($p < 0.05$ with Bonferroni correction); no other comparisons were close to significant ($p > 0.15$). The result of this analysis is consistent with the notion that the percept/stimulus of the second trial back (T_{-2}) did not influence the percept of the current trial (T_0), excluding when T_{-1} was fixation. Therefore, when eliminating trials as needed to balance the trial histories for each condition, we used only one previous trial in each case. If a trial was preceded by fixation, it would be categorized according to the percept/stimulus of the trial preceding prior fixation (e.g. if a sequence of trials was $E_L C_F$ → fixation → $E_L C_F$, it would be categorized as $E_L C_F$ → $E_L C_F$).

To further verify that the percept/stimulus of T_{-2} does not influence that of T_0, we additionally examined the trials already balanced according to the aforementioned methods to determine whether or not the two conditions being compared were preceded by disproportionate numbers of percepts/stimuli in T_{-2}. When grouping percepts into conditions according to category, the results indicate that the rivalrous face and house percepts are not preceded (in T_{-2}) by significantly different amounts of "face" trials (including both rivalrous and non-rivalrous faces) relative to "house" trials (including both rivalrous and non-rivalrous houses; $p > 0.05$). Specifically, across the trials used for the main analysis (after controlling for T_{-1}), rivalrous face percepts for T_0 were preceded by face percepts on T_{-2} in 17.3% of trials and by house percepts in 16.7% of trials ($p > 0.05$). Rivalrous house percepts for T_0 were preceded by face percepts on T_{-2} in 16.4% of trials and house percepts in 17.3% of trials ($p > 0.05$).

BOLD fMRI time courses for each condition were obtained with an FIR model. Correlation analysis was conducted on obtained time course signals for each time-point in each voxel with a standard multivariate pattern analysis method (Haxby et al., 2001). For normalization the mean response in each voxel across all conditions was subtracted from the response to each individual condition for each half of the data before the correlation values were calculated. Data were split into odd runs and even runs, and spatial patterns were extracted from each subset of data for each ROI for the six original conditions (rivalrous green face percept, rivalrous red face percept, rivalrous red house percept, rivalrous green house percept, non-rivalrous face stimulus, and non-rivalrous house stimulus). Within each ROI we then computed the Pearson correlation coefficients between the spatial patterns of the non-rivalrous face and house stimulus conditions, between the spatial patterns of the rivalrous face and house percept conditions, and between the spatial patterns of the rivalrous left- and right-eye dominant conditions. These correlations were computed for each subject and then averaged across subjects by condition. When analyzing pre-trial pattern activity, we only examined the one preceding time-point (TR_{-1}) because trial history was balanced for only one trial back. In the case of the post-trial patterns, we averaged the mean of the three time-points (TR_{1-3}) enveloping the peak of the time-course of the delayed hemodynamic response. Note that our method is equivalent to that used by Haxby et al. (2001), in which a given ROI is deemed to contain information about discrimination of two given stimuli if the pattern of response across voxels in that region is more similar for two response patterns produced by the same stimulus than for another pair produced by two different stimuli—that is, if mean($r[A1,A2]$, $r[B1,B2]$) > mean($r[A1,B2]$, $r[A2,B1]$). Successful discrimination of this sort between the spatial patterns of activation for two behaviorally defined conditions that occurs before stimulus onset can further be interpreted as a predictive neural correlate of post-stimulus behavior.

2.5. First psychophysical experiment

To test whether a stimulus duration of 500 ms sufficiently induces rivalry, we compared stimulus durations of 500 ms versus 1000 ms and asked subjects to report dominance ratings and perceptual stability outside of the scanner. Luminance values of the red and green channels for each stimulus were determined with the same staircase procedure from the original fMRI experiment. Other stimulus properties and procedures were identical to those used in the fMRI study, except as follows. All subjects completed two blocks, in which stimuli were presented for either 500 or 1000 ms. Over the course of each block, subjects viewed 50 presentations of each rivalrous stimulus and 10 presentations of each non-rivalrous stimulus. The presentation order of the trials was randomly counterbalanced within each block, and the order of the blocks was randomized across subjects. In addition to indicating whether they preferentially perceived the face or the house, subjects were also required to report perceptual stability—that is, whether or not the dominant percept remained consistently dominant throughout a given trial (2AFC). Finally, subjects rated the dominance of the dominant percept on a scale of 50% (no dominance) to 100% (complete dominance).

2.6. Second psychophysical experiment

We essentially repeated our original fMRI experiment outside of the scanner with a less dense stimulus presentation paradigm. Over the course of ten runs (approximately 1.2 h per subject), subjects viewed 50 presentations of each of the four rivalrous and non-rivalrous stimuli with a substantially longer inter-stimulus interval (ISI) of 19.5 s. As before, subjects were required to report whether they preferentially perceived the face or the house (2AFC). We computed the likelihoods that the rivalrous percept for a given trial was preceded by either the same perceived stimulus category or the same dominant eye, as reported in the previous trial.

3. Results

As expected, during viewing of the non-rivalrous stimuli, the fusiform face area (FFA) and the parahippocampal place area (PPA) were associated with stronger BOLD responses to the perceived images of the face relative to the house ($p = 0.003$) and the house relative to the face ($p < 0.001$), respectively, 4 s after the onset of the stimuli (Fig. 2). A slightly higher BOLD response to the house images was also observed in the foveal confluence (FC; $p = 0.047$). Also as expected, pattern analysis found that the response patterns in the FFA ($p = 0.014$), the PPA ($p < 0.001$), and the FC ($p = 0.005$) contained information relevant to the non-rivalrous stimuli from 2 to 6 seconds after stimulus onset—i.e., near the peak of the BOLD responses. One-sample *t*-tests comparing the pre-trial pattern information with 0 did not reveal significant information in any of the ROIs for non-rivalrous trials ($p > 0.05$), as expected.

For the rivalrous stimuli, when combining the data according to perceived category (whether the rivalrous stimulus was perceived as a face or a house), the mean BOLD response is in general less differentiable between conditions when compared with the non-rivalrous conditions. The FFA is still associated with a stronger BOLD response to a perceived (rivalrous) face than to a perceived (rivalrous) house 4 s after onset ($p = 0.035$), but the PPA and FC did not respond differentially for different rivalrous percepts ($p > 0.05$) (Fig. 3). As for the pattern analysis, our results show no category information in any of the ROIs after stimulus presentation ($p > 0.05$). This finding provides further evidence that our data are not contaminated by trial history because it is highly unlikely that a pattern of activity that does not encode the percept for one trial would encode the percept for future trials. In general, the preferential response is weaker than what was reported by Tong et al. (1998). We suspect that this situation resulted from the brevity (500 ms) of the rivalrous stimulus, such that the suppression was not complete. Nevertheless, even if the suppression is only partial, we can still inquire about the differences in pre-trial pattern activity that might engender a small perceptual/behavioral bias in the direction of the dominant image.

Indeed, the response patterns in the FFA contained information about the upcoming dominant percept 2 s before stimulus onset ($p = 0.009$). A paired-sample *t*-test between rivalrous and non-rivalrous pattern information at time $t = -2$ s in the FFA demonstrates that the predictive information is significantly greater in the rivalrous case ($p = .026$). We did not find analogous predictive pattern information in the PPA, perhaps because the face is the more attentionally salient of the two stimuli (such that its neural representation more strongly determines the percept than does the neural representation of the house stimulus) and also because houses are not optimal stimuli for eliciting PPA activity (Epstein & Kanwisher, 1998). The presence of pattern information in pre-trial activity was only observed in the FFA—and not in the FC—suggesting a greater causal role for higher extrastriate areas than early visual areas in determining the initial perceived category in binocular rivalry.

Thus, pre-trial pattern activity in the FFA is predictive of which stimulus category (face or house) will be dominant. What is the nature of this pre-trial activity, and how exactly does it lead to

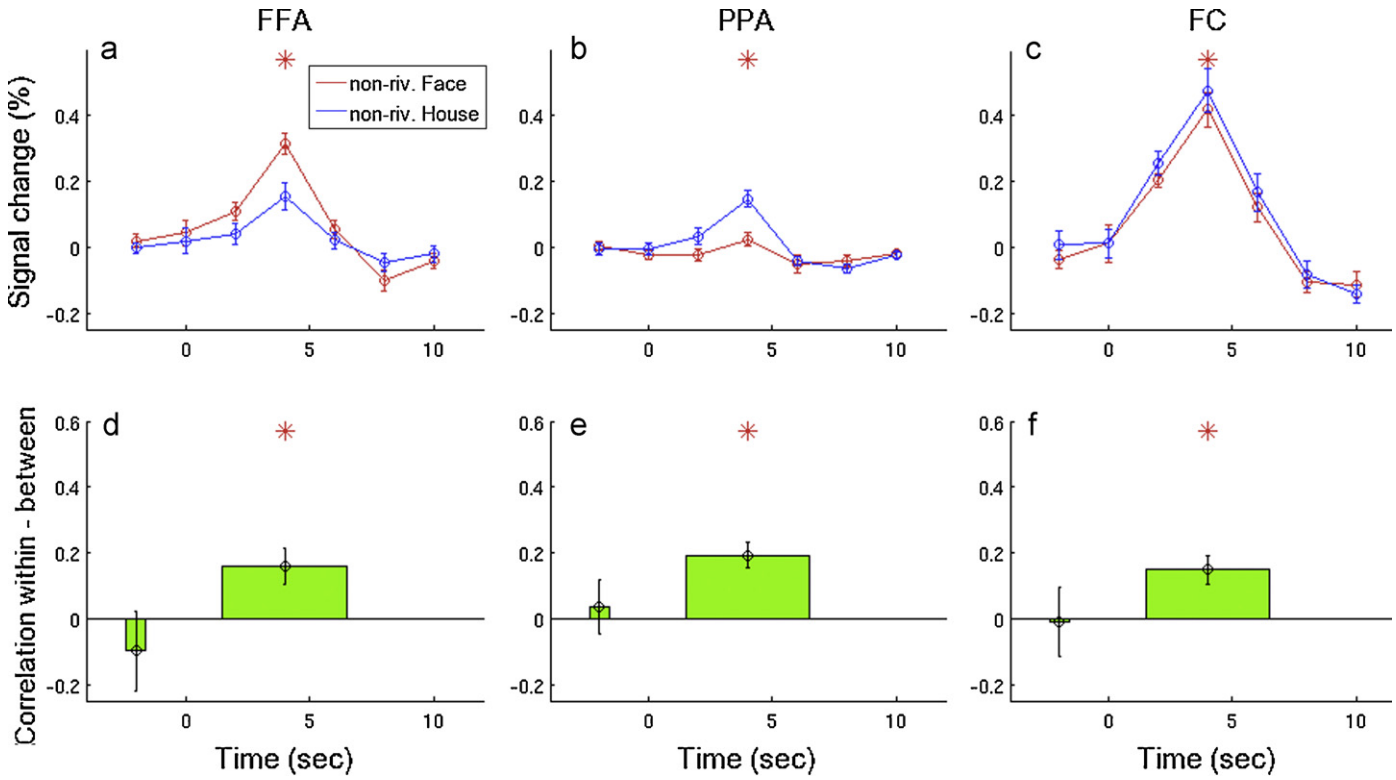

Fig. 2. Results for the non-rivalrous conditions. During viewing of the non-rivalrous stimuli, the FFA, the PPA, and the FC were associated with differential mean BOLD responses to the two stimuli 4 seconds after stimulus onset ($p = 0.003$, $p < 0.001$, and $p = 0.047$, respectively). Results from the pattern analysis show that the response patterns in the FFA, the PPA, and the FC contained stimulus-relevant information after the onset of the stimuli ($p = 0.014$, $p < 0.001$, and $p = 0.005$, respectively). No significant pattern information was found in the pre-trial activity of any of the ROIs ($p > 0.05$). Asterisks indicate significance, and error bars indicate standard errors across subjects.

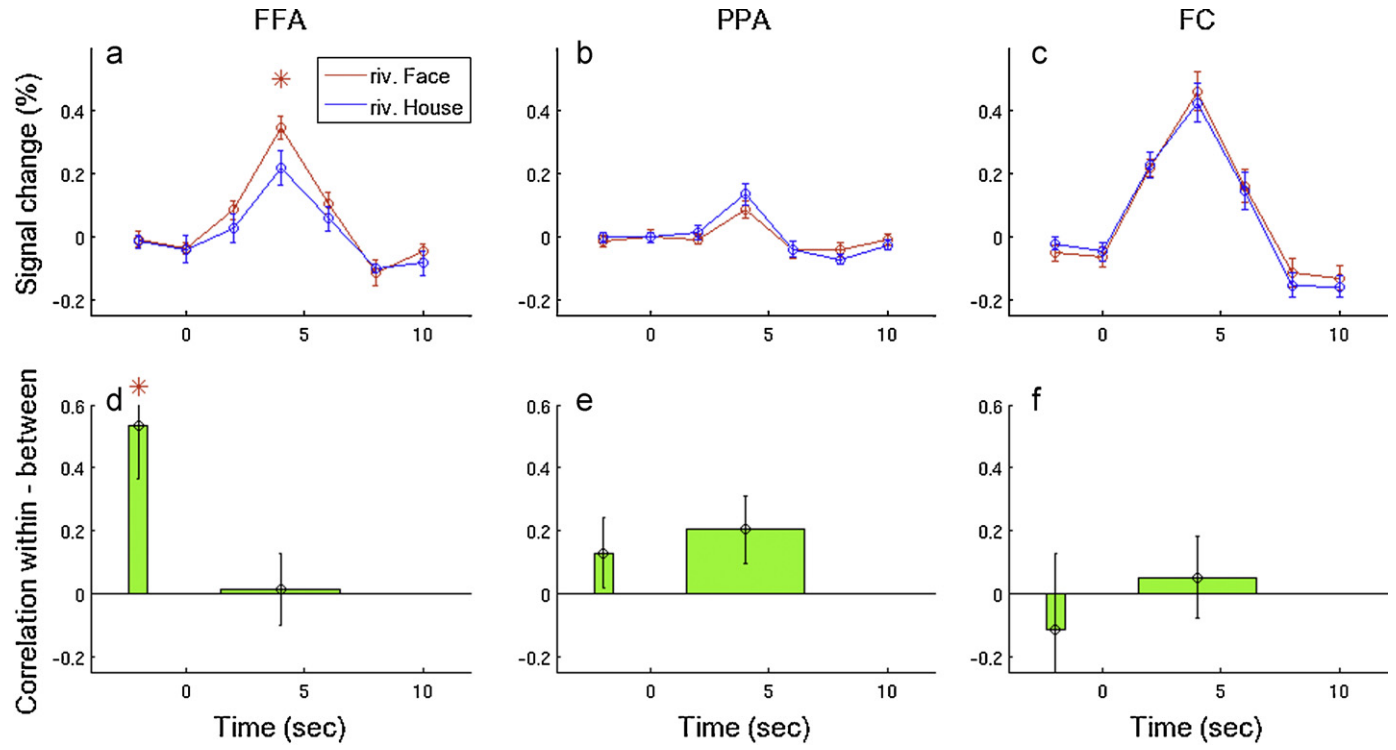

Fig. 3. Results for the rivalrous conditions (grouped according to stimulus category). Although the mean BOLD responses in (a) and (b) are weaker in general when compared with the non-rivalrous conditions, the FFA is nonetheless associated with a more robust BOLD response to the face percept 4 seconds after stimulus onset ($p = 0.035$). Response patterns in the FFA contained information about the upcoming dominant percept 2 s before stimulus onset ($p = 0.009$). Asterisks indicate significance, and error bars indicate standard errors across subjects.

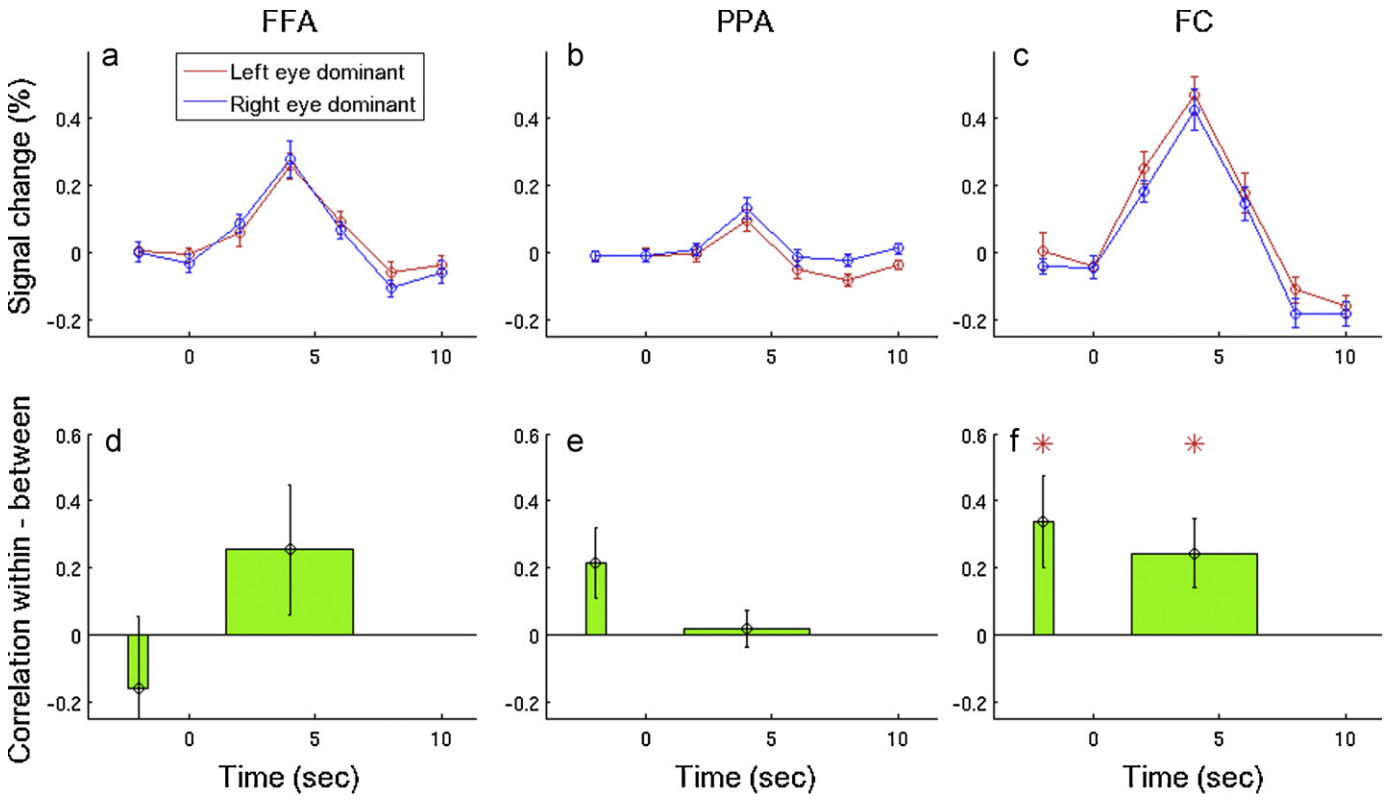

Fig. 4. Results for the rivalrous conditions (grouped according to eye dominance). No preferential mean BOLD response was observed in any of the areas ($p > 0.05$). Response patterns in the FC contained information about the upcoming dominant percept 2 s before stimulus onset ($p = 0.032$) and persisted after the onset of the stimuli ($p = 0.036$). Asterisks indicate significance ($p < 0.05$), and error bars indicate standard errors across subjects.

one percept over another? One intuitive hypothesis is that predictive patterns resemble and hence contribute to biasing neural responses toward the pattern of response that occurs when a given non-rivalrous stimulus is presented. However, for the FFA, evidence against this hypothesis comes from our finding that pre-trial activity for the rivalrous face trials is no more similar to the post-trial activity for the non-rivalrous face trials than to that for the non-rivalrous house trials ($p = 0.71$). In addition, pre-trial activity for the rivalrous face trials is no more similar to the post-trial activity for the same rivalrous face trials than to that for the rivalrous house trials ($p = 0.16$). Given that our other analyses show that the current percept is more strongly associated with a particular mean response in the FFA and PPA than with the pattern information in these regions during that perceptual experience, pre-trial pattern information apparently influences the subsequent percept by biasing subsequent post-trial mean responses, as opposed to patterns.

In a separate analysis we asked whether early or higher visual areas are involved in determining which eye will become dominant, regardless of which stimulus category is perceived. We combined the data according to eye dominance (whether the right- or left-eye input was perceived) and re-balanced the trials to ensure matched trial histories for eye dominance in T_{-1}. The results indicate that none of the ROIs exhibited a differential mean BOLD response (Fig. 4). However, results of our pattern analysis show that the response patterns in the FC contained information about the upcoming dominant percept 2 s before stimulus onset ($p = 0.032$). This pattern information persists after stimulus onset ($p = 0.036$) and is observed in the FC but not higher-level areas. Furthermore, pre-trial activity for the left-eye-dominant trials is more similar to the post-trial activity for the left-eye-dominant trials than to that for the right-eye-dominant trials ($p < 0.001$), suggesting that pre-trial activity patterns in the FC might reflect the activity of ocular dominance columns that can be maintained and amplified or instead otherwise bias post-stimulus eye dominance.

However, further analyses show that alternative accounts of these eye bias predictive phenomena are also possible. Specifically, we checked if the two conditions being compared, which were balanced based on T_{-1}, were preceded by disproportionate numbers of percepts/stimuli in T_{-2}. The results indicate that the left-eye-dominant trials and right-eye-dominant trials are preceded (in T_{-2}) by significantly different amounts of left-eye-dominant trials and right-eye-dominant trials overall ($p < 0.05$), respectively. This result reveals that, although the pre-trial pattern information in the FC is *indicative* of which eye will be dominant, this information may not be *predictive* because it could reflect the responses to previous trials (with hemodynamic lag).

To test whether a stimulus duration of 500 ms is sufficient to produce rivalry, we compared stimulus durations of 500 ms versus 1000 ms and asked subjects to report dominance ratings and perceptual stability outside of the scanner. Data averaged across subjects for both the rivalrous and non-rivalrous stimuli are shown in Fig. 5. Comparisons were made with a 1000 ms condition to address the concern that 500 ms could possibly be suboptimally brief. The results demonstrate that the dominance ratings for 500 ms presentations were around 85% (rated by the subjects on a scale of 50%, or no dominance, to 100%, or complete dominance) and were significantly greater than the no-dominance value of 50% ($p < 10^{-4}$). Moreover, perceptual stability for a 500 ms presentation duration was significantly greater ($p = 0.044$) than that for 1000 ms—that is, there were fewer perceptual switches within a 500 ms presentation.

One might argue that a slow event-related design would make for a good control experiment because there might not be any cross-trial effect with a longer ISI. To investigate this possibility, we ran another psychophysical experiment to test whether trial history

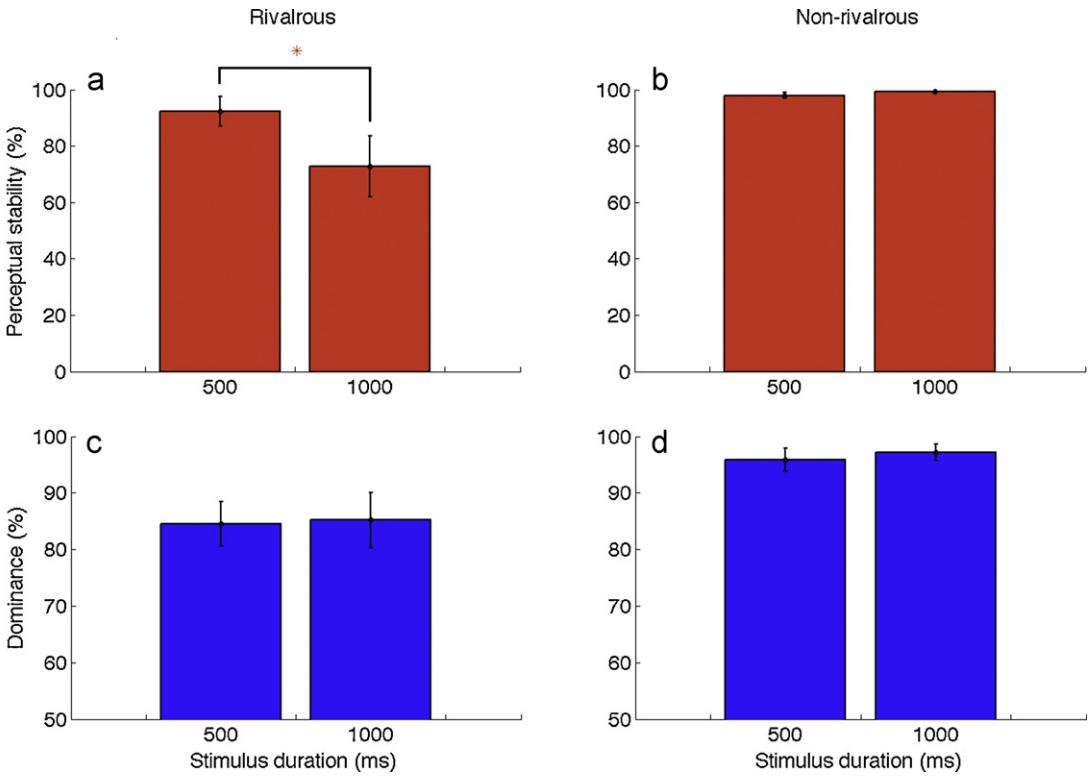

Fig. 5. Results of psychophysical experiments. Data averaged across subjects indicate that, for the rivalrous stimuli, perceptual stability (i.e., whether or not the dominant percept remained constant through the trial) was significantly greater for the 500 ms condition than for 1000 ms ($p = 0.044$). The dominance ratings (indicating perceived dominance of the dominant percept) for 500 ms presentations were around 85% and were significantly greater than the no-dominance value of 50% ($p < 10^{-4}$). For the control (non-rivalrous) stimuli, perceptual stability and dominance ratings were close to 100% as expected.

still affects the rivalrous percept of the current trial even when the ISI is 19.5 s long. Results demonstrate that the rivalrous percept for a given trial is more likely to be preceded by a rivalrous trial with the same perceived category ($p = 0.011$) and also more likely to be preceded by a rivalrous trial with the same dominant eye ($p = 0.012$). This means that very few trials in each condition (fewer than 7 trials) could be analyzed for each subject after applying post hoc trial balancing. These findings justify our choice of a faster presentation paradigm in lieu of a slow event-related design.

4. Discussion

How do rivalrous stimuli enter and exit states of dominance and suppression while vying for consciousness? Here we investigate not just the neural correlates of consciousness, but their neural antecedents, in an effort to ask which forms of neural activity play a role in determining what will be seen. Our results show that pre-stimulus fMRI pattern activity in the FFA predicts which category will be perceived, suggesting that pre-trial neural activity within higher-level visual cortex may serve to determine the initial dominant percept during binocular rivalry. Note that the differential pre-trial pattern activity is unlikely to be due to an attentional bias (Mitchell, Stoner, & Reynolds, 2004), mental imagery (Pearson, Clifford, & Tong, 2008), neural adaptation (Blake, 1989; Dayan, 1998; Freeman, 2005; Laing & Chow, 2002; Lankheet, 2006; Lehky, 1988; Mueller, 1990; Sugie, 1982; Wilson, 2003), or effects of perceptual stabilization after stimulus removal (Sterzer & Rees, 2008) because (1) these factors are presumably associated with the previous percepts/stimuli, which were controlled for after balancing trial histories, and (2) we did not find a pre-trial difference in mean activation. Moreover, our data indicate that the pre-trial pattern activity in the FFA does not resemble that of post-trial stimulation with either a rivalrous or non-rivalrous face ($p > 0.05$), suggesting that this predictive effect does not imply that states of awareness are determined by a simple amplification of a pre-trial sensory bias in pattern information. Instead, pre-trial pattern information likely influences the subsequent percept by biasing subsequent mean responses.

These results also provide intriguing clues about which kinds of information are represented in mean responses across whole ROIs and which kinds of information are represented in patterns of responses within those ROIs. Specifically, we found a dissociation in which pattern information predicts subsequent percepts, but those percepts themselves are more strongly manifested in mean responses. Thus, any differences between pattern information and mean responses are not merely consequences of differential statistical sensitivity. Several questions remain to be resolved in future studies. First, what might be the differences between the representations maintained by mean responses and those maintained by pattern information, given that pattern information is predictive and mean information reflects the ultimate percept? Is pattern information more reflective of neuronal activity at a sub-voxel scale (Op de Beeck, 2010), as suggested by performance in decoding grating orientation from patterns of activity in V1 (Haynes & Rees, 2006; Kamitani & Tong, 2005)? Is pattern information more likely to reflect subthreshold synaptic activity than spiking activity, such that overall metabolic demand is reduced (Heeger & Ress, 2002)? Does the percept-associated post-trial activation overshadow any underlying low-amplitude pattern information? Second, how does the observed pre-trial pattern difference come to exist? One account that has been proposed is the presence of low-frequency fluctuations in the temporal autocorrelation of neural signals (Haynes, 2011; Leopold, Murayama, & Logothetis, 2003). Another possible cause of the pre-trial pattern difference is varying degrees of noise across cortex at any given point in time. Such variations in noise could result from purely

endogenous neural noise, from neural adaptation that recovers randomly and unevenly, or from some combination of the two. These alternatives can be tested in future experiments by investigating, for example, whether trials in which the pre-trial FFA pattern predicts one percept (e.g., face) have a relatively greater post-trial stimulus-evoked response. Third, how precisely does this pre-trial pattern difference bias the subsequent percept, if not by mere amplification of an identical signal? One hypothesis is that some neurons within high-level visual areas might be more effective in biasing the subsequent percept than others (Heekeren, Marrett, & Ungerleider, 2008). For example, perhaps some neural populations are engaged in perceptual decision-making and are more spontaneously active. Such neural populations might engender a stronger post-stimulus mean BOLD response and bias the subsequent percept toward a region's preferred stimulus. In contrast, when other neural populations are more spontaneously active, this might lead to weaker post-stimulus mean BOLD activation and hence a bias in the other direction. In such situations, pattern information, but not mean responses, might predict later states of awareness, which would themselves be manifested as mean responses, not pattern information. These interpretations are also consistent with recent findings showing that neural patterns evolve across time (Crowe, Averbeck, & Chafee, 2010; Meyers, Freedman, Kreiman, Miller, & Poggio, 2008) and predict subsequent behaviors (Churchland, Cunningham, Kaufman, Ryu, & Shenoy, 2010). One can further test this hypothesis by examining how pattern information evolves across time and whether there is a threshold in the spatiotemporal trajectory of pattern evolution that contributes to different perceptual decisions.

Our finding that baseline fMRI pattern activity in the FFA predicts the subsequently perceived category is consistent with evidence favoring the interpretation of binocular rivalry as a high-level and representation-based process (Alais & Blake, 1999; Diaz-Caneja, 1928; Dörrenhaus, 1975; Logothetis, Leopold, & Sheinberg, 1996; Logothetis, 1998; Lumer et al., 1998; Tong et al., 1998; Yu & Blake, 1992). In contrast, other evidence supports the view of low-level and eye-based rivalry (Blake, Westendorf, & Overton, 1981; Fox & Check, 1968, 1972; Fukuda, 1981; Haynes & Rees, 2005; Lee, Blake, & Heeger, 2005; Meng, Remus, & Tong, 2005; Moutoussis, Keliris, Kourtzi, & Logothetis, 2005; O'Shea & Crassini, 1981; O'Shea, 1987; Polonsky et al., 2000; Smith, Levi, Harwerth, & White, 1982; Tong & Engel, 2001; Wales & Fox, 1970; Wunderlich, Schneider, & Kastner, 2005). Importantly, the two hypotheses are not mutually exclusive; it is likely that binocular rivalry involves parallel activity at multiple levels of cortical processing (Alais & Melcher, 2007). Low-level interocular competition may exist between monocular neurons in primary visual cortex (Blake, 1989; Tong, 2001) or the lateral geniculate nucleus (Lehky, 1988), and high-level inter-representation competition may exist in higher brain areas (Leopold & Logothetis, 1996; Logothetis et al., 1996). Further research and a more coherent model are necessary to resolve the ongoing debate about the regions that participate in rivalry (Blake & Logothetis, 2002; Freeman, 2005; Tong, Meng, & Blake, 2006; Wilson, 2003). Although our findings do not resolve this debate, they do establish that the neural events underlying rivalry may be initiated in higher-level extrastriate areas.

It has previously been shown that pre-trial activity in higher cortical areas can be analyzed as a neural predictor of subsequent memory (Turk-Browne, Yi, & Chun, 2006), perceptual decisions (Andrews, Schluppeck, Homfray, Mathews, & Blakemore, 2002; Hesselmann, Kell, Eger, & Kleinschmidt, 2008), motor decisions (Soon, Brass, Heinze, & Haynes, 2008), and moment-to-moment fluctuations in cognitive flexibility (Leber, Turk-Browne, & Chun, 2008). Our results show that the initial dominant percept during an episode of binocular rivalry is determined or biased by pre-trial activity in visual cortex. We speculate that the natural fluctuations of endogenous neural activity in visual cortex may play a pivotal role in determining the stochastic perceptual alternation that is inherent to the phenomenon of interocular suppression. By investigating not just the neural correlates of states of perceptual awareness, but their neural antecedents, we are moving closer to an understanding of the causal mechanisms underlying binocular rivalry and ultimately awareness.

References

Alais, D. & Blake, R. (1999). Grouping visual features during binocular rivalry. *Vision Research*, *39*, 4341–4353.

Alais, D. & Melcher, D. (2007). Strength and coherence of binocular rivalry depends on shared stimulus complexity. *Vision Research*, *47*, 2690–3279.

Andrews, T. J., Schluppeck, D., Homfray, D., Mathews, P. & Blakemore, C. (2002). Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. *NeuroImage*, *17*, 890–901.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, *96*, 145–167.

Blake, R. & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13–21.

Blake, R., Westendorf, D. H. & Overton, R. (1981). What is suppressed during binocular rivalry? *Perception*, *9*, 223–231.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Ryu, S. I. & Shenoy, K. V. (2010). Cortical preparatory activity: Representation of movement or first cog in dynamical machine? *Neuron*, *68*, 387–400.

Crowe, D. A., Averbeck, B. B. & Chafee, M. V. (2010). Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *Journal of Neuroscience*, *30*, 11640–11653 (2010)

Dayan, P. (1998). A hierarchical model of binocular rivalry. *Neural Computation*, *10*, 1119–1135.

Diaz-Caneja, E. (1928). Sur l'alternance binoculaire. *Annals Oculist*, (October), 721–731.

Dörrenhaus, W. (1975). Musterspezifischer visueller wettstreit. *Naturwissenschaften*, *62*, 578–579.

Donner, T. H., Sagi, D., Bonneh, Y. S. & Heeger, D. J. (2008). Opposite neural signatures of motion-induced blindness in human dorsal and ventral visual cortex. *Journal of Neuroscience*, *28*(41), 10298–10310.

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J. & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, *3*, 586–598.

Epstein, R. & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.

Fox, R. & Check, R. (1968). Detection of motion during binocular rivalry suppression. *Journal of Experimental Psychology*, *78*, 388–395.

Fox, R. & Check, R. (1972). Independence between binocular rivalry suppression and duration and magnitude of suppression. *Journal of Experimental Psychology*, *93*, 283–289.

Freeman, A. W. (2005). Multistage model for binocular rivalry. *Journal of Neurophysiology*, *94*, 4412–4420.

Fukuda, H. (1981). Magnitude of suppression of binocular rivalry within the invisible pattern. *Perceptual and Motor Skills*, *53*, 371–375.

Harrison, S. A. & Tong, F. (2009). Decoding reveals the content of visual working memory in early visual areas. *Nature*, *458*, 632–635.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L. & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.

Haynes, J. D. & Rees, G. (2005). Predicting the stream of consciousness from activity in human visual cortex. *Current Biology*, *15*, 1301–1307.

Haynes, J. D. & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*, 523–534.

Haynes, J. D. (2011). Decoding and predicting intentions. *Annals of the New York Academy of Sciences*, *1224*, 9–21.

Heeger, D. J. & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, *3*, 142–151.

Heekeren, H. R., Marrett, S. & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*, 467–479.

Hesselmann, G., Kell, C. A., Eger, E. & Kleinschmidt, A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *PNAS*, *105*, 10984–10989.

Hsieh, P. J., Caplovitz, G. P. & Tse, P. U. (2006). Bistable illusory rebound motion: Event-related functional magnetic resonance imaging of perceptual states and switches. *NeuroImage*, *32*(2), 728–739.

Hsieh, P. J. & Tse, P. U. (2009). Microsaccade rate varies with subjective visibility during motion-induced blindness. *PLoS ONE*, *4*, e5163.

Hsieh, P. J. & Tse, P. U. (2010). 'Brain-reading' of perceived colors reveals a feature mixing mechanism underlying perceptual filling-in in cortical area V1. *Human Brain Mapping*, *31*, 1395–1407.

Hsieh, P. J., Vul, E. & Kanwisher, N. (2010). Recognition alters the spatial pattern of fMRI activation in early retinotopic cortex. *Journal of Neurophysiology*, *103*, 1501–1507.

Kamitani, Y. & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*, 679–685.
Kanwisher, N., McDermott, J. & Chun, M. C. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.
Laing, C. R. & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience*, *12*, 39–53.
Lankheet, M. J. M. (2006). Unraveling adaptation and mutual inhibition in perceptual rivalry. *Journal of Vision*, *6*, 304–310.
Leber, A. B., Turk-Browne, N. B. & Chun, M. M. (2008). Neural predictors of moment-to-moment fluctuations in cognitive flexibility. *PNAS*, *105*, 13592–13597.
Lee, S. H., Blake, R. & Heeger, D. J. (2005). Traveling waves of activity in primary visual cortex during binocular rivalry. *Nature Neuroscience*, *8*, 22–23.
Lehky, S. R. (1988). An astable multivibrator model of binocular rivalry. *Perception*, *17*, 215–228.
Leopold, D. A. & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*, 549–553.
Leopold, D. A., Wilke, M., Maier, A. & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*, 605–609.
Logothetis, N. K., Leopold, D. A. & Sheinberg, D. L. (1996). What is rivaling during binocular rivalry? *Nature*, *380*, 621–624.
Leopold, D. A., Murayama, Y. & Logothetis, N. K. (2003). Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging. *Cerebral Cortex*, *13*, 422–433.
Logothetis, N. K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London B*, *353*, 1801–1818.
Lumer, E. D., Friston, K. J. & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*, *280*, 1930–1934.
Meng, M., Remus, D. A. & Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nature Neuroscience*, *8*, 1248–1254.
Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K. & Poggio, T. (2008). Dynamic population coding of category information in inferior temporal and prefrontal cortex. *Journal of Neurophysiology*, *100*, 1407–1419.
Mitchell, J. F., Stoner, G. R. & Reynolds, J. H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature*, *429*, 410–413.
Moutoussis, K., Keliris, G., Kourtzi, Z. & Logothetis, N. (2005). A binocular rivalry study of motion perception in the human brain. *Vision Research*, *45*, 2231–2243.
Mueller, T. J. (1990). A physiological model of binocular rivalry. *Visual Neuroscience*, *4*, 63–73.
O'Shea, R. P. & Crassini, B. (1981). The sensitivity of binocular rivalry suppression to changes in orientation assessed by reaction-time and forced-choice techniques. *Perception*, *10*, 283–293.
O'Shea, R. P. (1987). Chronometric analysis supports fusion rather than suppression theory of binocular vision. *Vision Research*, *27*, 781–791.
Op de Beeck, H. P. (2010). Probing the mysterious underpinning of multi-voxel fMRI analyses. *NeuroImage*, *50*(2), 567–571.
Pearson, J. & Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Science*, *12*, 334–341.
Pearson, J., Clifford, C. W. & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, *18*, 982–986.
Polonsky, A., Blake, R., Braun, J. & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, *3*, 1153–1159.
Schoth, F., Waberski, T. D., Krings, T., Gobbele, R. & Buchner, H. (2007). Cerebral processing of spontaneous reversals of the rotating Necker cube. *Neuroreport*, *18*(13), 1335–1338.
Smith, E. L., Levi, D. M., Harwerth, R. S. & White, J. M. (1982). Color vision is altered during the suppression phase of binocular rivalry. *Science*, *218*, 802–804.
Soon, C. S., Brass, M., Heinze, H. J. & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543–545.
Sterzer, P. & Rees, G. (2008). A neural basis for perceptual stabilization in binocular rivalry. *Journal of Cognitive Neuroscience*, *20*, 389–399.
Sugie, N. (1982). Neural models of brightness perception and retinal rivalry in binocular vision. *Biological Cybernetics*, *43*, 13–21.
Tong, F. (2001). Competing theories of binocular rivalry: A possible resolution. *Brain and Mind*, *2*, 55–83.
Tong, F. & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, *411*, 195–199.
Tong, F., Meng, M. & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, *10*, 502–511.
Tong, F., Nakayama, K., Vaughan, J. T. & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, *21*, 753–759.
Turk-Browne, N. B., Yi, D. J. & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron*, *49*(6), 917–927.
Wales, R. & Fox, R. (1970). Increment detection thresholds during binocular rivalry suppression. *Perception and Psychophysics*, *8*, 90–94.
William, M. A., Baker, C. I., Op de Beeck, H. P., Shim, W. M., Dang, S., Triantafyllou, C. & Kanwisher, N. (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*, *11*, 1439–1445.
Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F. & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *Journal of Neuroscience*, *24*(12), 2898–2904.
Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 14499–14503.
Wolfe, J. M. (1983). Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli. *Perception*, *12*, 447–456.
Wunderlich, K., Schneider, K. & Kastner, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate neucleus. *Nature Neuroscience*, *8*, 1595–1602.
Yu, K. & Blake, R. (1992). Do recognizable figures enjoy an advantage in binocular rivalry? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1158–1173.