Behavioral/Systems/Cognitive

Cortical Dynamics Underlying Face Completion in Human Visual System

Juan Chen,[1] **Tiangang Zhou,**[2] **Hua Yang,**[3] **and Fang Fang**[1]

[1]Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, [2]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, and [3]YuanPei College, Peking University, Beijing 100871, People's Republic of China

In natural images, visual objects are typically occluded by other objects. A remarkable ability of our visual system is to complete occluded objects effortlessly and see whole, uninterrupted objects. How object completion is implemented in the visual system is still largely unknown. In this study, using a backward masking paradigm, we combined psychophysics and functional magnetic resonance imaging to investigate the temporal evolvement of face completion at different levels of the visual processing hierarchy. Human subjects were presented with two kinds of stimuli that were designed to elicit or not elicit the percept of a completed face, although they were physically very similar. By contrasting subjects' behavioral and blood oxygenation level-dependent (BOLD) responses to completed and noncompleted faces, we measured the psychophysical time course of the face completion and its underlying cortical dynamics. We found that face completion manifested its effect between 50 and 250 ms after stimulus onset. Relative to noncompleted faces, completed faces induced weaker BOLD response at early processing phases in retinotopic visual areas V1 and V2 and stronger BOLD response at late processing phases in occipital face area and fusiform face area. Attending away from the stimuli largely abolished these effects. These findings suggest that face completion consists of two synergetic phases: early suppression in lower visual areas and late enhancement in higher visual areas; moreover, attention is necessary to these neural events.

Introduction

Great strides have been made in understanding the neural mechanisms of visual object recognition (Grill-Spector and Malach, 2004; Kanwisher, 2010). So far, object recognition has been studied mainly using individual objects presented alone. However, visual objects rarely occur in isolation in natural scenes. It is common for one object to occlude another object in natural images. A striking ability of human vision is the recognition of objects even when the sensory information specifying objects is optically incomplete due to occlusion. We have little difficulty completing occluded objects and seeing whole, uninterrupted objects.

How object completion is implemented in the visual cortex remains elusive. Evidence from human brain imaging studies suggest that (only) high-level visual areas selective for objects are likely candidates to mediate the operation of object completion effects (Doniger et al., 2000; Lerner et al., 2002; Stanley and Rubin, 2003; Hegdé et al., 2008; Grützner et al., 2010). However, psychophysical and electrophysiological studies suggest that early

visual cortical areas also play an important functional role in object completion (Nakayama et al., 1989; Sugita, 1999; Bakin et al., 2000; Pillow and Rubin, 2002). The discrepancy could be due to sluggish temporal response of the functional magnetic resonance imaging (fMRI) method. In the fMRI studies by Lerner et al. (2002) and Hegdé et al. (2008), occluded objects were presented for hundreds of milliseconds. This completion-related activation in object-selective areas as detected by fMRI might have reflected the perceptual consequence of object completion, rather than the process of object completion.

Psychophysical and event-related potential (ERP) studies have shown that object completion is not instantaneous; instead, it manifests its effect within a temporal window shortly after stimulus onset (Sekuler and Palmer, 1992; Murray et al., 2001; Johnson and Olshausen, 2005; Chen et al., 2009). In this study, we attempted to investigate the temporal evolvement of face completion at different levels of the visual hierarchy. To circumvent the low temporal resolution deficit of the fMRI method, we used a backward masking paradigm to present occluded faces with various durations, which rendered it possible to interrupt the visual processing of occluded faces and follow in some detail the temporal evolvement of face completion (Grill-Spector et al., 2000; Bar et al., 2001; Lamme et al., 2002).

Visual stimuli were constructed by presenting identical face fragments stereoscopically either behind or in front of a textured occluder (Fig. 1). In the first condition, the stimuli were perceptually completed and organized into a coherent face. In the second condition, they were perceived as disjoint face fragments hovering above the textured occluder (Nakayama et al., 1989;

A

L5   R5   Occluder

B

Face behind occluder (FBO)   Face in front of occluder (FIO)   Face only (FO)

C

Face behind occluder (FBO)   Face in front of occluder (FIO)
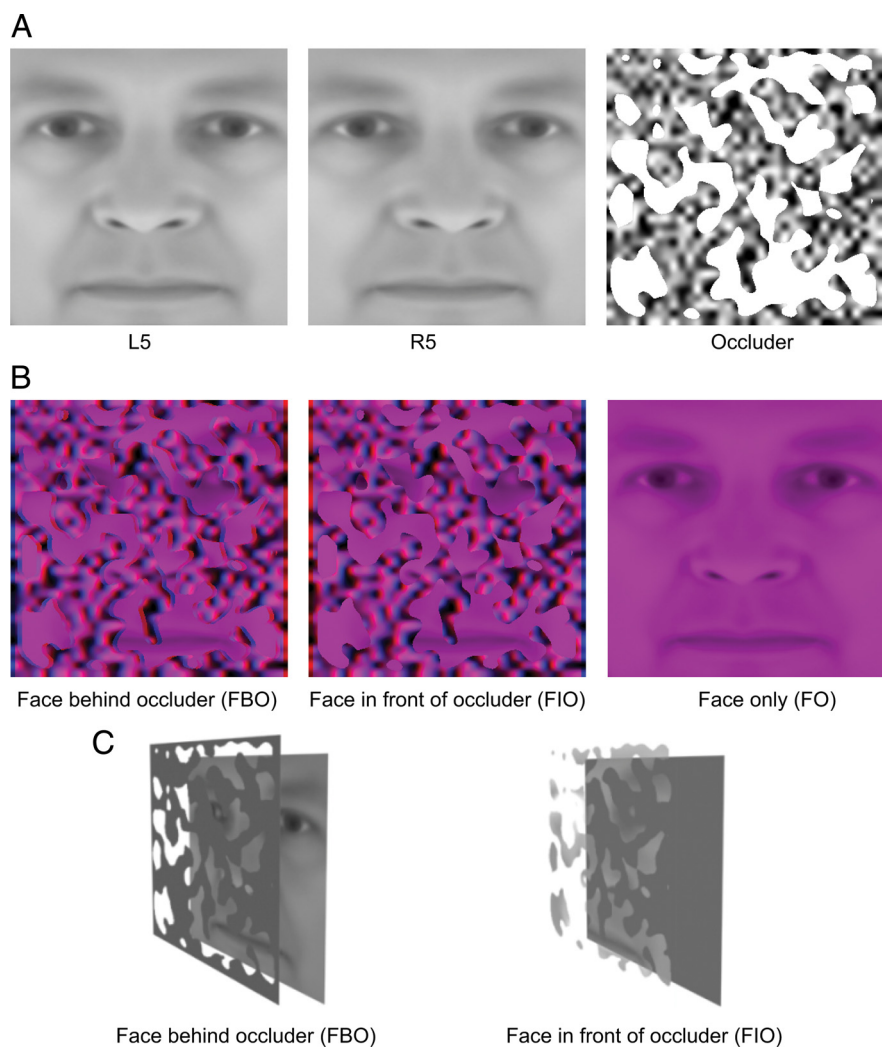
Figure 1. Stimuli. **A**, Sample images used to construct stimuli in the psychophysical and fMRI experiments. The left and the middle images are 5° side views of a face, which tilt to left or right, respectively. The right image is a textured occluder with irregular holes. **B**, Sample stimuli in the psychophysical and fMRI experiments. They were generated by masking a face image with a textured occluder and were presented stereoscopically by using red/blue anaglyphic glasses. Disparity information specified that the face image could be either behind (FBO, left) or in front of (FIO, middle) the occluder. An isolated purple face (FO, right) was also presented in the experiments. **C**, Perceptual effects of FBO and FIO stimuli. FBO stimulus is perceptually completed and organized into a coherent face (left), while FIO stimulus is perceived as disjoint fragments floating over a textured plane (right).

angle and were presented against a gray background. Occluded faces were generated by masking a 5° face side view with a textured occluder (Fig. 1A) and were presented stereoscopically by using red/blue anaglyphic glasses. Approximately 35% of the face area was exposed to subjects through the holes of the occluder. Disparity information specified that the occluder could be either in front of or behind the face image (or face fragments) (Fig. 1B). Face fragments were always at zero disparity. The occluder was at either +0.18° or −0.18° disparity. When the face fragments were stereoscopically presented behind the textured occluder [face behind occluder (FBO) condition], they were perceptually completed and organized into a coherent face by observers (Fig. 1C, left). However, when the same fragments were presented stereoscopically in front of the textured occluder [face in front of occluder (FIO) condition], they were perceived as disjoint fragments floating over the textured plane (Fig. 1C, right). The FBO and FIO stimuli were identical in two dimensions, the key difference is the face recognition advantage generated by the stereoscopic occlusion for FBO stimuli (Nakayama et al., 1989). In all these three experiments, complete faces without occlusion were also used [face only (FO) condition] and they were presented in purple (with only red and blue channels on) to match the color of the FIO and FBO stimuli (Fig. 1B, right).

The 5° side view of a face was generated by projecting a three-dimensional (3D) face model with a 5° in-depth rotation angle onto the monitor plane with the front view as the initial position. Both left and right rotations were executed. The 3D face models were generated by FaceGen Modeller 3.1, and a total of 40 models were used in this study. We generated 80 occluders, each of which had holes with random shapes and positions. Any combination of face models and occluders was used for both the FBO and FIO stimuli.

In the third fMRI experiment, only the textured occluder was presented stereoscopically at a near (near occluder condition) or far (far occluder condition) depth. Their disparities (±0.18°) were the same as those in other experiments. No face fragments were presented.

In all the experiments, backward masks were used to control stimulus duration. The masks were generated by convolving a random noise pattern (pixel size = 0.23° × 0.23°) with a two-dimensional Gaussian function ($\sigma = 0.23°$). They also subtended 10.6° × 10.6° of visual angle.

Designs. In the psychophysical experiment, we adopted the performance-based measure developed by Murray et al. (2001) to unfold the psychophysical time course of face completion. FIO, FBO, and FO stimuli were presented on an Iiyama HM204DT 22 inch monitor, with a spatial resolution of 1024 × 768 and a refresh rate of 75 Hz. The viewing distance was 74 cm. Their head position was stabilized using a chin rest and a head rest. Throughout the experiment, subjects were asked to fixate a small dot presented at zero disparity and at the center of the monitor. Each trial started with a 1000 ms blank interval. Then a face stimulus (FIO, FBO, or FO) was presented at the center of the monitor with duration of 50, 150, 250, 350, or 450 ms, followed by a 300 ms mask. Subjects pressed one of the two response keys to indicate the view direction of the face stimulus, either left or right (Fig. 2A).

There were 15 experimental conditions in the psychophysical experiment: five durations (50, 150, 250, 350, and 450 ms) × three stimulus

Fang and He, 2005). In the psychophysical experiment and the first fMRI experiment, by contrasting subjects' behavioral and blood oxygenation level-dependent (BOLD) responses in these two conditions, we measured the psychophysical time course of the face completion and its underlying cortical dynamics. In the second fMRI experiment, we investigated the role of attention in the face completion. The third fMRI experiment was performed to rule out alternative explanations to the data in the first experiment.

Materials and Methods

Subjects. Eight human subjects (6 female and 2 male) participated in the psychophysical and the first two fMRI experiments. Four of them (3 female and 1 male) participated in the third fMRI experiment. All of them were right-handed, reported normal or corrected-to-normal vision, and had no known neurological or visual disorders. Ages ranged from 19 to 33 years. They gave written, informed consent in accordance with the procedures and protocols approved by the human subjects review committee of Peking University.

Stimuli. Face stimuli in the psychophysical and the first two fMRI experiments were identical, which subtended 10.6° × 10.6° of visual

types (FIO, FBO, and FO). The experiment consisted of eight sessions, and a session consisted of five blocks of 60 trials, one block for a duration condition. In each block, there were 20 trials for each of the three stimulus types and the stimulus duration was fixed. Both the order of the five blocks in a session and the order of the trials in a block were randomized. All data from the eight sessions were pooled together for analysis.

The first fMRI experiment was conducted to investigate how face completion evolved at different levels of the visual hierarchy. A block design was adopted. There are 12 experimental conditions: four durations (50, 150, 250, and 350 ms) × three stimulus types (FIO, FBO, and FO). The experiment consisted of sixteen 360 s functional scans. Each scan consisted of twelve 12 s stimulus blocks (one for each condition) interleaved with twelve 18 s blank intervals. The order of the experimental conditions in a scan was randomized. A fixation point was presented at zero disparity and at the center of the monitor. The fixation point became dimmer during the last two seconds of a blank interval to signal an upcoming stimulus block. A stimulus block contained six 2 s trials. In a trial, a face stimulus (FIO, FBO, or FO) was presented at the center of the gray screen for a fixed duration (50, 150, 250, or 350 ms), followed by a 300 ms mask and then by a blank screen. Subjects were asked to attend to the stimulus and press one of the two response keys to indicate the view direction of the face stimulus, either left or right.

The second fMRI experiment was conducted to investigate the role of attention in face completion in the visual cortex. Its design was identical to that of the first fMRI experiment except that subjects performed a highly attention-demanding rapid serial visual presentation (RSVP) task at fixation in stimulus blocks, rather than judged the face view direction. In a stimulus block, this attention task required subjects to count the number of targets (Xs) in a stream of rapidly presented distractor letters (Z, L, N, and T). Each letter subtended 0.27° of visual angle and was presented for 150 ms. Subjects needed to report the number of targets observed at the end of each stimulus block by pressing one of four response keys corresponding to the number of target Xs presented (1–4).

In the third fMRI experiment, eight experimental conditions were included: four durations (50, 150, 250, and 350 ms) × two stimulus types (near occluder and far occluder). The design was similar to the first two fMRI experiments. The experiment consisted of twelve 240 s functional scans. Each scan consisted of eight 12 s stimulus blocks (one for each condition) interleaved with eight 18 s blank intervals. The order of the experimental conditions in a scan was randomized. A fixation point was presented at zero disparity and at the center of the monitor. The fixation point became dimmer during the last two seconds of a blank interval to signal an upcoming stimulus block. A stimulus block contained six 2 s trials. In a trial, a stimulus (near occluder or far occluder) was presented for a fixed duration (50, 150, 250, or 350 ms), followed by a 300 ms mask and then by a blank screen. The position of the stimulus was shift to the left or right of the fixation point by 0.18°. Subjects were asked to attend to the stimulus and press one of the two response keys to indicate the shift direction of the stimulus, either left or right.

Retinotopic visual areas (V1, V2, and V3) were defined by a standard phase-encoded method developed by Sereno et al. (1995) and Engel et al. (1997), in which subjects viewed rotating wedge and expanding ring stimuli that created traveling waves of neural activity in visual cortex. A block-design scan was used to define the regions of interest (ROIs), including face-selective areas and responsive areas in V1, V2, and V3. Subjects viewed images of faces, non-face objects, and texture patterns (scrambled faces), which had the same size as the stimuli used in our main experiments and were presented at the center of the screen. Images appeared at a rate of 2 Hz in blocks of 12 s, interleaved with 12 s blank blocks. Each image was presented for 300 ms, followed by a 200 ms blank

A Experimental procedure

Blank
1000ms

Stimulus
50-450ms

Mask
300ms

Wait for response
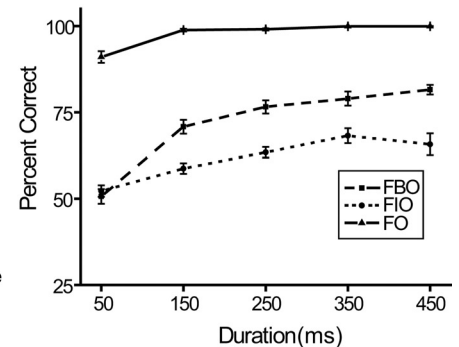
Time
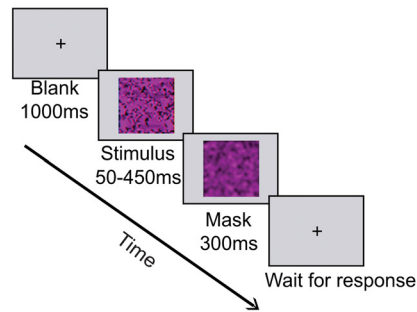
B Result

Figure 2. Design and psychophysical result. *A*, Schematic description of a trial in psychophysical and fMRI experiments. A face stimulus (FIO, FBO, or FO) was presented with a duration of 50, 150, 250, 350, and 450 ms, followed by a 300 ms mask. Subjects were asked to judge the view direction of the face stimulus, either left or right. *B*, Psychophysical result. The performance of view direction judgment was plotted as a function of stimulus duration for FIO, FBO, and FO stimuli. Error bars denote 1 SEM calculated across subjects.

interval. Each block type was repeated 5 times in the scan, which lasted 360 s. Subjects performed a one-back task during scanning.

MRI data acquisition. In the scanner, the stimuli were back-projected via a video projector (refresh rate, 60 Hz; spatial resolution, 1024 × 768) onto a translucent screen placed inside the scanner bore. Subjects viewed the stimuli through a mirror located above their eyes. The viewing distance was 83 cm. Functional MRI data were collected using a 3T Siemens Trio scanner with a 12-channel phase-array coil. BOLD signals were measured with an echoplanar imaging sequence (echo time, 30 ms; repetition time, 2000 ms; field of view, 196 × 196 mm^2; matrix, 64 × 64; flip angle, 90; slice thickness, 3 mm; gap, 0 mm; number of slices, 33; slice orientation, axial). The bottom slice was positioned at the bottom of the temporal lobes. A high-resolution 3D structural dataset (3D magnetization-prepared rapid-acquisition gradient echo; 1 × 1 × 1 mm^3 resolution) was collected in the same session before the functional runs. All the subjects underwent five sessions, one for retinotopic mapping and localizing face-selective areas, two for the first experiment and two for the second experiment. Four of the subjects underwent an extra session for the third experiment.

MRI data processing and analysis. The anatomical volume for each subject in the retinotopic mapping session was transformed into the AC-PC (anterior commissure–posterior commissure) space and then inflated using BrainVoyager QX. Functional volumes in all the sessions for each subject were preprocessed, including 3D motion correction, linear trend removal, and high-pass (0.015 Hz) (Smith et al., 1999) filtering using BrainVoyager QX. Head motion within any fMRI session was <2 mm for all subjects. The images were then aligned to the anatomical volume in the retinotopic mapping session and transformed into the AC-PC space. The first 6 s of BOLD signals were discarded to minimize transient magnetic saturation effects.

A general linear model (GLM) procedure was used for ROI analysis. The ROIs in V1, V2, and V3 were defined as areas that responded more strongly to the textured patterns (scrambled faces) than blank screen ($p < 10^{-8}$, uncorrected) and confined by the V1/V2/V3 boundaries defined by the retinotopic mapping scan. Face-selective areas were defined as areas that responded more strongly to faces than non-face objects ($p < 10^{-4}$, uncorrected). Five face-selective areas [with their Talairach coordinates (x, y, z)] were found in all subjects [right fusiform face area (rFFA): 36 ± 1, −46 ± 1, −15 ± 1; right occipital face area (rOFA): 37 ± 1, −71 ± 1, −7 ± 12; left occipital face area (lOFA): −37 ± 2, −69 ± 3, −7 ± 13; right superior temporal sulcus (rSTS): 47 ± 1, −50 ± 2, 10 ± 2; lSTS: −46 ± 2, −53 ± 2, 7 ± 3], while lFFA (−41 ± 1, −45 ± 1, −17 ± 1) was found in 7 (of 8) subjects, according to the above criterion.

The BOLD signals induced by the stimulus blocks were calculated separately for each ROI and each subject. For each fMRI run, the time course of fMRI signal intensity was first extracted by averaging the data across all the voxels within the predefined ROI and then normalized by the average of the last two time points of all 18 s blank intervals in that

run. The peak response in an ROI was extracted by averaging the response within a 7–12 s interval after the start of the stimulus block and then averaged according to different experimental conditions.

Results

Psychophysical results

Subjects' performance of view direction judgment was plotted as a function of stimulus duration for the FIO, FBO, and FO stimuli, respectively (Fig. 2B). For the FO stimulus, subjects had no difficulty judging the view direction of a face at all durations. Even with only 50 ms exposure, their performance could reach 92%. For the FIO and FBO, stimuli, subjects' performance improved as the stimulus duration increased, but their overall performance significantly dropped down, compared with the FO stimulus. A repeated-measures ANOVA of percentage correct was performed with stimulus type and duration as within-subject factors. Both the main effects of stimulus type ($F_{(2,14)} = 357.01, p < 0.001$) and duration ($F_{(4,28)} = 81.84, p < 0.001$) were significant, which were consistent with our observation.

To reveal the time course of face completion, we took a close look at the performance in the FIO and FBO conditions and their difference. The performance in the FBO condition, compared with the FIO condition, can be taken as a measure of face completion. When the performance in the FBO condition is better than that in the FIO condition, we attribute this to face completion. When the performance in the FBO condition is no better than that in the FIO condition, we take this to mean that face completion has not occurred. The extent of face completion as a function of stimulus duration was measured and defined as the time course of face completion (Murray et al., 2001). At 50 ms duration, there was no significant difference between the FIO and FBO stimuli ($t_{(7)} = 0.83, p = 0.44$). At longer durations, subjects performed significantly better for the FBO stimuli than for the FIO stimuli (150 ms: $t_{(7)} = 5.91, p < 0.001$; 250 ms: $t_{(7)} = 18.49, p < 0.001$; 350 ms: $t_{(7)} = 6.73, p < 0.001$; 450 ms: $t_{(7)} = 5.43, p < 0.001$). In other words, the performance functions for the FIO and FBO stimuli diverged after 50 ms, which suggested that the face completion started to manifest its effect after 50 ms.

To investigate when face completion terminated, we run multiple paired t tests to compare the performance at different duration conditions for the FBO stimuli. Significant performance difference was observed between 50 ms and 150 ms conditions ($t_{(7)} = 11.65, p < 0.001$) and between 150 ms and 250 ms conditions ($t_{(7)} = 4.51, p < 0.01$), but not between 250 ms and 350 ms conditions ($t_{(7)} = 1.20, p = 0.27$) or between 350 ms and 450 ms conditions ($t_{(7)} = 1.82, p = 0.11$). These results show that subjects' performance with the FBO stimuli saturated at 250 ms and suggest that face completion terminated before 250 ms. Overall psychophysical data suggested that face completion took effect between 50 and 250 ms after stimulus onset.

fMRI results

The first fMRI experiment was designed to investigate how face completion evolved in the visual processing hierarchy. In other words, we attempted to reveal how low-level (V1, V2, and V3) and high-level (OFA, STS, and FFA) visual areas responded during the process of face completion. Since there was no qualitative difference in the fMRI data between the two hemispheres, we collapsed the data from the two hemispheres for further analyses. BOLD responses in lSTS and rSTS were very weak (<0.1% signal change) to both FIO and FBO stimuli at all durations. They were not included in this study.

BOLD responses in V1, V2, V3, OFA, and FFA were plotted as a function of stimulus duration for the FIO, FBO, and FO stimuli,
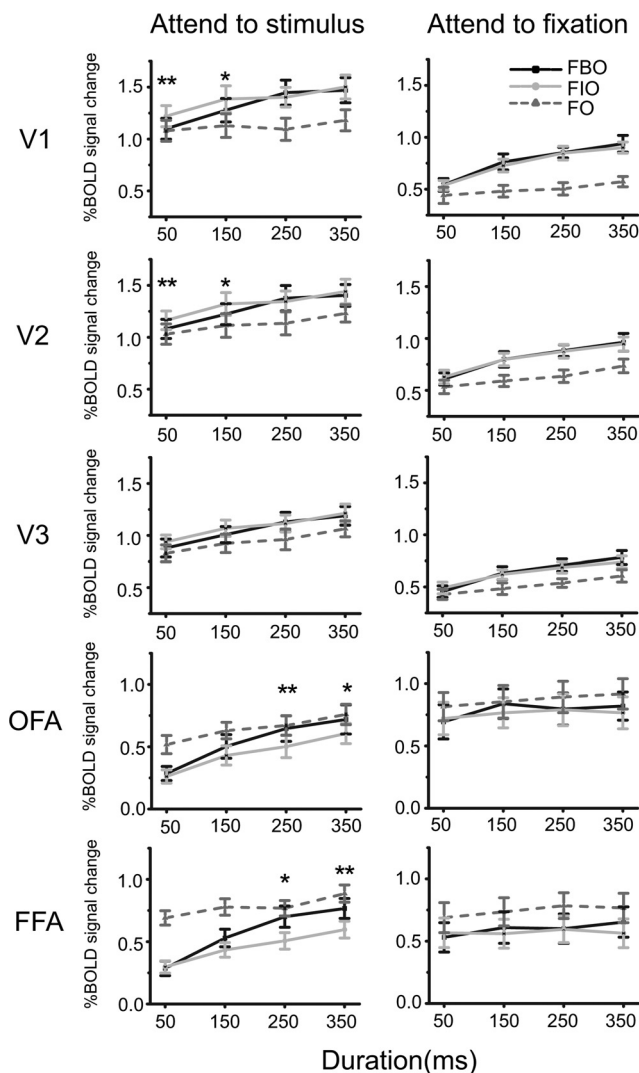

Figure 3. Results in the first (left column) and the second (right column) fMRI experiments. For each cortical area, BOLD responses to FIO, FBO, and FO stimuli were plotted as a function of stimulus duration. Asterisks indicate a statistically significant difference between BOLD responses to FBO and FIO stimuli (*$p < 0.05$, **$p < 0.01$). Error bars denote 1 SEM calculated across subjects for each condition.

respectively (Fig. 3, left column). Statistical analyses focused on the comparison between the FIO and the FBO conditions. For each area, a repeated-measures ANOVA of BOLD response was performed with stimulus type (FBO and FIO) and duration (50, 150, 250, and 350 ms) as within-subject factors. The main effect of duration was significant in all the areas (V1: $F_{(3,21)} = 25.04, p < 0.001$; V2: $F_{(3,21)} = 38.39, p < 0.001$; V3: $F_{(3,21)} = 41.52, p < 0.001$; OFA: $F_{(3,21)} = 36.46, p < 0.001$; FFA: $F_{(3,21)} = 62.45 p < 0.001$). BOLD responses to the FIO and FBO stimuli generally increased with stimulus duration. A significant increase was found in the following comparisons: 50 ms vs 150 ms for FBO in V1; 150 ms vs 250 ms for FBO in V1; 50 ms vs 150 ms for FIO in V1; 50 ms vs 150 ms for FBO in V2; 150 ms vs 250 ms for FBO in V2; 50 ms vs 150 ms for FIO in V2; 250 ms vs 350 ms for FIO in V2; 50 ms vs 150 ms for FBO in V3; 150 ms vs 250 ms for FBO in V3; 50 ms vs 150 ms for FIO in V3; 250 ms vs 350 ms for FIO in V3; 50 ms vs 150 ms for FBO in OFA; 150 ms vs 250 ms for FBO in OFA; 250 ms vs 350 ms for FBO in OFA; 50 ms vs 150 ms for FIO in OFA; 150 ms vs 250 ms for FIO in OFA; 250 ms vs 350 ms for FIO in OFA; 50 ms vs 150 ms for FBO in FFA; 150 ms vs 250
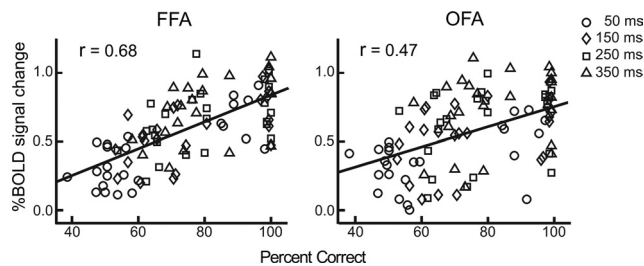
Figure 4. Correlations between psychophysical data and BOLD signals at OFA and FFA. Data from all eight subjects in all 12 experimental conditions were pooled together for computing correlation coefficients. Four durations are marked with different symbols.

Figure 5. Results in the third fMRI experiment. BOLD responses to near and far occluders in V1 and V2 were plotted as a function of stimulus duration. Error bars denote 1 SEM calculated across subjects for each condition.

ms for FBO in FFA; 50 ms vs 150 ms for FIO in FFA; 150 ms vs 250 ms for FIO in FFA; 250 ms vs 350 ms for FIO in FFA (all $t_{(7)} >$ 2.51, $p < 0.05$).

The main effect of stimulus type was significant in V1, V2, OFA, and FFA (V1: $F_{(1,7)} = 5.09$, $p = 0.05$; V2: $F_{(1,7)} = 6.63$, $p < 0.05$; OFA: $F_{(1,7)} = 11.26$, $p < 0.05$; FFA: $F_{(1,7)} = 8.28$, $p < 0.05$), but not in V3 ($F_{(1,7)} = 3.074$, $p = 0.12$). *Post hoc* analyses showed that V1 and V2 responded stronger to the FIO stimuli than to the FBO stimuli at short durations (50 and 150 ms) (V1: 50 ms: $t_{(7)} = 4.34$, $p < 0.01$; 150 ms: $t_{(7)} = 2.36$, $p < 0.05$; V2: 50 ms: $t_{(7)} = 4.11$, $p < 0.01$; 150 ms: $t_{(7)} = 2.75$, $p < 0.05$), while OFA and FFA responded stronger to the FBO stimuli than to the FIO stimuli at long durations (250 and 350 ms) (OFA: 250 ms: $t_{(7)} = 4.01$, $p < 0.01$; 350 ms: $t_{(7)} = 3.05$, $p < 0.05$; FFA: 250 ms: $t_{(7)} = 3.46$, $p < 0.05$; 350 ms: $t_{(7)} = 4.18$, $p < 0.01$). No significant response difference was found in other conditions. These findings suggest that both low- and high-level visual areas were involved in the process of face completion, but they responded at different temporal phases with opposite activation patterns. When the FIO stimuli were presented longer than 50 ms, subjects' performance was significantly above chance level (all $t_{(7)} > 5.73$, $p < 0.001$). Note that subjects' behavioral data in the magnet could replicate their psychophysical results described above.

To examine the link between psychophysical data (Fig. 2*B*) and fMRI data (Fig. 3, left column), for each cortical area, we pooled all subjects' data in all the 12 conditions and calculated the correlation coefficient between their psychophysical and fMRI data. The correlation was significant at OFA ($n = 96$, $r = 0.47$, $p < 0.01$) and FFA ($n = 96$, $r = 0.68$, $p < 0.01$), and the correlation difference between OFA and FFA was significant ($z = 2.17$, $p < 0.05$) (Fig. 4). Since the analysis above used data across both durations and subjects, it is not clear therefore whether the correlation reflects differences across subjects or across durations or both. To address this issue, we performed an additional analysis. For each of eight subjects, we calculated correlation coefficients across all 12 experimental conditions between psychophysical data and BOLD signals in FFA and OFA, respectively. Thus, we got eight coefficients for FFA and OFA, respectively. We then transformed all these 16 coefficients to Fisher Z scores so that they followed a normal distribution and could be compared with a t test (Fischer and Whitney, 2009). Paired t test showed that the correlation in FFA is significantly higher than that in OFA ($t_{(7)} = 2.74$, $p = 0.014$), which suggests that FFA activity is more correlated with psychophysical data than OFA activity. Note that each point in Figure 4 is not independent of the other because the effect of face completion accumulated over time, it is possible that the correlations might not be so significant as calculated.

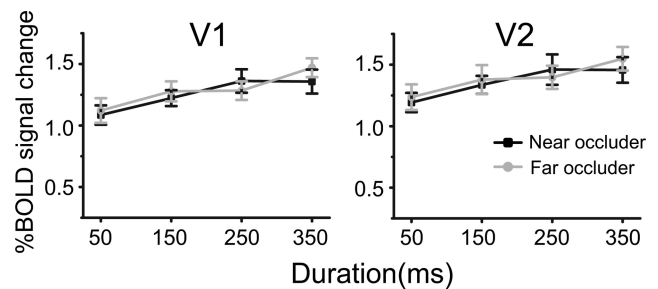In the second fMRI experiment, subjects' attention was directed to a very demanding RSVP task at fixation, instead of the face stimuli. To investigate the role of attention in face completion, we performed similar repeated-measures ANOVAs as those in the first fMRI experiment. This attentional manipulation largely abolished differential responses to the FIO and the FBO stimuli as observed in the first fMRI experiment (Fig. 3, right column). The main effect of stimulus type was not significant in any of cortical areas (all $F_{(1,7)} < 4.62$ and $p > 0.07$), which indicates that attention was necessary to the cortical dynamics underlying face completion.

The third fMRI experiment was performed to examine whether the differential responses to the FIO and FBO stimuli in V1 and V2 was due to the absolute disparity difference between the near and far occluders. We run paired t tests to compare the BOLD responses to the FIO and the FBO stimuli. In both V1 and V2, there was no significant difference for all stimulus durations (all $t < 1.28$ and $p > 0.29$), which rules out disparity difference as an alternative explanation for the first fMRI experiment (Fig. 5).

Finally, to examine whether the activation pattern observed in the first fMRI experiment can be generalized to another task, we performed an additional experiment and collected data from five subjects. In this experiment, we used a new set of face images and occluders. Occluders were generated in the same way as before. Face images were generated by rotating front face views with a 3° in-plane angle (left or right). Then, FBO and FIO stimuli were constructed similarly as before and were presented with durations of 50, 150, 250, and 350 ms. Subjects were asked to judge the in-plane orientations of the faces (left or right tilted). Note that, in the first fMRI experiment, subjects were asked to report the view directions of faces (i.e., in-depth orientation). We found that, with the new stimuli and the new task, V1 and V2 responded stronger to the FIO stimuli than to the FBO stimuli at short durations (50 and 150 ms) (all $t > 3.24$, $p < 0.05$), while OFA and FFA responded stronger to the FBO stimuli than to the FIO stimuli at long durations (250 and 350 ms) (all $t > 4.45$, $p < 0.05$), which replicated the basic pattern found in the first fMRI experiment.

Discussion

Using a backward masking paradigm, we examined the cortical dynamics underlying face completion with fMRI and psychophysics. Our data provide evidence that both early visual cortical areas (V1 and V2) and high-level face-selective areas (OFA and FFA) were involved in face completion, but they responded at different temporal phases with opposite activation patterns. We also show that attention was necessary to these neural events.

The psychophysical experiment showed that face completion manifested its effect between 50 and 250 ms after stimulus onset, which replicated our previous finding (Chen et al., 2009). It

should be noted that the time course for perceptual completion varies across different studies. For example, the completion in Murray et al. (2001) took place before 50 ms. The discrepancy could be attributed to task and stimulus differences. For example, completion time was found to depend on how much of the stimulus occluded - the more areas occluded, the longer time course needed (Shore and Enns, 1997). Here, we emphasizes that the time course measured in the current study might be specific to our task and stimuli. Based on the psychophysical measures, in the fMRI experiments, stimuli were presented with several durations that were designed to elicit face completion to various degrees. We found that BOLD responses in OFA and FFA were closely correlated with the psychophysical measures. Specifically, completed faces elicited significantly stronger BOLD responses than noncompleted faces when they were presented for 250 and 350 ms, but not for 50 and 150 ms. These results are consistent with previous fMRI studies by Lerner et al. (2002) and Hegdé et al. (2008). They presented occluded objects for hundreds of milliseconds without backward masking. Object completion related activation was found in object-selective cortical areas in both the ventral and the dorsal processing streams. Here we extended previous findings by showing that face completion was implemented progressively in the high-level visual cortex. The correlation between the psychophysical and the fMRI data in FFA was significantly higher than that in OFA. OFA is selective for face parts (i.e., eyes, nose, mouth) (Pitcher et al., 2007) and is thought to be at a lower position in the face processing hierarchy than FFA (Haxby et al., 2000). It is likely that OFA was responsible for completing face parts and FFA took a step forward to complete face configuration. Note that face configure could provide more reliable information for computing face orientation and determining subjects' behavioral performance. Harris and Aguirre (2008) did not find face completion effect in face-selective regions. A possible explanation for this discrepancy is that the perceptual contrast between the FBO and FIO stimuli in our study seems to be greater than that in their study since we used randomly positioned and irregular holes that made the perceptual grouping of face fragments much more difficult.

In addition to the activation by completed faces in OFA and FFA, we also found that completed faces induced weaker BOLD signals in V1 and V2 than noncompleted faces, only with stimulus durations of 50 and 150 ms. How was this decrease related to face completion? Murray et al. (2004) have suggested that perceptual grouping involves increases in activity in high-level visual areas that code for spatial patterns (e.g., objects, surfaces, and textures) along with decreases in activity in early visual areas that code for local, individual elements of the pattern (e.g., local orientation or direction of motion). They proposed that this inverse relationship in neural activity between high-level and early visual areas reflects an 'efficient code' of visual information. As high-level visual areas converge on a single, global hypothesis for the individual elements in a visual scene, early visual areas no longer need to represent the individual elements. Their view is consistent with predictive coding models (Mumford, 1992; Rao and Ballard, 1999) and has received support from fMRI and MEG studies (Murray et al., 2002; Summerfield et al., 2006; Furl et al., 2007; Harrison et al., 2007; Fang et al., 2008). In our study, perceptual grouping of face fragments into a coherent face (the FBO condition) increased activity in high-level visual areas and decreased activity in early visual areas. Can predictive coding models provide a good explanation of this response pattern? It should be noted that predictive coding models suggest that feedback from higher areas operates to reduce activity in lower areas. Pre-

dictive coding models usually posit a subtractive comparison between hypotheses generated in higher areas and incoming sensory input in lower areas. In these models, reduced activity occurs when the predictions of higher areas match incoming sensory information. However, we found that the decreased activity in V1 and V2 occurred before the increased activity in OFA and FFA, which means that the activity reduction in lower areas cannot be attributed to feedback from higher areas and renders the explanation from predictive coding models unlikely.

A more likely explanation to the completion-related response reduction in V1 and V2 is figure-ground segmentation in early visual processing phase. Zipser et al. (1996) and Marcus and Van Essen (2002) showed that responses of neurons in V1 and V2 could be enhanced by a small figure presented against a large background. The figural enhancement could occur as early as at the onset peak of neuronal response with latency of 40–80 ms (Marcus and Van Essen, 2002). In our study, the FIO stimuli contained many small figures presented against a large textured ground, but the FBO stimuli did not have such a clear figure-ground configuration (Fig. 1*C*). This difference could explain why the FIO stimuli induced a stronger response in V1 and V2 than the FBO stimuli. Indeed, the data presented by Zipser et al. (1996), their Figure 8, confirmed our postulation. The suppression effect in V1 and V2 suggests an important role of early visual cortical areas in face completion (Nakayama et al., 1989; Sugita, 1999; Bakin et al., 2000; Pillow and Rubin, 2002). Once V1 and V2 segmented figures from background in the FIO stimuli in the early processing phase, the segregated figures would be treated as independent objects and would not be further processed in the late phase for grouping or completion. On the other hand, the FBO stimuli did not suffer from such a processing constraint.

The completion-related response reduction in V1 and V2 only occurred with short stimulus durations. Lerner et al. (2002) and Hegdé et al. (2008) presented occluded objects for hundreds of milliseconds and did not find such a reduction. In another study by Lerner et al. (2004), they presented occluded objects as short as 60 ms, but no reduction was found in early cortical areas. It should be noted that their study differed from ours in a number of important respects, including the type of objects (objects vs faces), occluders (pictorial vs stereoscopic, vertical bas vs random holes), and so on. The most important difference is the extent of face completion with brief stimulus presentations. In the study by Lerner et al. (2004), a significant amount of completion has been done with 60 ms stimulus duration. But in our study, face completion manifested little effect at 50 ms. These evidence further suggests that the response reduction in early visual cortical areas is associated with the very early phase of face completion.

In the second fMRI experiment, we show that attention was necessary to the cortical dynamics underlying face completion. When subjects attended away from the face stimuli, both response suppression in the early visual areas and response enhancement in the high visual areas elicited by face completion were almost completely abolished. An fMRI study by Kouider et al. (2009) showed that attention could modulate the processing of backward masked face in high-level visual areas, even when the face was presented too brief (50 ms) to be aware. Here, we showed that early visual processing in early visual areas could also be modulated by attention, which resonates with the finding that ERPs could be modulated by spatial attention as early as at 70–80 ms after stimulus onset (Martínez et al., 1999; Frey et al., 2010). Note that attending away from the face stimuli might affect not only face completion, but also face perception itself. It is worthwhile to separate these two effects in future. The third experiment

rules out disparity difference as an alternative explanation for the neural events found in the first fMRI experiment. Indeed, to the best of our knowledge, no existing electrophysiological evidence would predict differential responses in early visual areas to the near and the far occluders (Cumming and DeAngelis, 2001; Parker, 2007).

In summary, the present study combined psychophysics and fMRI to examine the spatiotemporal dynamics of face completion in the visual processing hierarchy. We found that face completion involved early suppression in V1 and V2 and late enhancement in OFA and FFA. We also showed that attention is necessary to this response pattern. In future research, it will be of great interest to examine whether this pattern is also the neural substrate of other kinds of perceptual completion.

References

Bakin JS, Nakayama K, Gilbert CD (2000) Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. J Neurosci 20:8188–8198.

Bar M, Tootell RB, Schacter DL, Greve DN, Fischl B, Mendola JD, Rosen BR, Dale AM (2001) Cortical mechanisms specific to explicit visual object recognition. Neuron 29:529–535.

Chen J, Liu B, Chen B, Fang F (2009) Time course of amodal completion in face perception. Vision Res 49:752–758.

Cumming BG, DeAngelis GC (2001) The physiology of stereopsis. Annu Rev Neurosci 24:203–238.

Doniger GM, Foxe JJ, Murray MM, Higgins BA, Snodgrass JG, Schroeder CE, Javitt DC (2000) Activation timecourse of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. J Cogn Neurosci 12:615–621.

Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7:181–192.

Fang F, He S (2005) Viewer-centered object representation in the human visual system revealed by viewpoint aftereffect. Neuron 45:793–800.

Fang F, Kersten D, Murray SO (2008) Perceptual grouping and inverse fMRI activity patterns in human visual cortex. J Vis 8:2.1–9.

Fischer J, Whitney D (2009) Attention narrows position tuning of population responses in V1. Curr Biol 19:1356–1361.

Frey HP, Kelly SP, Lalor EC, Foxe JJ (2010) Early spatial attentional modulation of inputs to the fovea. J Neurosci 30:4547–4551.

Furl N, van Rijsbergen NJ, Treves A, Friston KJ, Dolan RJ (2007) Experience-dependent coding of facial expression in superior temporal sulcus. Proc Natl Acad Sci U S A 104:13485–13489.

Grill-Spector K, Malach R (2004) The human visual cortex. Annu Rev Neurosci 27:649–677.

Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. Nat Neurosci 3:837–843.

Grützner C, Uhlhaas PJ, Genc E, Kohler A, Singer W, Wibral M (2010) Neuroelectromagnetic correlates of perceptual closure processes. J Neurosci 30:8342–8352.

Harris A, Aguirre GK (2008) The representation of parts and wholes in face-selective cortex. J Cogn Neurosci 20:863–878.

Harrison LM, Stephan KE, Rees G, Friston KJ (2007) Extra-classical receptive field effects measured in striate cortex with fMRI. Neuroimage 34:1199–1208.

Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233.

Hegdé J, Fang F, Murray SO, Kersten D (2008) Preferential responses to occluded objects in the human visual cortex. J Vis 8:16.1–16.

Johnson JS, Olshausen BA (2005) The recognition of partially visible natural objects in the presence and absence of their occluders. Vision Res 45:3262–3276.

Kanwisher N (2010) Functional specificity in the human brain: a window into the functional architecture of the mind. Proc Natl Acad Sci U S A 107:11163–11170.

Kouider S, Eger E, Dolan R, Henson RN (2009) Activity in face-responsive brain regions is modulated by invisible, attended faces: evidence from masked priming. Cereb Cortex 19:13–23.

Lamme VAF, Zipser K, Spekreijse H (2002) Masking interrupts figure-ground signals in V1. J Cogn Neurosci 14:1044–1053.

Lerner Y, Hendler T, Malach R (2002) Object-completion effects in the human lateral occipital complex. Cereb Cortex 12:163–177.

Lerner Y, Harel M, Malach R (2004) Rapid completion effects in human high-order visual areas. Neuroimage 21:516–526.

Marcus DS, Van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. J Neurophysiol 88:2648–2658.

Martínez A, Anllo-Vento L, Sereno MI, Frank LR, Buxton RB, Dubowitz DJ, Wong EC, Hinrichs H, Heinze HJ, Hillyard SA (1999) Involvement of striate and extrastriate visual cortical areas in spatial attention. Nat Neurosci 2:364–369.

Mumford D (1992) On the computational architecture of the neo-cortex: II. The role of the cortico-cortical loops. Biol Cybern 66:241–251.

Murray RF, Sekuler AB, Bennett PJ (2001) Time course of amodal completion revealed by a shape discrimination task. Psychon Bull Rev 8:713–720.

Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL (2002) Shape perception reduces activity in human primary visual cortex. Proc Natl Acad Sci U S A 99:15164–15169.

Murray SO, Schrater P, Kersten D (2004) Perceptual grouping and the interactions between visual cortical areas. Neural Netw 17:695–705.

Nakayama K, Shimojo S, Silverman GH (1989) Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. Perception 18:55–68.

Parker AJ (2007) Binocular depth perception and the cerebral cortex. Nat Rev Neurosci 8:379–391.

Pillow J, Rubin N (2002) Perceptual completion across the vertical meridian and the role of early visual cortex. Neuron 33:805–813.

Pitcher D, Walsh V, Yovel G, Duchaine B (2007) TMS evidence for the involvement of the right occipital face area in early face processing. Curr Biol 17:1568–1573.

Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87.

Sekuler AB, Palmer SE (1992) Perception of partly occluded objects: a microgenetic analysis. J Exp Psychol Gen 121:95–111.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268:889–893.

Shore DI, Enns JT (1997) Shape completion time depends on the size of the occluded region. J Exp Psychol Hum Percept Perform 23:980–998.

Smith AM, Lewis BK, Ruttimann UE, Ye FQ, Sinnwell TM, Yang Y, Duyn JH, Frank JA (1999) Investigation of low frequency drift in fMRI signal. Neuroimage 9:526–533.

Stanley DA, Rubin N (2003) fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. Neuron 37:323–331.

Sugita Y (1999) Grouping of image fragments in primary visual cortex. Nature 401:269–272.

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.

Zipser K, Lamme VA, Schiller PH (1996) Contextual modulation in primary visual cortex. J Neurosci 15:7376–7389.