NeuroImage
doi:10.1006/nimg.2002.1243

Activity in the Fusiform Gyrus Predicts Conscious Perception of Rubin's Vase–Face Illusion

Fn1-2 Timothy J. Andrews,*,†,[1] Denis Schluppeck,*,[2] Dave Homfray,† Paul Matthews,† and Colin Blakemore*

*University Laboratory of Physiology, Parks Road, Oxford, United Kingdom; and †FMRIB,
John Radcliffe Hospital, Oxford, United Kingdom

Received December 31, 2001

We localized regions in the fusiform gyrus and superior temporal sulcus that were more active when subjects viewed photographs of real faces than when they viewed complex inanimate objects and other areas in the parahippocampal gyrus and the lateral occipital lobe that showed more activity during the presentation of nonface objects. Event-related functional magnetic resonance imaging was then used to monitor activity in these extrastriate visual areas while subjects viewed Rubin's vase–face stimulus and indicated switches in perception. Since the spontaneous shifts in interpretation were too rapid for direct correlation with hemodynamic responses, each reported percept (faces or vase) was prolonged by suddenly adding subtle local contrast gradients (embossing) to one side or the other of the figure–ground boundary, stabilizing the percept. Under these conditions, only face-selective areas in the fusiform gyrus responded more strongly during the perception of faces. To control for effects of the physical change to Rubin's stimulus (i.e., addition of embossing), we compared activity when the face contours were embossed after the subject had just reported the onset of perception of either faces or vase. Activity in the fusiform face area responded more strongly under the first condition, despite the fact that the physical stimulus sequences were identical. Moreover, on a trial-to-trial basis, the activity was statistically predictive of the subjects' responses, suggesting that the conscious perception of faces could be made explicit in this extrastriate visual area.

Key Words: **consciousness; objects; extrastriate visual cortex; choice probability; signal detection, FMRI.**

INTRODUCTION

In constructing a representation of the visual world, the brain has to cope with the fact that any given two-dimensional retinal image could be the projection of countless object configurations in the three-dimensional world. Although in most situations this inherent ambiguity is resolved by the visual system, there are occasions when human vision alternates between different interpretations of a stimulus. Common examples of such multistable or ambiguous stimuli include figure–ground reversals (Rubin, 1915), transparent three-dimensional objects (Necker, 1832; Purves and Andrews, 1997), and binocular rivalry (Blake, 1989). Ambiguous figures offer a potentially fruitful tool to probe the relationship between neural activity and sensory perception, because, in the absence of any extrinsic changes in the stimulus, perceptual alternations must be due to specific brain processes that lead to conscious vision.

The spontaneous alternations in perception that occur when different images are presented to the two eyes (binocular rivalry) has many features in common with that experienced when viewing other ambiguous stimuli (Logothetis *et al.,* 1996; Andrews and Purves, 1997). A number of reports have suggested that activity in relatively "high" areas of visual cortex correlates with changes in perception that occur during binocular rivalry in both monkeys (Leopold and Logothetis, 1996; Sheinberg and Logothetis, 1997) and humans (Tong *et al.,* 1998). However, recent evidence suggests that the simple rivalry between contours of different orientation depends on inhibitory or suppressive interactions occurring in primary visual cortex (Polonsky *et al.,* 2000; Tong and Engel, 2001), as suggested by Sengpiel and Blakemore (1994). If this is the case, then any changes in the activity of neurons in higher visual areas could be attributed to suppression at the input stage.

This recent controversy suggests that the mechanism underlying rivalry might be quite different, in nature and location, from that causing shifts in the

---

[1] To whom reprint requests should be addressed at present address: Department of Psychology, University of Durham, UK. E-mail: tim.andrews@durham.ac.uk. Fax: +44-1642-385945.

[2] Present address: Department of Psychology, UCLA, Los Angeles, CA.

1053-8119/02 $35.00

perception of other ambiguous figures (Andrews, 2001; Blake and Logothetis, 2002). In our study, we investigated human cortical activity while subjects were viewing Rubin's classical vase–face stimulus, where different stimulus interpretations (faces and vase) are clearly competing. We took advantage of the fact that inanimate objects and faces are known to be analyzed in different areas of extrastriate visual cortex (for review see Tanaka, 1997). However, the selectivity of neural responses to these different classes of stimuli does not in itself demonstrate that the conscious perception of a face or an object is made explicit in these visual areas. It could be that this activity represents a divergence of processing before the level at which percepts arise. Indeed, it is also possible that explicit representations of faces and objects cannot be localized to particular areas in the brain, but are widely distributed (Haxby *et al.,* 2001). One way to distinguish between these possibilities is to monitor fMRI activity in visual cortical areas while subjects indicate changes in the perception of Rubin's stimulus.

In a previous study, Kleinschmidt *et al.* (1998) examined activity resulting from ambiguous stimuli, but they asked a rather different question, namely, "which areas are specifically related to perceptual switches, regardless of the particular perceptual interpretation?" They reported that regions in the occipital, parietal, and frontal lobes were more active during intervals when a shift occurred compared to periods of perceptual stability. However, their analysis did not discriminate the direction and nature of perceptual changes and conflated data from scans using different ambiguous stimuli. More recently, Hasson *et al.* asked whether face- and object-selective regions of visual cortex could discriminate between the different perceptual interpretations of Rubin's vase–face stimulus. They found that the fusiform gyrus was more active when the vase–face stimulus was biased toward the face by the use of color or texture. However, this does not alone imply that the perception of a face is made explicit in this area. It could be, for example, that this activity simply represents a response to the change in the stimulus.

Our purpose was to extend the approaches used by Kleinschmidt *et al.* (1998) and Hasson *et al.* (2001) by defining activity in specific visual areas during shifts of face and object awareness. We hypothesized that a cortical area that makes "explicit" in its activity the interpretation of a face (the sine qua non of conscious awareness of a face: see Crick and Koch, 1995) would show more activity for a perceptual transition from vase to face than for a shift from face to vase. Conversely, areas directly involved in or leading to the awareness of inanimate objects ought to display an opposite pattern of activity. We further posited that, if an area is involved in a specific aspect of visual awareness, the trial-by-trial variation in activity should correlate with the subjects' perceptual responses.

MATERIALS AND METHODS

Subjects

All three observers were right-handed healthy adults with normal or corrected to normal visual acuity. Informed consent was obtained from all subjects and the study was approved by the Central Oxford Research Ethics Committee (COREC 98.161). Stimuli were back-projected (Focus LP1000, Unicol Engineering, Oxford, UK) onto a screen placed at a distance of 280 cm from the subject's eyes. Subjects lay supine in the magnet bore and viewed the back-projection screen outside the bore through prism glasses (Wardray-Premise, Thames Ditton, UK). Data were collected from subjects over three to four sessions.

Imaging Parameters

All experiments were carried out using the Siemens-Varian 3 Tesla MRI scanner at the FMRIB center in Oxford. A Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, head, radiofrequency coil tuned to 127.4 MHz. A gradient-echo EPI sequence was used for image collection. Sixteen contiguous axial slices were employed to cover the brain (TR, 2 ss; TE, 30 ms; FOV, 256 × 256 mm; in-plane resolution, 4 × 4 mm; slice thickness, 7 mm). T1-weighted structural images were acquired with a 3D Turbo Flash Sequence at a resolution of 1 × 1 mm within slice and 3.5 mm between slices. Image segmentation to extract brain was carried out using BET, FMRIB's Brain Extract Tool (Smith *et al.,* 2000; www.fmrib.ox.ac.uk/fsl). In order to facilitate anatomical localization of the foci of activation, statistical maps from the echo-planar imaging were registered to high-resolution structural images of the subjects. Additionally, the statistical maps were registered onto a standard image in Talairach space (Montreal Neurological Institute, MNI average 152 T1 brain). Registration was carried out using FLIRT (Jenkinson and Smith, 2000; www.fmrib.ox.ac.uk/fsl).

Experiment 1.1

To discriminate areas in visual cortex selectively activated for faces and objects, a localizer scan was carried out in each session. The stimuli were gray scale photographs of actual faces and objects (approx. 8° × 8°). Examples are shown in Fig. 1a. Front and profile views of faces were taken from a database of the Psychological Image Collection at Stirling (PICS, http://pics.psych.stir.ac.uk/) and were not familiar to any of the subjects. Photographs of inanimate objects were obtained from various sources, including the PICS da-

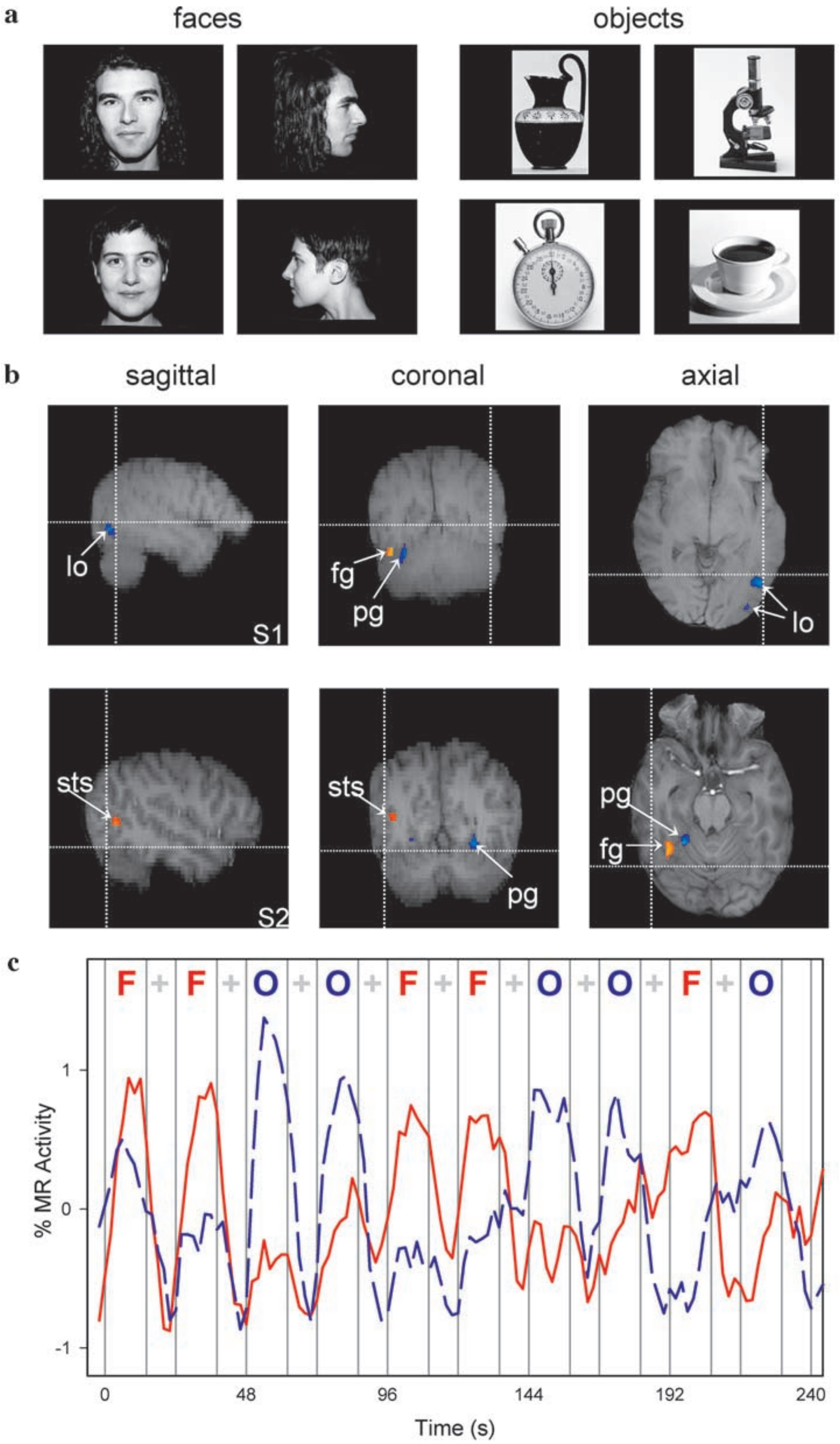

FIG. 1. Localizer scan (Experiment 1.1). (a) Examples of faces and objects used to localize selective regions in visual cortex. (b) Location of areas in visual cortex that showed selective responses to faces (orange-red) or objects (blue) in two subjects, S1 and S2. Areas in the right fusiform gyrus (fg) were significantly more active when subjects looked at faces than when they viewed an assortment of common inanimate objects. Face-specific responses were also found in the right superior temporal sulcus (sts) of some subjects. Conversely, areas in the parahippocampal gyrus (pg) bilaterally and on the lateral aspect of the occipital lobe (lo) of the left hemisphere were activated by nonface objects. These scan images follow radiological convention, with the left hemisphere shown on the right. The axial section is defined by the

tabase, from Microsoft clip art, and from a book on antique vases. During each localizer scan, subjects were presented with alternating blocks of faces or of nonface objects in rapid sequence (12 images per 14-s block). There were no significant differences in the average luminance of the object and face images. Face and object blocks were separated by 10-s presentations of a gray screen, of the same space-averaged luminance.

Analysis of the localizer scans was carried out using FEAT, the FMRIB Easy Analysis Tool (www.fmrib.ox.ac.uk/fsl) integrated into MEDx (Sensor Systems, VA). Statistical analysis was carried out using FILM (FMRIB's Improved Linear Model) with local autocorrelation correction (Woolrich *et al.,* 2000). The initial four TRs (8 s) of data from each scan were discarded to minimize the effects of magnetic saturation and visual adaptation. The following prestatistics processing was applied to all EPI scans: 3D motion correction, using AIR (Woods *et al.,* 1998); spatial smoothing using a Gaussian kernel of FWHM 5.0 mm; mean-based intensity normalization of all volumes by the same factor; nonlinear high-pass temporal filtering (Gaussian-weighted LSF straight line fitting, with $\sigma = 7.5$ s). Z (Gaussianized T) statistic images were thresholded using resel (corrected Bonferroni) thresholding with a corrected significance threshold of $P < 0.05$ (Friston *et al.,* 1995; Forman *et al.,* 1995). Areas defined as face-selective included all voxels that responded significantly more to faces than to objects, whereas object-selective areas included voxels that responded more to inanimate objects than to faces at this level of significance. To define the face and object voxels for further analysis, the statistical images from the localizer experiments were registered onto the event-related EPI data set using FLIRT.

Experiment 1.2

To determine the temporal characteristics of the response in the face- and object-selective areas to single presentations of faces and objects, we devised an event-related paradigm for two subjects. An event involved a single presentation of a face or an object for 2 s followed by a gray screen of the same average luminance for 8 s. In each scan, 20 faces and 20 objects were randomly interleaved. The time series of the resulting filtered MR data at each voxel was converted from units of image intensity to units of fractional signal change (% MR activity). The statistical images from previously performed localizer scans were used as masks. The time-course plots were also normalized to the activity at stimulus onset. Signals were then averaged separately for the face and object events in the face- and object-selective areas.

Experiment 2.1

We next determined whether the activities in the regions of interest from Experiment 1 were selective for versions of the vase–face stimulus, for which perception was biased toward either the vase or faces interpretation. Subjects were initially presented with the "standard" vase–face stimulus (8° × 10°). They were instructed to fixate a small cross in the middle of the stimulus (see Fig. 3) and to indicate perceptual switches between *faces* and *vase* by pressing one of two buttons as soon as each new interpretation was perceived. After 2 s of normal viewing, the stimulus display software was programmed to present an embossed version, with subtle contrast gradients added to one side or the other of the boundaries between *vase* and *faces* (see Fig. 3). This embossing reduced the ambiguity of the stimulus, biasing perception toward one interpretation or the other. In this experiment the direction of embossing was contingent on the direction of the perceptual switch just reported by the subject. If a *vase-to-faces* transition had just been signaled, the face side of each boundary was embossed, and hence the subject continued to perceive the faces. Whenever the reported perceptual switch was from *faces-to-vase,* the edges of the vase contour were embossed, thus prolonging the perception of the vase. The embossed image was presented immediately after a reported change in perception and subjects were also asked to indicate whether the embossing biased perception in the predicted way. The embossed image was presented for 2 s, followed by an equiluminant gray screen, with a fixation cross in the same position, for 8 s before the normal ambiguous figure was again displayed.

Two strategies were employed to determine activity when subjects viewed unambiguous versions of the vase-face stimulus. The first involved analyzing the time series of activity following a spontaneous onset of a particular perception, as signaled by the button press, followed by embossing of one contour or the other. The MR activity for each voxel falling within the areas previously defined by the localizer scan was converted into units of fractional signal change and normalized to the level at the time of the perceptual

---

horizontal dashed line on the sagittal and coronal sections, whereas the sagittal section is given by the vertical dashed lines in the axial and coronal sections, and the coronal section is defined by the horizontal and vertical lines in the axial and saggital sections respectively. (c) MR time course during localizer scans, showing the average activity of all face-selective voxels (red line) and object-selective voxels (blue dashed line) averaged across three subjects. Subjects viewed a randomized sequence of blocks of faces (F) and of objects (O). Each block, lasting 14 s, consisted of 12 different face or object images. Interleaved between the blocks were 10-s presentations of a gray screen (+).

switch. A repeated-measures ANOVA was used to determine whether there was a significant increase in activity in the 6 s following the perceptual transition that lead to the presentation of the embossed image.

The second strategy involved a simple average of the integrated MR activity following the initial change in perception. On a trial-by-trial basis, the change in percentage of MR signal was integrated from 0 to 6 s (3 TRs) after the initial perceptual transition. An ANOVA was then performed to determine statistical difference between the means of the two conditions (*vase-to-face, face-to-vase*).

Experiment 2.2

To control for effects of the physical change in stimulus, we next looked at responses to conditions in which the sequence of stimuli was identical, but perception was different. Activations associated with a particular embossing were compared, depending on whether the immediately preceding, spontaneous perceptual switch was confirmed or canceled by the embossing. Consider the following sequences:

1. Switch from *vase–faces,* followed by embossed faces identical stimulus sequence.
2. Switch from *faces–vase,* followed by embossed faces

identical stimulus sequence.

Although the stimulus presentations are absolutely identical, in (1) the spontaneous onset of faces perception is confirmed and strengthened by the embossing, whereas for (2) the embossing disconfirms the vase perception and eventually leads to the faces being perceived. As in Experiment 2.1, the embossed image was presented for 2 s and was followed by a gray fixation screen (8 s), before the next ambiguous stimulus. To confirm that the embossed image had the anticipated effect, subjects reported their perception throughout the trial.

To assess whether the activity of previously defined areas of visual cortex could predict perception when viewing the ambiguous stimulus, we calculated choice probability, a measure derived from signal detection theory (Green and Swets, 1974), which has also been used to characterize the responses of single neurons (e.g., Britten *et al.,* 1996; Dodd *et al.,* 2001). The MR activity from a region of interest was integrated on a trial-by-trial basis (0–6 s after stimulus onset) and histograms were constructed for the two different conditions, i.e., trials in which the initial percept was confirmed and prolonged and trials when the direction of embossing caused it to change. The proportion of trials in which the MR activity correctly (hits) and incorrectly (false alarms) predicted the direction of perceptual change was calculated. The hits were plotted against the false alarms in a receiver operating characteristic (ROC) curve. The area under the ROC curve is a nonparametric measure of discriminability and provides an index of the link between neural activity, as measured by fMRI and conscious perception. The reliability of the choice probability values was assessed using randomization testing. The values of change in MR activity of individual trials were randomly reallocated to one of the two conditions. With this randomized set, histograms were again plotted and ROC curves constructed. This process was repeated 1000 times and a distribution of area values obtained. The original value was accepted as statistically significant if it was higher than the 95th percentile of the distribution of randomized values (corresponding to $P < 0.05$; Manly, 1997).

To determine whether switches in perception of Rubin's vase–face stimulus might be correlated with changes in gaze (which would alter the retinal stimulus and might generate MR signals that could contaminate the results), we monitored eye position outside the scanner in one of the subjects (S3). Horizontal and vertical eye position was recorded using a Fourward Technologies Dual Purkinje Image Eye Tracker (Crane and Steele, 1985) while a sequence of stimuli identical to that described above was viewed. Eye position was sampled at 200 Hz via a CIO-DAS802 digital to analogue converter (Adept Scientific Plc, UK) during 10 *vase-to-faces* switches and 10 *faces-to-vase* switches. We then compared the average horizontal and vertical eye position in the second preceding a *vase-to-faces* or a *faces-to-vase* switch using a *t* test.

RESULTS

Experiment 1.1

Spatially discrete face-selective and object-selective areas were found in all subjects (Fig. 1b). A small region in the right fusiform gyrus showed significant activation for faces versus nonface objects. Face-selective responses were also detected in a region of the right superior temporal sulcus. Object-selective responses were found bilaterally in the parahippocampal gyrus and the lateral aspect of the left occipital lobe. The size and location of statistically significant activations are listed in Tables 1 and 2 and were consistent with previous reports (Sergent *et al.,* 1992; Puce *et al.,* 1995; Malach *et al.,* 1995; Kanwisher *et al.,* 1997; McCarthy *et al.,* 1997; Grill-Spector *et al.,* 1999). These regions of interest were used as a mask in subsequent analyses. The average time courses of activation in the face- and object-selective areas during the localizer scans with unambiguous images are shown in Fig. 1c.

T1-2

Experiment 1.2

The block design used in Experiment 1.1 provided a good signal-to-noise ratio and a reliable method to localize face and object areas. However, for subsequent parts of this study, it was important to determine

TABLE 1

Extent and Location (Talairach Coordinates) of Face-Selective Areas

	Fusiform gyrus				Superior temporal sulcus			
	x	*y*	*z*	Volume (cm^3)	*x*	*y*	*z*	Volume (cm^3)
S1	44	−59	−15	1.1				
S2	38	−50	−28	0.6	45	−44	5	1.7
S3	38	−56	−32	0.8	56	−34	6	0.9

whether there was also reliable activation of voxels to presentations of single stimuli. Figure 2 shows that the face and object areas were indeed both activated by single presentations of faces and objects, respectively. Following the presentation of a face, there was an increase in MR activity of voxels in face-selective areas that reached a maximum after 4–6 s. The magnitude of this pattern was significantly larger than that when an object was presented. Object-selective voxels showed an inverse pattern of activity responding more to a brief presentation of an object than to a face. These results demonstrate that the face and object selectivity in the areas previously defined in a blocked design is also high in an event-related design.

Experiment 2.1

We now monitored activity in the face- and object-selective areas, already defined with unambiguous images, when subjects viewed the Rubin vase–face stimulus (Fig. 3). When viewing the "standard" Rubin stimulus, subjects reported alternations between perception of *vase* and *faces* every few seconds (mean *vase* duration (seconds) ± SD, S1, 2.3 ± 0.6; S2 4.0 ± 0.3; S3, 1.9 ± 0.3) (mean *faces* duration (seconds) ± SD, S1, 3.3 ± 0.9; S2 4.0 ± 0.4; S3, 3.2 ± 1.1). After a 2-s delay, the software triggered the presentation of an embossed version of the stimulus, immediately after the report of a spontaneous change in perception (*vase-to-faces* or *faces-to-vase*). The particular embossing (emphasizing either vase or faces) was selected on the basis of the perceptual transition that had just been reported and always reinforced the new perceptual interpretation (i.e., vase–faces → embossed face; faces-to-vase → embossed vase): see Fig 3. The embossed image biased perception in the predicted way for all subjects in all trials. The interval prior to the presentation of the embossed stimulus was dependent on the length of time taken to report the expected perceptual transition and therefore varied from trial to trial, but was about 6–7 s (mean ± SD, 6.7 ± 1.2 s).

Figure 4 shows the pattern of activity in face-selective areas following this procedure. Activity in the fusiform gyrus increased significantly in all subjects following a *vase-to-faces* perceptual transition (S1, $F > 12$, $P < 0.0001$; S2, $F > 3$, $P < 0.05$; S3, $F > 6$, $P < 0.001$), but only showed a significant increase in two of the three subjects following a *faces-to-vase* switch (S1, $F > 5$, $P < 0.005$; S3, $F > 2.5$, $P < 0.05$). In all three subjects, voxels in the right fusiform gyrus were more active following the *vase-to-faces* perceptual transition (prolonged by embossing of the faces) than for *faces-to-vase* transitions (prolonged by embossing of the vase) (Fig. 4a). However, the difference in the summed MR activity following the *vase-to-faces* switch was significantly greater than that following a *faces-to-vase* switch in only two of the three subjects (S1, $F > 10$; $P < 0.001$; S3; $F > 3.5$; $P < 0.05$).

The activity in the superior temporal sulcus is shown in Fig. 4b. Although the MR activity increased to a maximum 4–6 s following the *vase-to-faces* transition (prolonged by embossing of the faces) (S2, $F > 4$, $P < 0.01$; S3, $F > 3.5$, $P < 0.01$), the integrated MR activity

TABLE 2

Extent and Location (Talairach Coordinates) of Object-Selective Areas

	Parahippocampal gyrus				Lateral occipital lobe			
	x	*y*	*z*	Volume (cm^3)	*x*	*y*	*z*	Volume (cm^3)
S1	30	−64	−14	3.1				
	−34	−64	−20	1.1	−39	−76	−2	2.3
S2	28	−46	−12	1.1				
	−26	−52	−10	1.3				
S3	28	−62	−20	0.7				
	−30	−54	−32	0.6	−24	−89	−11	1.2

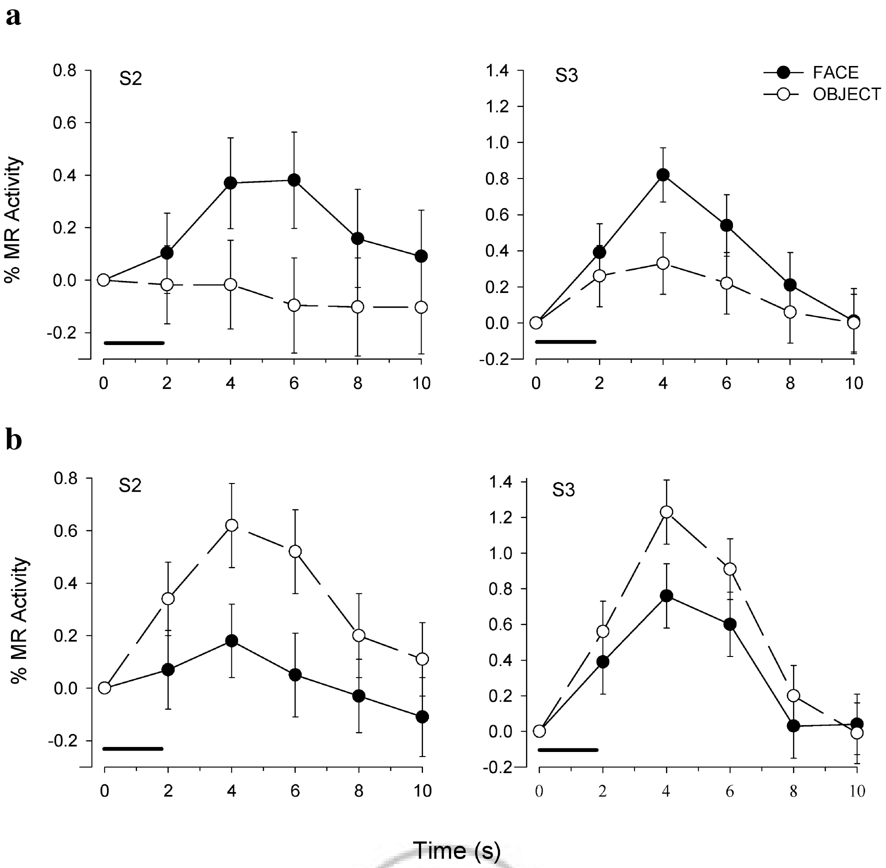

FIG. 2. (a) Face-selective and (b) object-selective areas defined using the block design (see Fig. 1) also show selective activation in an event-related paradigm (Experiment 1.2). The data represent the MR activity for individual 2-s presentations of faces (filled circles, solid lines) and objects (unfilled circles, dashed lines), each followed by 8 s of a blank gray screen. Each curve is the average result from 20 presentations in two subjects, S2 and S3. Error bars represent SEM.

was not significantly different than the *faces-to-vase* transitions (followed by embossing of the vase) (S2, $F < 0.3$, $P > 0.7$; S3, $F < 0.2$, $P > 0.7$). Data are shown for the superior temporal sulcus in only two subjects, because S1 showed no consistent activity in this area for face images during the localizer scan.

Object-selective areas in visual cortex did not show a clear difference in their responses following *faces-to-vase* or *vase-to-face* transitions. The average patterns of activity in the parahippocampal gyrus and lateral F5 occipital complex are shown for all subjects in Fig. 5. There was a nonselective increase in activity following the presentation of the embossed image in all subjects (S1–S3, parahippocampal gyrus, $F > 3$, $P < 0.05$; lateral occipital complex, $F > 3$, $P < 0.05$).

Experiment 2.2

To control for effects of the physical change in stimulus, we next looked at responses to conditions in which the sequence of stimuli was identical, but perception was different. Activations associated with a particular embossing were compared, depending on whether the immediately preceding, spontaneous perceptual switch was canceled or confirmed by the embossing (Fig. 6a). Figure 6b shows the averaged activ- F6 ity in the fusiform gyrus, integrated over 6 s, when the embossed faces followed either a *vase-to-faces* transition (filled bars) or *faces-to-vase* switch. To analyze the data, we applied a method adapted from signal detection theory to take into account the trial-to-trial variability in signal within subjects. Analysis of the data showed that the choice probabilities were significant in the fusiform gyrus for each subject (S1, 0.61, $P < 0.01$; S2, 0.62, $P < 0.05$; S3, 0.61, $P < 0.05$). Choice probability values > 0.5 indicate that an increase in the activity of that particular area correctly predicts the vase-to-faces transition (Fig. 6c).

To determine whether changes in gaze are necessary to affect a change in perception, we monitored horizontal and vertical eye position outside the scanner in one subject while viewing the vase–face stimulus. Figure 7 F7 shows that switches in perception can occur in the absence of any changes in horizontal ($t = 1.2$; $P > 0.25$) or vertical ($t = 0.57$; $P > 0.55$) eye position.

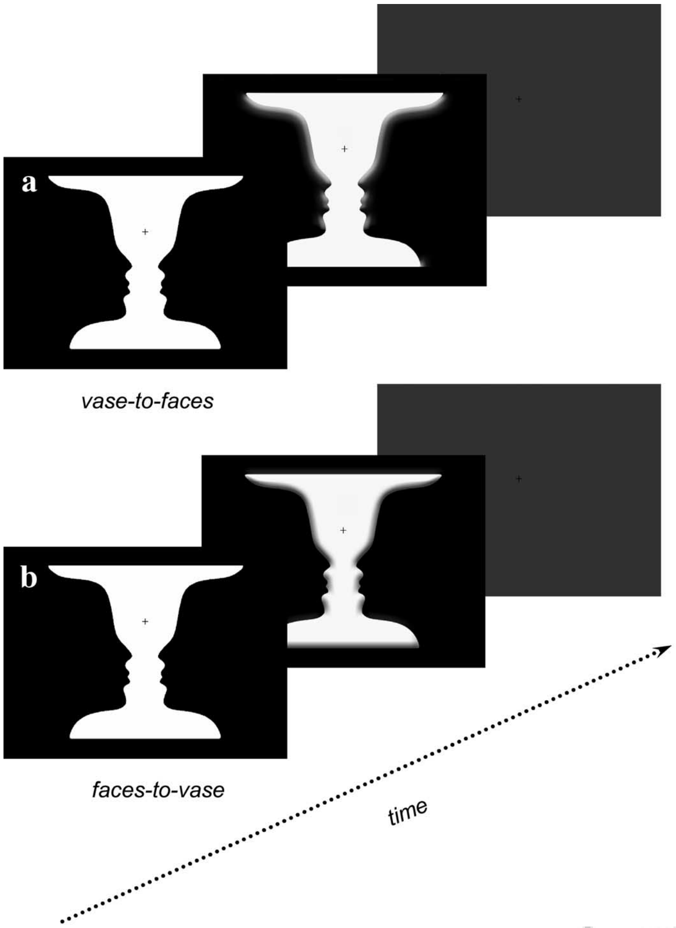

FIG. 3. Stimulus and task for Experiment 2.1. Subjects were initially presented with the Rubin's ambiguous vase–face stimulus and instructed to indicate alternations in perception between the vase and faces by button presses. Since the periods of spontaneous perceptual dominance were short compared with the time constant of the BOLD response, we designed an event-related paradigm in which the sudden onset of perception of either (a) the faces or (b) the vase was prolonged by adding subtle local contrast gradients (embossing) to one edge or other of the figure–ground boundary. In (a), following a *vase-to-faces* transition, the standard image was replaced by an embossed-face version of the same stimulus, whereas in (b), subsequent to a *faces-to-vase* change, an embossed-vase version replaced the standard.

DISCUSSION

We confirm previous studies, locating regions of visual cortex selective for unambiguous faces in the fusiform gyrus and the superior temporal sulcus (Sergent *et al.,* 1992; Haxby *et al.,* 1994; Puce *et al.,* 1995; Kanwisher *et al.,* 1997; Allison *et al.,* 1999). Object-selective areas were found in the parahippocampal gyrus (Epstein and Kanwisher, 1998; Kanwisher *et al.,* 1997; Ishai *et al.,* 1999) and the lateral occipital lobe (Malach *et al.,* 1995; Grill-Spector *et al.,* 1999). These regions of interest were selectively activated when subjects viewed photographs of faces and objects presented either in blocks or as single events.

To determine whether the perception of faces and objects is also made explicit in these areas, in the sense that activity correlates with conscious perception regardless of the physical stimulus, we monitored activity when subjects viewed Rubin's ambiguous vase–face stimulus. We predicted that sudden transitions to a perceptual interpretation preferred by a region (*vase-to-faces* for face-selective regions; *faces-to-vase* for object-selective regions) might be reflected by increased activation compared to switches away from the preferred percept (*faces-to-vase* for face-selective regions; *vase-to-faces* for object-selective regions). Since the frequency of spontaneous perceptual change was too rapid to be followed by the underlying BOLD response, we devised the procedure of adding local contrast gradients to the edges of the ambiguous stimulus (see Fig. 3), so as to prolong perception of either the vase or the faces after a perceptual transition. The activity of face-selective voxels in the fusiform gyrus did indeed discriminate between the alternative perceptions of Rubin's stimulus in this paradigm (Fig. 4). Greater activity was detected following *vase-to-faces* transitions than during *faces-to-vase* changes.

These results are consistent with a recently published report showing that the fusiform gyrus is more active when the vase–face stimulus is biased toward the face by the use of color or texture (Hasson *et al.,* 2001; see also Andrews *et al.,* 2000). The data are also, in part, consistent with an earlier fMRI study of binocular rivalry in which complex objects (houses) and faces were presented independently to the two eyes and changes in perception from *house-to-face* were associated with increased activity in the fusiform gyrus (Tong *et al.,* 1998). However, the authors also reported that changes from *face-to-house* resulted in a decrease in MR signal, while we did not find a systematic decrease following perceptual switches to the nonpreferred (*vase*) percept. One explanation for the disparity between these studies is differences in study design. In our study, the stimulus was changed, albeit subtly, by addition of the embossing during the measurement period. This change may have caused a nonspecific increase in activity that was not evident in the rival paradigm, in which the stimulus remained unchanged. However, Kleinschmidt *et al.* (1998) also report only increases in activity in the fusiform gyrus during changes in perception when subjects viewed Rubin's vase–face stimulus, although they did not discriminate the direction of perceptual changes. This leads us to speculate that the competitive interactions underlying binocular rivalry between complex figures might, in some circumstances, employ a different mechanism than that involved in the interpretation of conventional ambiguous figures (cf. Andrews, 2001).

Activity in voxels in the superior temporal sulcus that were selective for unambiguous faces were not statistically significant predictors for the two percep-

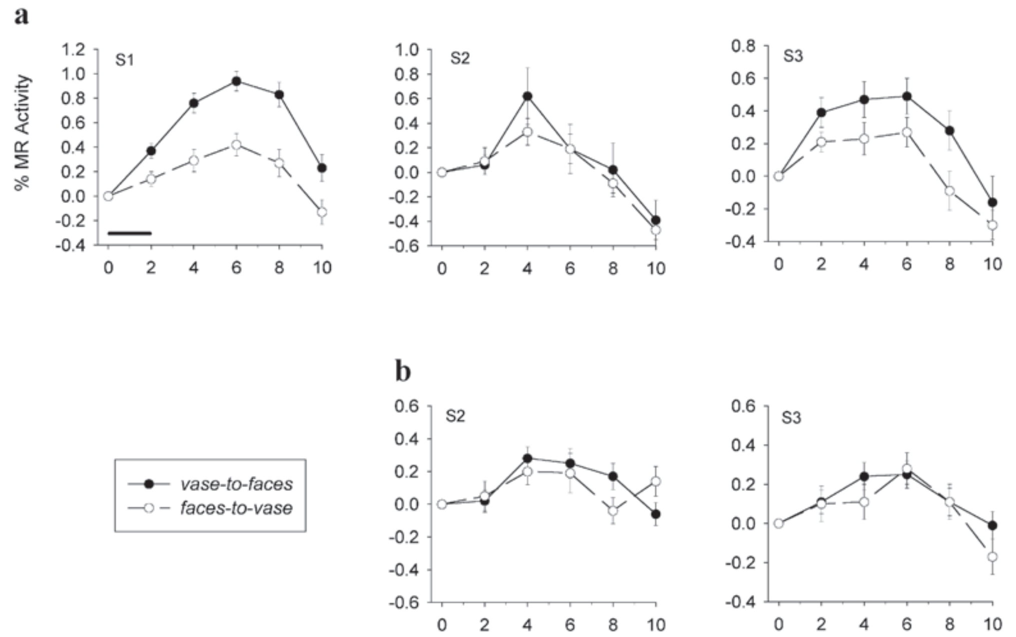

FIG. 4. Face-selective voxels in the fusiform gyrus, but not the superior temporal sulcus, reflect the perceptual interpretation of a face in the Rubin figure, when prolonged by the embossing technique (Experiment 2.1). Each curve represents the mean time course of MR signal averaged over many presentations (~40) and for all face-selective voxels in either the fusiform gyrus (a) or the superior temporal sulcus (b). The activity following spontaneous *vase-to-faces* transitions, prolonged by embossing the faces (filled circles, solid line), was compared with that following *faces-to-vase* transitions, reinforced by embossing of the vase (unfilled circles, interrupted line). No data are shown for the superior temporal sulcus in subject 1, because no significant activity was detected there during the localizer scans. The horizontal line represents the duration of the embossed image. Error bars represent SEM.

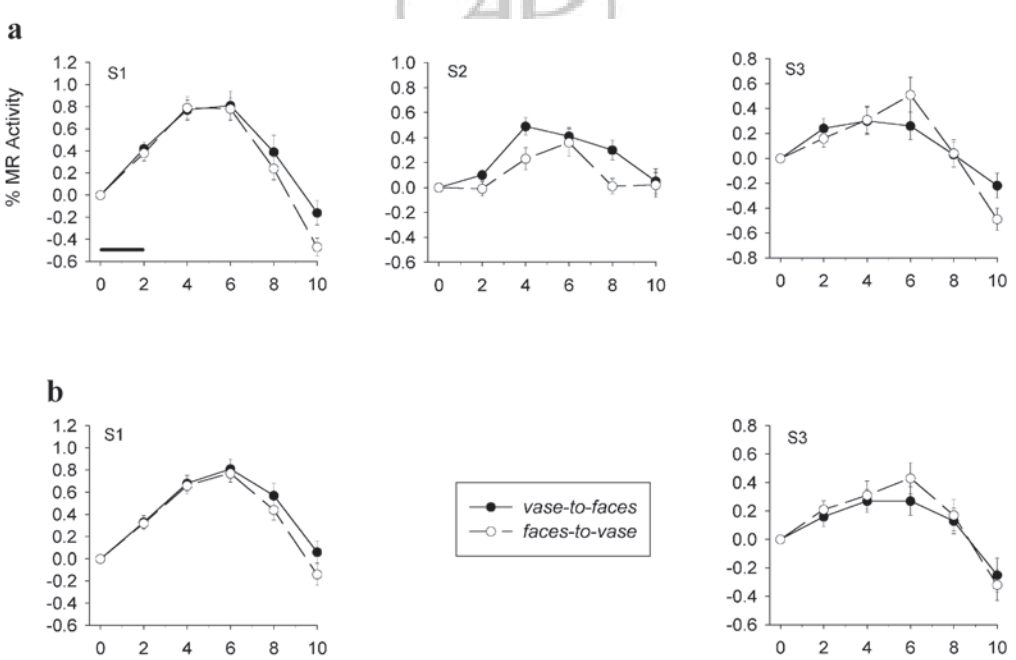

FIG. 5. Object-selective voxels in extrastriate visual cortex do not show selectivity for the perceptual interpretation of the vase in the Rubin figure, when prolonged by the embossing technique (Experiment 2.1). The results from Experiment 2.1 are displayed as in Fig. 4, except that data are from object-selective voxels in either the parahippocampal gyrus (a) or the lateral occipital complex (b). Subject 2 showed no significant activity in the lateral occipital complex during the localizer scans. The horizontal line represents the duration of the embossed image. Error bars represent SEM.

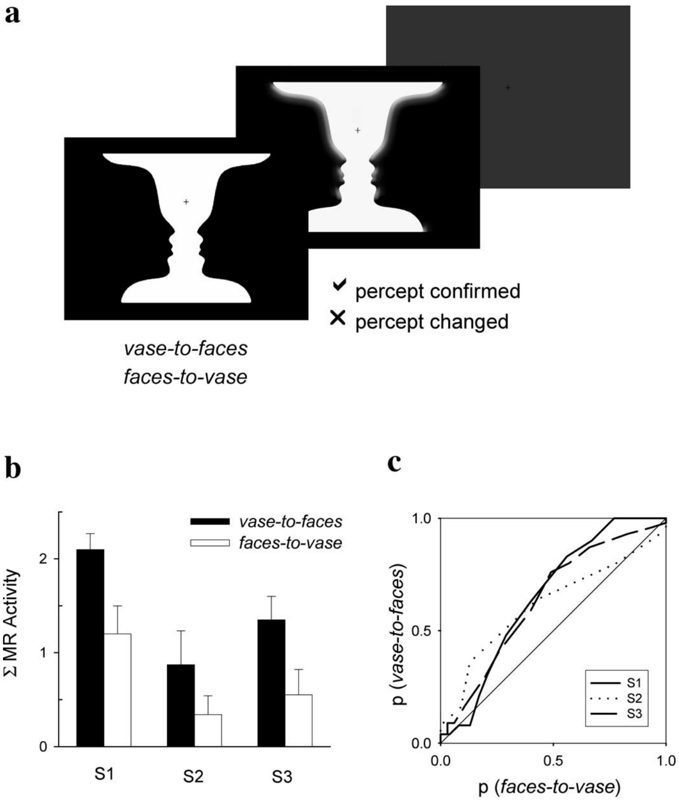

FIG. 6. The face-selective area in the right fusiform gyrus shows different activity for identical physical stimulus sequences, depending on the perceptual interpretation (Experiment 2.2). (a) From the stream of data, responses were selected for all sequences in which a change of perception for the standard Rubin's stimulus, signaled by a button press, was followed by embossing of the *face* contours, thus biasing subsequent percepton toward the faces percept. (b) Each bar represents the integrated MR response for 6 s after the spontaneous shift of perception, which was followed by the embossed faces, averaged over many presentations (~40). Data are from all face-selective voxels in the right fusiform gyrus of each subject (S1, S2, S3). Filled bars show results when the preceding spontaneous perceptual transition had been from *vase-to-faces,* for which the embossing prolonged the perceptual interpretation of the faces. Unfilled bars show results for with the preceding perceptual switch had been from *faces-to-vase,* in which the embossing contradicted the existing perceptual interpretation. Note that the sequence of physical images was identical under both conditions, yet more activity occurred if the perception of faces in the ambiguous figure was confirmed by the embossing than if it was changed. Error bars represent SEM. (c) To assess whether the activity of the fusiform gyrus could predict perception when viewing the ambiguous stimulus on a trial-by-trial basis, we calculated the proportion of trials in which the MR activity correctly (hits) and incorrectly (false alarms) predicted the direction of perceptual change. The hits were plotted against the false alarms in a receiver operating characteristic curve. Analysis of the data showed significant choice probabilities were apparent in the fusiform gyrus for each subject.

tual interpretations of Rubin's image. One explanation for why this area is able to distinguish between photographs of faces and objects, but is not able to discriminate between the vase and faces in the ambiguous stimulus is the paucity of explicit detail in the latter. It is possible that areas in the superior temporal sulcus are more concerned with the details of facial structure (Haxby *et al.,* 2000; Allison *et al.,* 2000). For example, changeable aspects of the face such as direction of eye gaze (Perrett *et al.,* 1985; Hoffman and Haxby, 2000), facial expression (Hasselmo *et al.,* 1989; Perrett and Mistlin, 1990), and lip movement (Calvert *et al.,* 1997) all evoke selective responses in the superior temporal sulcus. Consistent with these findings, lesions to the superior temporal sulcus in nonhuman primates do not impair face recognition (Heywood and Cowey, 1992).

Areas selective for unambiguous inanimate objects were similarly unable to discriminate the direction of perceptual change when viewing Rubin's stimulus. Perceptual transitions to the preferred percept (*faces-to-vase*) did not produce more activity than shifts to the nonpreferred percept (*vase-to-faces*) in either the parahippocampal gyrus or the lateral occipital lobe. Again, the most parsimonious explanation for this result is that the vase representation is a less salient percept than the photographs of objects that were used to define this area.

The observation that activity in the fusiform gyrus was selective for the different conscious interpretations of Rubin's stimulus when prolonged by embossing (Experiment 2.1; see also Hasson *et al.,* 2001) does not alone imply that the perception of a face is made explicit in this area. It could be, for example, that this activity simply reflects differential responsiveness to the relatively unambiguous embossed image, rather than to the initial spontaneous switch to perception of faces. To control for the change in the stimulus, in the critical Experiment 2.2, we compared activity in the fusiform gyrus when the sequence of stimuli was identical, but perception was different. We found that even when the physical stimulation remained the same, more activity was recorded in the fusiform gyrus when a *vase-to-faces* transition preceded the presentation of an embossed face than when a *faces-to-vase* switch was initially reported.

To determine whether face-selective areas in the fusiform gyrus were predictive of the subjects' responses on individual trials, we employed the analyses of choice probability (see Britten *et al.,* 1996). This method allowed us to take into account the trial-to-trial variability in signal for each subject; this aspect of the data is usually neglected when considering only the average activity from many subjects (cf. Parker and Newsome, 1988). We found that the face-selective area in the fusiform gyrus was statistically predictive of the subjects' responses with significant choice probabilities (>0.6), for all subjects. The implication is that activity in the fusiform gyrus could make a decision that leads directly to the perception of a face. Our results are consistent with those of other studies that have shown that the responses of regions within the temporal lobe are modulated by selective attention to faces (Wojciulik

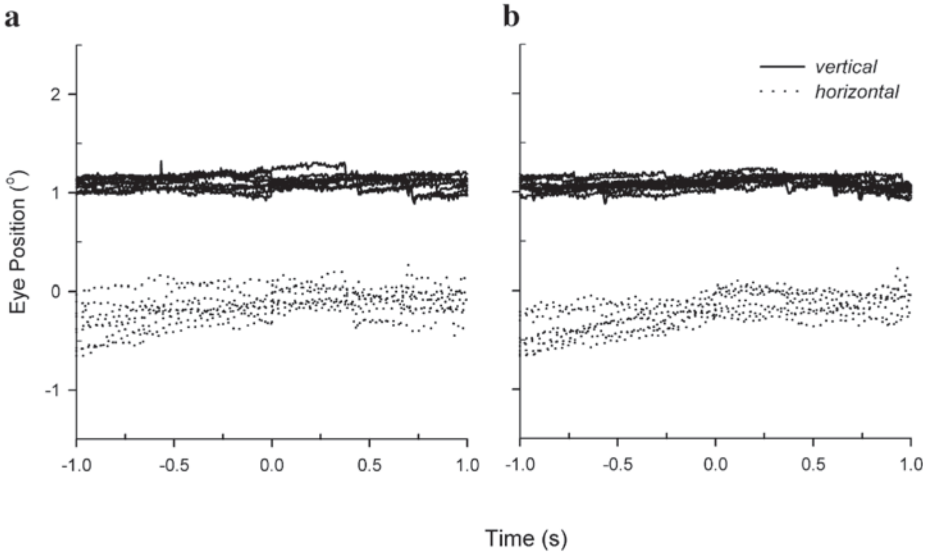

FIG. 7. Horizontal and vertical eye positions while viewing Rubin's vase–face stimulus. Switches in perception from (a) *vase-to-faces* or (b) *faces-to-vase* were aligned to 0 on the abscissa. Following the report of a perceptual switch, an embossed image was presented (see Experiment 2). The graphs show that changes in eye position were not necessary to effect a switch in perception, nor were they triggered by a switch when viewing the ambiguous figure. Each line represents the horizontal or vertical eye position on an individual trial and eye position is plotted relative to a calibration grid.

et al., 1998; O'Craven *et al.,* 1999) or when a "Mooney" image of a face (a difficult-to-recongnize, high-contrast version) becomes recognizable after the subject views a photographic version of the same image (Tovee *et al.,* 1996; Dolan *et al.,* 1997). More generally, it could be that this area is involved directly in the awareness of a broader range of specialized object categories (Kanwisher, 2000; Tarr and Gauthier, 2000).

Together these results strongly suggest that activity in the fusiform gyrus "face area" reflects the perceived rather than merely the retinal stimulus. We show that on a trial-to-trial basis, the MR activity was statistically predictive of the subjects' responses. Using other ambiguous figures in which different stimulus representations compete for perceptual dominance might allow us to make similar inferences about the role of other sensory areas in the resolution of perceptual uncertainty.

ACKNOWLEDGMENTS

We thank Caroline Andrews, Ileana Antonopoulou, Peter Hobden, Dave Flitney, Eugene McSorley, and Steve Smith for their for help during the course of the study. We are also grateful to Steve Engel, Dale Purves, and two anonymous reviewers for providing helpful criticism of the manuscript. Functional imaging was carried out at the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB). D.S. held studentships from the Oxford McDonnell–Pew Centre for Cognitive Neuroscience and the Studienstiftung des Deutschen Volkes and the work was supported by grants from the Medical Research Council and the Oxford McDonnell–Pew Centre for Cognitive Neuroscience. Parts of this work have already been published in abstract form (Andrews *et al.,* 2000).

REFERENCES

Allison, T., Puce, A., Spencer, D. D., and McCarthy, G. 1999. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* **9:** 415–430.

Allison, T., Puce, A., and McCarthy, G. 2000. Social perception from visual cues: Role of the STS region. *Trends Cog. Sci.* **4:** 267–278.

Andrews, T. J. 2001. Binocular rivalry and visual awareness. *Trends Cog. Sci.* **5:** 407–409.

Andrews, T. J., Schluppeck, D., and Blakemore, C. 2000. Ambiguous figures and neural correlates of perceptual awareness in human visual cortex. *Soc. Neurosci. Abstr.* **686:** 9.

Andrews, T. J., and Purves, D. 1997. Similarities in normal and binocularly rivalrous viewing. *Proc. Natl. Acad. Sci. USA* **94:** 9905–9908.

Blake, R. 1989. A neural theory of binocular rivalry. *Psychol. Rev.* **96:** 145–167.

Blake, R., and Logothetis, N. K. 2002. Visual competition. *Nature Rev. Neurosci.* **3:** 1–11.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., and Movshon, J. A. 1996. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13:** 87–100.

Calvert, G. A., Bullmore, E., Brammer, M. J., Campbell, R., Iversen, S. D., Woodruff, P., McGuire, P., Williams, S., and David, A. S. 1997. Silent lipreading activates the auditory cortex. *Science* **276:** 593–596.

Crane, H. D., and Steele, C. M. 1985. Generation-V dual Purkinje-image eyetracker. *Appl Optics* **24:** 527–537.

Crick, F., and Koch, C. 1995. Are we aware of neural activity in primary visual cortex? *Nature* **375:** 121–123.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., and Friston, K. J. 1997. How the brain learns to see objects and faces in an impoverished context. *Nature* **389:** 596–599.

Dodd, J. V., Krug, K., Cumming, B. G., and Parker, A. J. 2001. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J. Neurosci.* **21:** 4809–4821.

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

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. 1995. Improved assessment of significant activation in functional magnetic resonance imaging (FMRI): Use of a cluster-size threshold. *Magn. Reson. Med.* **33:** 636–647.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., and Frackowiak, R. S. J. 1995. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Map.* **2:** 189–210.

Green, D. M., and Swets, J. A. 1974. *Signal Detection Theory and Psychophysics.* Krieger, Melbourne, FL.

Grill-Spector, K., Kushnir, T., Edelman, S., Galia, A., Itchak, Y., and Malach, R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* **24:** 187–203.

Hasselmo, M. E., Rolls, E. T., and Baylis, G. C. 1989. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of monkey. *Behav. Brain Res.* **32:** 203–208.

Hasson, U., Hendler, T., Bashat, D. B., and Malach, R. 2001. Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *J. Cog. Neurosci.* **13:** 744–753.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., and Grady, C. L. 1994. The functional organisation of the human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. *J. Neurosci.* **14:** 6336–6353.

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. 2000. The distributed human neural system for face perception. *Trends Cog. Sci.* **4:** 223–233.

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

Heywood, C. A., and Cowey, A. 1992. The role of the 'face-cell' area in the discrimination and recognition of faces by monkeys. *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* **335:** 31–37.

Hoffman, E. A., and Haxby, J. V. 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neurosci.* **3:** 80–84.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., and Haxby, J. V. 1999. Distributed representation of objects in the human ventral visual pathway. *Proc. Natl. Acad. Sci. USA* **96:** 9379–9384.

Jenkinson, M., and Smith, S. M. 2000. *A Global Optimisation Method for Robust Affine Registration of Brain Images.* Technical Report.

Kanwisher, N. 2000. Domain specificity in face perception. *Nat. Neurosci.* **3:** 759–763.

Kanwisher, N., McDermott, J., and Chun, M. M. 1997. The fusiform face area: A module in extrastriate cortex specialized for face perception. *J. Neurosci.* **17:** 4302–4311.

Kleinschmidt, A., Buchel, C., Zeki, S., and Frackowiak, R. S. 1998. Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc. R. Soc. London B Biol. Sci.* **265:** 2427.

Leopold, D. A., and Logothetis, N. K. 1996. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* **379:** 549–553.

Logothetis, N. K., Leopold, D. A., and Sheinberg, D. L. 1996. What is rivalling during binocular rivalry? *Nature* **380:** 621–624.

Malach, R., Reppas, J. B., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., and Tootell, R. B. H. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. USA* **92:** 8135–8138.

Manly, B. F. J. 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology,* 2nd ed. Chapman and Hall, London.

Necker, L. A. 1832. Observations on some remarkable phenomena seen in switzerland and optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *London Edinburgh Philos. Mag. J. Sci.* **1:** 329–337.

O'Craven, K. M., Downing, P. E., and Kanwisher, N. 1999. FMRI evidence for objects as the units of attentional selection. *Nature* **401:** 584–587.

Parker, A. J., and Newsome, W. T. 1988. Sense and the single neuron: Probing the physiology of perception. *Annu. Rev. Neurosci.* **21:** 227–277.

Perrett, D. I., and Mistlin, A. J. 1990. Perception of facial characteristics by monkeys. In *Comparative Perception, Complex Signals* (W. C. Stebbins and M. A. Berkley, Eds.), pp. 187–215. Wiley, New York.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., and Jeeves, M. A. 1985. Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proc. R. Soc. London B* **223:** 293–317.

Polonsky, A., Blake, R., Braun, J., and Heeger, D. J. 2000. Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neurosci.* **3:** 1153–1159.

Puce, A., Allison, T., Gore, J. C., and McCarthy, G. 1995. Face-sensitive regions in human extra-striate cortex studied by functional MRI. *J. Neurophysiol.* **74:** 1192–1199.

Purves, D., and Andrews, T. J. 1997. The perception of transparent three-dimensional objects. *Proc. Natl. Acad. Sci. USA* **94:** 6517–6522.

Rubin, E. 1915. *Visuell Wahrgenommene Figuren.* Gyldenalske Boghandel, Copenhagen.

Sengpiel, F., and Blakemore, C. 1994. Interocular control of neuronal responsiveness in cat visual cortex. *Nature* **368:** 847–850.

Sergent, J., Ohta, S., and MacDonald, B. 1992. Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain* **115:** 15–36.

Sheinberg, D. L., and Logothetis, N. K. 1997. The role of temporal cortical areas in perceptual organization. *Proc. Natl. Acad. Sci. USA* **94:** 3408–3416.

Smith, S. 2000. Robust automated brain extraction. In *Sixth Int. Conf. on Functional Mapping of the Human Brain,* p. 610. **AQ: 1**

Tanaka, K. 1997. Mechanisms of visual object recognition: Monkey and human studies. *Curr. Opin. Neurobiol.* **7:** 523–529.

Tarr, M. J., and Gauthier, I. 2000. FFA: A flexible fusiform area for subordinate visual processing automatized by expertise. *Nature Neurosci.* **3:** 764–769.

Tong, F., Nakayama, K., Vaughan, J., and Kanwisher, N. 1998. Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* **21:** 753–759.

Tong, F., and Engel, S. 2001. Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* **411:** 95–199.

Tovee, M. J., Rolls, E. T., and Ramachandran, V. S. 1996. Visual learning in neurons of the primate temporal visual cortex. *Neuroreport* **7:** 2757–2760.

Wojciulik, E., Kanwisher, N., and Driver, J. 1998. Covert visual attention modulates face-specific activity in the human fusiform gyrus: FMRI study. *J. Neurophysiol.* **79:** 1574–1578.

Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., and Mazziotta, J. C. 1998. Automated image registration: 1. General methods and intrasubject, intramodality validation. *J. Comput. Assist. Tomogr.* **22:** 141–154.

Woolrich, M., Ripley, B., Brady, J., and Smith, S. 2000. Nonparametric estimation of temporal autocorrelation in fMRI. In *Sixth Int. Conf. on Functional Mapping of the Human Brain,* p. 610. **AQ: 2**