Representation of the Perceived 3-D Object Shape in the Human Lateral Occipital Complex

Zoe Kourtzi, Michael Erb[1], Wolfgang Grodd[1] and Heinrich H. Bülthoff

Max Planck Institute for Biological Cybernetics and [1]University Clinics, Tübingen, Germany

We used human functional magnetic resonance imaging (fMRI) to test whether the human lateral occipital complex (LOC), an area known to be involved in the analysis of visual shape, represents the perceived 3-D shape of objects or simply their 2-D contours. We employed an fMRI adaptation paradigm, in which repeated presentation of a stimulus results in decreased responses compared to responses to different stimuli. We found adaptation in the LOC for images of objects with the same perceived 3-D shape structure but different 2-D contours that resulted from small rotations of the objects in the frontal plane or in depth. However, no adaptation was observed in the LOC for images of objects that had the same 2-D contours but differed in their perceived 3-D shape; namely, 2-D silhouettes versus 3-D shaded images of objects, or convex versus concave objects. Differences in the fMRI adaptation responses across subregions in the LOC suggest that different neural populations in the LOC may mediate different mechanisms for the processing of object features.

Introduction

Successful interactions with our world entail that we use information about the 3-D structure of objects to recognize them and guide our actions promptly and accurately towards them. This is an especially complex task for the visual brain, given that the visual input is registered only in two dimensions on our retina. The present study used human functional magnetic resonance imaging (fMRI) to ask whether cortical regions involved in shape analysis represent the perceived 3-D object shape or simply the 2-D image contours.

Specifically, several human imaging studies have shown that the lateral occipital complex (LOC), a region in the lateral occipital cortex extending anterior in the temporal cortex, is involved in shape processing (Malach et al., 1995; Kanwisher et al., 1996; Grill-Spector et al., 2001; Haxby et al., 2001). Recent studies have suggested that the LOC represents the perceived object shape rather than low-level image features (Grill-Spector et al., 1998, 1999; Malach et al., 1998; Kourtzi and Kanwisher, 2001; Avidan et al., 2002a) and may mediate processes of object recognition (Grill-Spector et al., 2000; Bar et al., 2001).

But how abstract are the shape representations in the LOC? There is accumulating evidence that not only information about 2-D shapes, but also about 3-D objects is processed in the LOC (Mendola et al., 1999; Kourtzi and Kanwisher, 2000; Gilaie-Dotan et al., 2001; Moore and Engel, 2001; Kourtzi et al., 2002). What are the neural mechanisms underlying processing of 3-D objects in the LOC? Are the same or different neural populations in the LOC involved in the processing of 2-D and 3-D shape information of objects? Are objects with different perceived 3-D shape structure (i.e. convex versus concave) represented by different neural populations in the LOC?

We used human fMRI to address these questions. We employed an adaptation paradigm in which lower responses are observed for stimuli that have been presented repeatedly than for different stimuli (Miller et al., 1991, 1996; Buckner et al., 1998; Malach et al., 1998; Wiggs and Martin, 1998; James et al., 1999). Because neural adaptation depends critically on the sameness of two stimuli, fMRI adaptation paradigms provide a sensitive tool that allows us to overcome the spatial resolution limitations of conventional fMRI paradigms and discern neural populations selective for different stimulus dimensions (Grill-Spector and Malach, 2001; Avidan et al., 2002b).

In particular, experiment 1 tested for adaptation in the LOC between 2-D silhouettes and 3-D shaded images of objects. Experiment 2 tested for adaptation in the LOC across changes in the 3-D shape structure of objects (convex versus concave stimuli) compared to changes in their 2-D contours but not the 3-D shape (30° rotated images of objects in the frontal plane or in depth). We showed adaptation for stimuli with the same but not with different perceived 3-D shape structure in anterior rather than posterior regions of the LOC, suggesting that different neural populations in the LOC may mediate different mechanisms for the processing of object features.

Materials and Methods

Subjects

Ten students from the University of Tübingen participated in this study. Each subject participated in the two experiments in the same scanning session. The data from two subjects were excluded due to excessive head movement.

Visual Stimuli

The stimuli used in all the experiments were 300 × 300 pixel images. For the LOC localizer scans we used grayscale images and line drawings of novel and familiar objects, as well as scrambled versions of each set. The scrambled images were created by dividing the intact images in a 20 × 20 square grid and by scrambling the positions of each of the resulting squares. The grid lines were present in both the intact and the scrambled images.

In experiment 1, a set of 18 novel objects was used. As shown in Figure 1a, for each object we rendered a 2-D silhouette, a stereo image, i.e a silhouette that was rendered as a red-green anaglyph and appeared in front of the fixation plane through red-green glasses (disparity of 0.2°) and a shaded 3-D image. In experiment 2, we used a set of 32 shaded novel objects different from those used in experiment 1. As shown in Figure 1b, for each object we rendered a convex and a concave shaded image that had the same 2-D contours, a 2-D rotated image that differed from the original image in a ±30° frontal plane rotation and a 3-D rotated image that differed from the original image in a ±30° rotation in depth about the vertical axis. For all objects, the 3-D rotation resulted in views where the same object parts were visible. The use of large sets of different stimuli in the two experiments ensured that the effects observed would be due to differences among the experimental conditions rather than the low-level features of the stimuli.

Procedure and Design

Each subject was run in one session consisting of ten scans: two LOC localizer scans and four event-related adaptation scans for each one of the
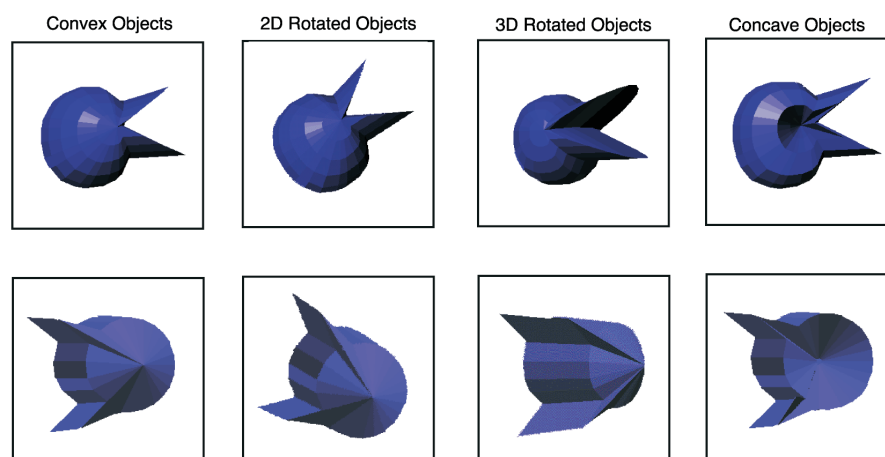
a. Experiment 1
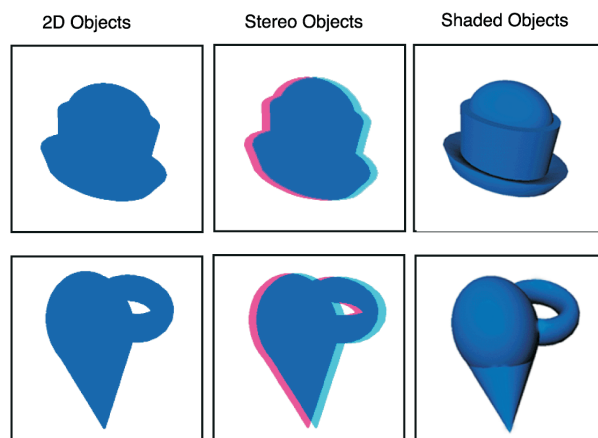
b. Experiment 2

Figure 1. Stimuli. Examples of the stimuli used in: (*a*) experiment 1: 2-D silhouettes, stereo images depicted as red–green anaglyphs and shaded images of two objects; and (*b*) experiment 2: convex objects, rotated images in the frontal plane (30°), rotated images in depth (30°) and concave images of two objects.

two experiments. The order of the scans was counterbalanced across subjects.

For the LOC localizer scans we used a blocked design with sixteen 16 s stimulus epochs and interleaved fixation periods as described in previous studies (Kourtzi and Kanwisher, 2000, 2001). Twenty different images of the same type were presented in each epoch (Fig. 2*a*). Each image was presented for 250 ms with a blank interval of 550 ms between images. Each of the four stimulus types (grayscale images and line drawings of objects, as well as scrambled versions of each set) were presented in different epochs within each scan, in a design that balanced for the order of conditions. The subjects performed a one-back matching task; i.e. they were instructed to press a button whenever they saw two identical pictures in a row. Two or more consecutive repetitions occurred in each epoch. This one-back matching task was used to engage observers' attention on all the stimulus conditions used.

Experiments 1 and 2 used an event-related (Buckner *et al.*, 1998; Rosen *et al.*, 1998) adaptation paradigm as described in previous studies (Kourtzi and Kanwisher, 2000, 2001). The event-related scans consisted of one epoch of experimental trials and two 8 s fixation epochs, one at the beginning and one at the end of the scan. Each scan consisted of 18 experimental trials for each of the five conditions and 18 fixation trials. A new trial began every 3 s and consisted of a pair of images presented sequentially. Each image was presented for 300 ms with a blank interval of 400 ms between images. A blank interval of 2 s was introduced

between trials. Previous human fMRI (Kourtzi and Kanwisher, 2000, 2001) and neurophysiological studies (Lisberger and Movshon, 1999; Mueller *et al.*, 1999) have reported adaptation effects for short stimulus presentations similar to those used in the current study. Specific stimuli (i.e. the different image types rendered for each object) were counterbalanced across conditions such that each specific stimulus appeared equally often in each condition. Thus, differences in the response between conditions were due to the relationship between the two images in a trial and not due to the stimuli used in each condition. As in previous studies (Kourtzi and Kanwisher, 2000, 2001), the order of presentation was counterbalanced so that trials from each condition, including the fixation condition, were preceded equally often by trials from each of the other conditions. Subjects were instructed to passively view the images while fixating.

Five conditions that were defined by the type of stimuli presented in each trial were tested in experiment 1. In the 'identical' condition, the same image (i.e. 2-D, stereo or shaded object) was presented twice, while in the 'different' condition, images (i.e. 2-D, stereo or shaded object) of two different objects were presented. To investigate the representation of 2-D versus 3-D shapes in the LOC, we tested conditions where the stimuli in each trial differed in their 3-D information. In particular, we tested responses to (i) images of the same object presented at different depth planes in a trial (2-D versus stereo object) and (ii) images of the same object with different 3-D shape structure (2-D versus shaded objects and

a. Stimuli for the localization of the LOC
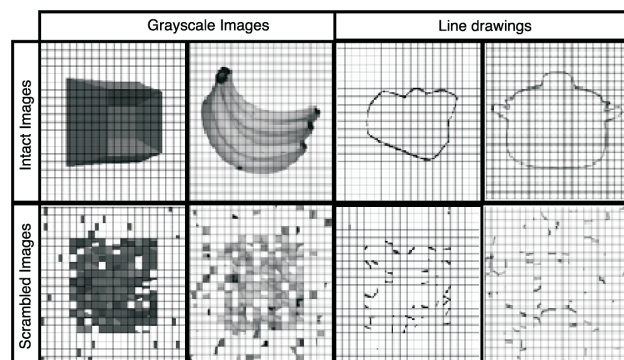
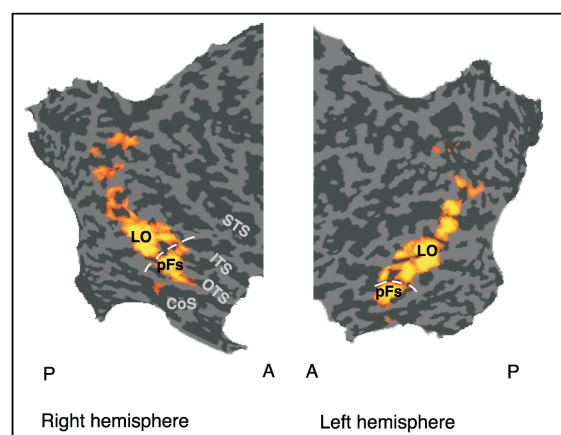## b. Functional activation maps of the LOC

Figure 2. Functional localization of the LOC. (*a*) Stimuli for the localization of the LOC: examples of intact and scrambled grayscale images and line drawings of familiar and novel objects used to localize the LOC in each subject individually. (*b*) Functional activation maps for one subject showing the LOC. The functional activations are superimposed on flattened cortical surfaces of the right and left hemispheres. The sulci are coded in darker gray than the gyri and the anterior–posterior orientation is noted by A and P. Major sulci are labeled: STS, superior temporal sulcus; ITS, inferior temporal sulcus; OTS, occipitotemporal sulcus; CoS, collateral sulcus. The LOC was defined as the set of all voxels in the ventral occipitotemporal cortex that were activated more strongly ($P < 10^{-4}$) by intact than by scrambled images of objects. The posterior (LO: mean Talairach coordinates for right hemispheres 37.9, −69.4, −8.1 and left hemispheres −38.1, −72.2, −6.7) and (LO) and anterior (pFs: mean Talairach coordinates for right hemispheres 31.5, −46.8, −15.8 and left hemispheres −35.2, −48.9, −15.4) subregions of the LOC were identified. LO was located at the posterior part of the inferior-temporal sulcus, while pFs in the posterior fusiform gyrus. The white dotted lines indicate the border between the two subregions.

stereo versus shaded object) presented in a trial. The 2-D images consisted of silhouettes and the stereo objects of 2-D silhouettes of objects that appeared stereoscopically in the front plane. The silhouette stimuli were presented first in a trail and were followed by the shaded or stereoscopic stimuli, to avoid possible 3-D interpretation of the silhouettes that may occur after presentation of the shaded or stereoscopic images of the same objects.

Similarly, experiment 2 consisted of five conditions. In the identical condition, the same shaded 3-D object was presented twice, while in the different condition, two different shaded 3-D objects were presented. To investigate the representation of 3-D shapes in the LOC, we tested conditions where the stimuli in each trial differed or not in their perceived 3-D shape. Two conditions where the stimuli in each trial had the same perceived 3-D shape structure were tested: (i) 3-D rotation and (ii) 2-D rotation condition. In the 3-D rotation condition, the two images

of a 3-D object differed by a 30° rotation in depth, while in the 2-D rotation condition, the two images of a 3-D object differed by 30° rotation in the frontal plane. In comparison, an additional condition was tested in which the stimuli in a trial differed in their 3-D shape, but not their contours. This condition consisted of a convex and a concave image of the same 3-D shaded object presented in the same trial (convex versus concave object condition). Different adaptation effects across conditions could not be simply attributed to the low-level differences between the stimuli in each trial, because all the image manipulations tested (rotations, or curvature changes) resulted in image changes, but only the change in the curvature direction (convex versus concave) resulted in differences in the perceived 3-D shape of objects.

One could suggest that any adaptation effect, namely decreased signal for identical images, would indicate reduced attention to repeated stimuli compared to different images. The event-related design used in this experiment, with randomized presentation of trials from all conditions, prevents observers from anticipating the condition of the upcoming trial. Moreover, a previous experiment (Kourtzi and Kanwisher, 2001) showed similar shape adaptation effects for passive viewing and a matching task that engaged the subjects' attention similarly across trials where the identical or different stimuli were presented.

MRI Acquisition
For all experiments, scanning was carried out on the 1.5 T Siemens scanner at the University Clinics in Tübingen, Germany. A gradient echo pulse sequence (T_R = 2 s for the localizer scans; T_R = 1 s for the event-related scans; T_E = 40 ms) was used. Eleven near-coronal slices (parallel to the brainstem, 5 mm thick with 3.00 × 3.00 mm in-plane resolution) were collected with a head coil.

Data Analysis
fMRI data were processed using the BrainVoyager 4.6 software package. Preprocessing of all the functional data included head movement correction and removal of linear trends. The 2-D functional images were aligned to 3-D anatomical data and the complete data set was transformed to Talairach coordinates, inflated, unfolded and flattened.

For each individual subject, the LOC (Talairach coordinates: mean for right hemispheres 39.1, −65.6, −12.0 and left hemispheres −41.9, −64.8, −2.7) was identified as the set of all voxels in the ventral occipitotemporal cortex that were activated more strongly ($P < 10^{-4}$) by intact than by scrambled images of objects presented in the two localizer scans, as described previously (Kourtzi and Kanwisher, 2000, 2001) (Fig. 2). 3-D statistical maps were calculated by correlating the signal time course with a reference function for each voxel based on the hemodynamic response properties (Boynton *et al.*, 1996; Cohen *et al.*, 1997; Dale and Buckner, 1997). Also, two subregions of the LOC were identified on the functional maps, as in previous studies (Grill-Spector *et al.*, 2000): the LO (lateral occipital) at the posterior part of the inferior-temporal sulcus and the pFs (posterior fusiform) in the posterior fusiform gyrus. The magnitude of the response in this ROI was then measured for each subject in each condition in the two event-related experiments.

Specifically, for each event-related scan, the fMRI response was extracted by averaging the data from all the voxels within the LOC. First, for each scan we calculated the signal intensity for each trial in each condition and then we averaged the signal across trials in each condition at each of 11 corresponding time points (seconds) and converted these time courses to percentage signal change relative to the fixation trials, as described previously (Kourtzi and Kanwisher, 2000, 2001). We then averaged the time courses for each condition across scans for each subject and then across subjects. Because of the hemodynamic lag in the blood oxygen level dependent (BOLD) fMRI response, the peak in overall response and, therefore, the differences across conditions are expected to occur at a lag of several seconds after stimulus onset (Boynton *et al.*, 1996; Cohen *et al.*, 1997; Dale and Buckner, 1997). To find the latencies where any adaptation effects occurred, we conducted an analysis of variance (ANOVA) for each experiment with factors of condition (identical, different) and time point (measurements made at latencies of 0–10 s after trial onset). We observed significant main effects of condition {experiment 1 [$F(1,77)$ = 51.66, $P < 0.001$]; experiment 2 [$F(1,77)$ = 10.15, $P < 0.001$]} and time {experiment 1 [$F(10,77)$ = 21.19, $P < 0.001$]; experiment 2 [$F(10,77)$ = 7.19, $P < 0.01$]}. A significant interaction

between condition and time (11 time points) was observed for each experiment {experiment 1 [$F_{(10,77)}$ = 2.05, P < 0.05]; experiment 2 [$F_{(10,231)}$ = 2.52, P < 0.05]}. This analysis verified that adaptation occurred and that it varied with latency. Follow-up contrast analyses run separately on each time point tested for a significantly lower response for the identical compared to the different conditions. This adaptation effect was found only for time point 4 {experiment 1 [$F_{(1,77)}$ = 30.73, P < 0.001]; experiment 2 [$F_{(1,77)}$ = 3.02, P = 0.05]}, time point 5 {experiment 1 [$F_{(1,77)}$ = 27.03, P = 0.001]; experiment 2 [$F_{(1,77)}$ = 4.42, P = 0.05]} and time point 6 {experiment 1 [$F_{(1,77)}$ = 13.06, P < 0.01]; experiment 2 [$F_{(1,77)}$ = 12.98, P < 0.01]}, but not for the onset of a trial, i.e. time point 0 {experiment 1 [$F_{(1,77)}$ = 2.17, P = 0.18]; experiment 2 [$F_{(1,77)}$ = 1.1, P = 0.33]}. The average of the response at time points 4, 5 and 6 was therefore taken as the measure of response magnitude for each condition in subsequent analyses.

The identical condition, where the two stimuli in a pair were identical and the different condition, where the two stimuli were different objects, provided the upper and lower reference points to which we could compare the response for the critical conditions in each experiment.

Results

Experiment 1: Adaptation between 2-D and 3-D Images

As shown in Figure 3, adaptation was observed for identical images of objects presented repeatedly, as reported previously (Kourtzi and Kanwisher, 2000, 2001). Specifically, a repeated ANOVA on the mean signal of the LOC at the peak of the hemodynamic response (4–6 s after trial onset) across conditions showed a main effect [$F_{(4,92)}$ = 16.92, P < 0.001] for condition (identical, different, 2-D versus stereo object, 2-D versus shaded object, stereo versus shaded object). Subsequent contrast analysis showed that responses at the peak of the trials were significantly lower [$F_{(1,92)}$ = 36.24, P < 0.001] for the identical condition than for the different condition.

More importantly, in the critical conditions, adaptation was observed between 2-D and stereo images of the same object, but not between images of objects that differed in their 3-D shape. In particular, the peak responses in the 2-D versus stereo object condition were significantly lower [$F_{(1,92)}$ = 18.21, P < 0.001] than the responses in the different condition but not significantly different [$F_{(1,92)}$ = 3.07, P = 0.08] from the responses in the identical condition. However, no adaptation was observed between 2-D silhouettes of objects and their shaded 3-D images. That is, the responses in the 2-D versus shaded object condition were significantly higher [$F_{(1,92)}$ = 27.03, P < 0.001] than the responses in the identical condition, but not significantly different from the responses in the different condition [$F_{(1,92)}$ < 1, P = 0.67]. Moreover, no adaptation was observed between stereoscopically defined silhouettes of objects and shaded 3-D images of the same objects. That is, the responses in the stereo versus shaded object condition were significantly higher [$F_{(1,92)}$ = 44.88, P < 0.001] than the responses in the identical condition, but not significantly different from the responses in the different condition [$F_{(1,92)}$ < 1, P = 0.49].

Finally, a repeated measures ANOVA on the critical conditions for this experiment (2-D versus stereo object, 2-D versus shaded object, stereo versus shaded object) showed significant differences across these conditions [$F_{(2,46)}$ = 10.47, P < 0.001]. Specifically, responses to the 2-D versus stereo object condition were significantly lower [$F_{(1,46)}$ = 19.12, P < 0.001] than responses to the 2-D versus shaded object and stereo versus shaded object conditions. However, responses to the 2-D versus shaded object condition were not significantly different [$F_{(1,46)}$ = 1.83, P = 0.18] from responses to the stereo versus shaded object condition. These findings are summarized in Figure 5a, where we plot an adaptation index for all conditions. This index

was calculated by dividing the percentage signal change for each condition by the percentage signal change for the identical condition. A ratio of 1 indicates adaptation, as in the 2-D versus stereo object condition [$F_{(1,92)}$ = 4.21, P = 0.07], while responses significantly higher than 1 indicate no adaptation, as in the 2-D versus shaded object and stereo versus shaded object conditions [$F_{(1,92)}$ = 33.74, P < 0.001].

In sum, these results showed adaptation between 2-D and stereoscopically presented silhouettes that differ in the cues that define their 2-D contours. This finding is consistent with previous studies showing responses in the LOC to images of objects independent of the cues that define them (Grill-Spector *et al.*, 1998; Mendola *et al.*, 1999). However, no adaptation was observed between 2-D silhouettes and 3-D shaded images of objects that differ in their perceived 3-D shape structure. Thus, these findings suggest that different neural populations in the LOC may represent 2-D and 3-D images of objects. Alternatively, given the ambiguous 3-D nature of silhouettes (Sinha and Poggio, 1996; Moore and Engel, 2001), it is possible that the 2-D silhouettes and 3-D shaded images were perceived by the observers and encoded by the neural populations in the LOC as different objects. Therefore, the results of this experiment may only suggest that the sameness of the perceived 3-D shape is necessary for adaptation in the LOC. Is the sameness of the perceived 3-D shape of objects also sufficient for adaptation in the LOC? Experiment 2 investigated this question by testing adaptation responses to images of objects that had the same 2-D contours but different perceived 3-D shape structure (i.e. convex versus concave objects) compared to objects that had different 2-D contours but the same perceived 3-D shape structure (i.e. rotated objects).

Experiment 2: Adaptation across Changes in the 3-D Shape

As shown in Figure 4, adaptation was observed for identical images. That is, similar to experiment 1, a repeated ANOVA on the mean peak signal in the LOC across conditions showed a main effect [$F_{(4,92)}$ = 7.91, P < 0.001] for condition (identical, 3-D rotation, 2-D rotation and convex versus concave object). Subsequent contrast analysis showed that responses at the peak of the trials were significantly lower [$F_{(1,92)}$ = 13.99, P < 0.001] for the identical condition than for the different condition.

More importantly, in the critical conditions, adaptation was observed between images that shared the same perceived 3-D shape. That is, we observed adaptation when the two stimuli in a trial were images of the same object rotated for 30° in the frontal plane or in depth. In particular, the peak responses in 3-D and 2-D rotation conditions were significantly lower [3-D rotation, $F_{(1,92)}$ = 14.03, P < 0.001; 2-D rotation, $F_{(1,92)}$ = 14.42, P < 0.001] than the responses in the different condition, but not significantly different [3-D rotation, $F_{(1,92)}$ < 1, P = 0.99; 2-D rotation, $F_{(1,92)}$ < 1, P = 0.95] from the responses in the identical condition. However, no adaptation was observed between images of objects that differed in their perceived 3-D shape (convex versus concave images of the same object). Specifically, the responses in the convex versus concave object condition were significantly higher [$F_{(1,92)}$ = 9.29, P < 0.01] than the responses in the identical condition, but not significantly different from the responses in the different condition [$F_{(1,92)}$ < 1, P = 0.49].

A repeated ANOVA showed significant differences across the critical conditions [$F_{(1,46)}$ = 7.14, P < 0.01]. Specifically, responses to the convex versus concave object condition were significantly higher [$F_{(1,46)}$ = 14.28, P < 0.001] than responses
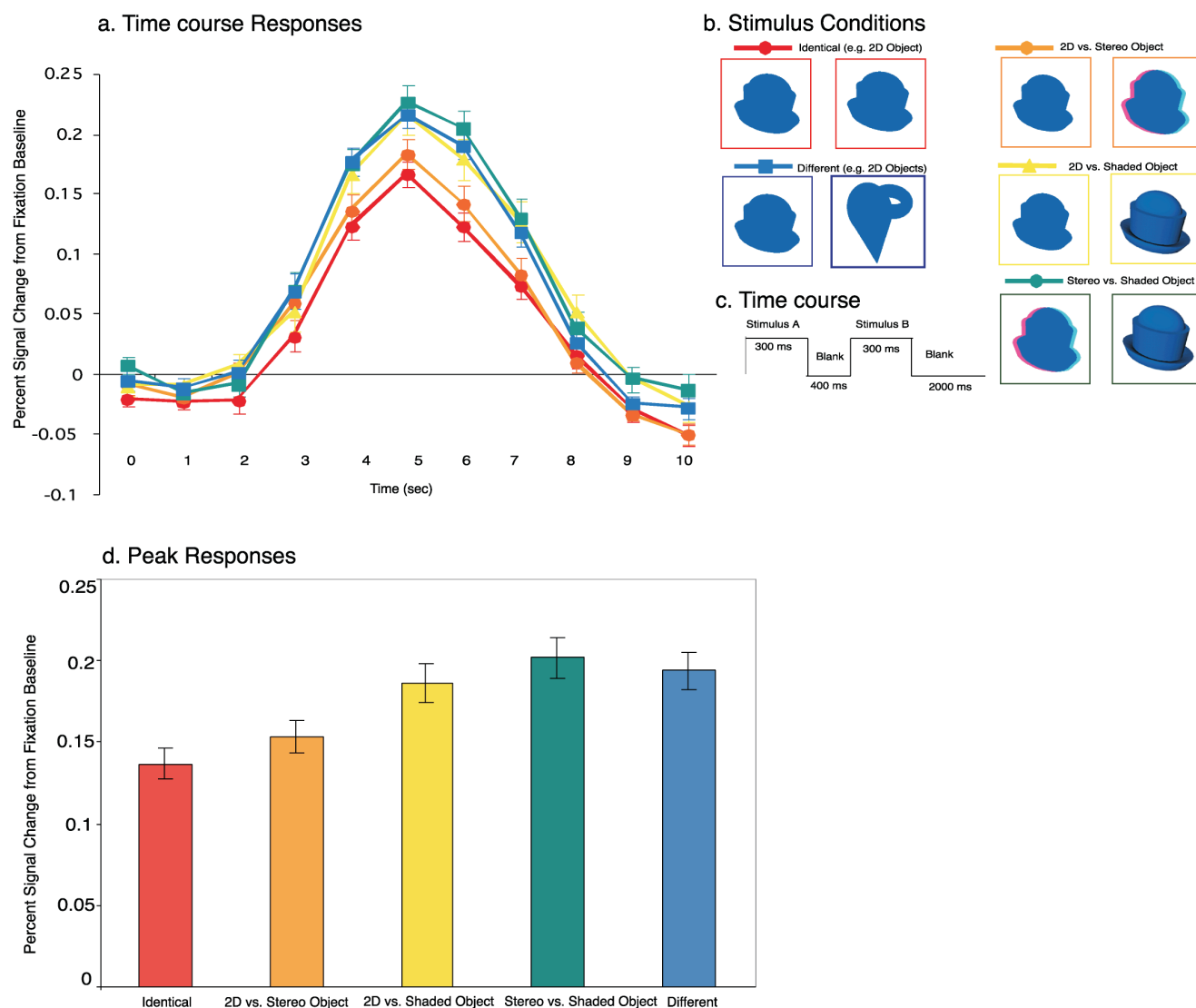
Figure 3. Results for experiment 1. Average percentage signal increases (from the fixation baseline trials) in the LOC across the stimuli (*b*) presented in the time course of a trial (*c*) in experiment 1 plotted for (*a*) 11 time points starting at stimulus onset (time = 0 s) and (*d*) the peak time points (time = 4–6 s). The error bars indicate standard errors on the percentage signal change averaged across scans and subjects.

to the 2-D and 3-D rotation conditions. However, responses to the 2-D rotation condition were not significantly different [$F(1,46) < 1$, $P = 0.95$] from the responses to the 3-D rotation condition. Similar to experiment 1, we summarize these findings in Figure 5*b* by plotting an adaptation index for all conditions. A ratio not significantly different from 1 indicated adaptation in the 2-D and 3-D rotation conditions [$F(1,92) = 3.68$, $P = 0.98$], while responses significantly higher than 1 indicated no adaptation in the convex versus concave object condition [$F(1,92) = 12.09$, $P < 0.001$].

Thus, experiment 2 showed two main findings. First, we observed no adaptation in the LOC for objects that have the same 2-D contours but different perceived 3-D shape structure (convex versus concave objects). In contrast, we observed adaptation in the LOC for images of objects that have the same perceived 3-D shape even when they have different 2-D contours as a result of frontal plane or depth rotation. These results are consistent with previous psychophysical (Bülthoff and Edelman, 1992; Edelman and Bülthoff, 1992), neurophysiological (Logothetis *et al.*, 1994, 1995) and imaging (Grill-Spector *et al.*,

1999; Gauthier *et al.*, 2002; Vuilleumier *et al.*, 2002) studies that provide evidence for generalization across small orientation changes but view-dependent representations across large orientation changes. Taken together, these findings suggest that the sameness of the perceived 3-D shape is sufficient for adaptation in the LOC.

Adaptation Responses in the Anterior and Posterior Regions of the LOC

To further investigate the encoding of 3-D object shape in the LOC, we tested for adaptation responses in two different subregions: a posterior one in the lateral occipital cortex (LO) and an anterior one in the pFs. Figure 5*c,d* shows the adaptation index across conditions in these two regions of interest for each experiment. A ratio of 1 indicates adaptation, while responses significantly higher than 1 indicate no adaptation.

For experiment 1, analysis on adaptation index data in the LO showed adaptation, that is responses not significantly different from 1, for the 2-D versus stereo object [$F(1,92) < 1$, $P = 0.29$], but no adaptation for the 2-D versus shaded object and stereo
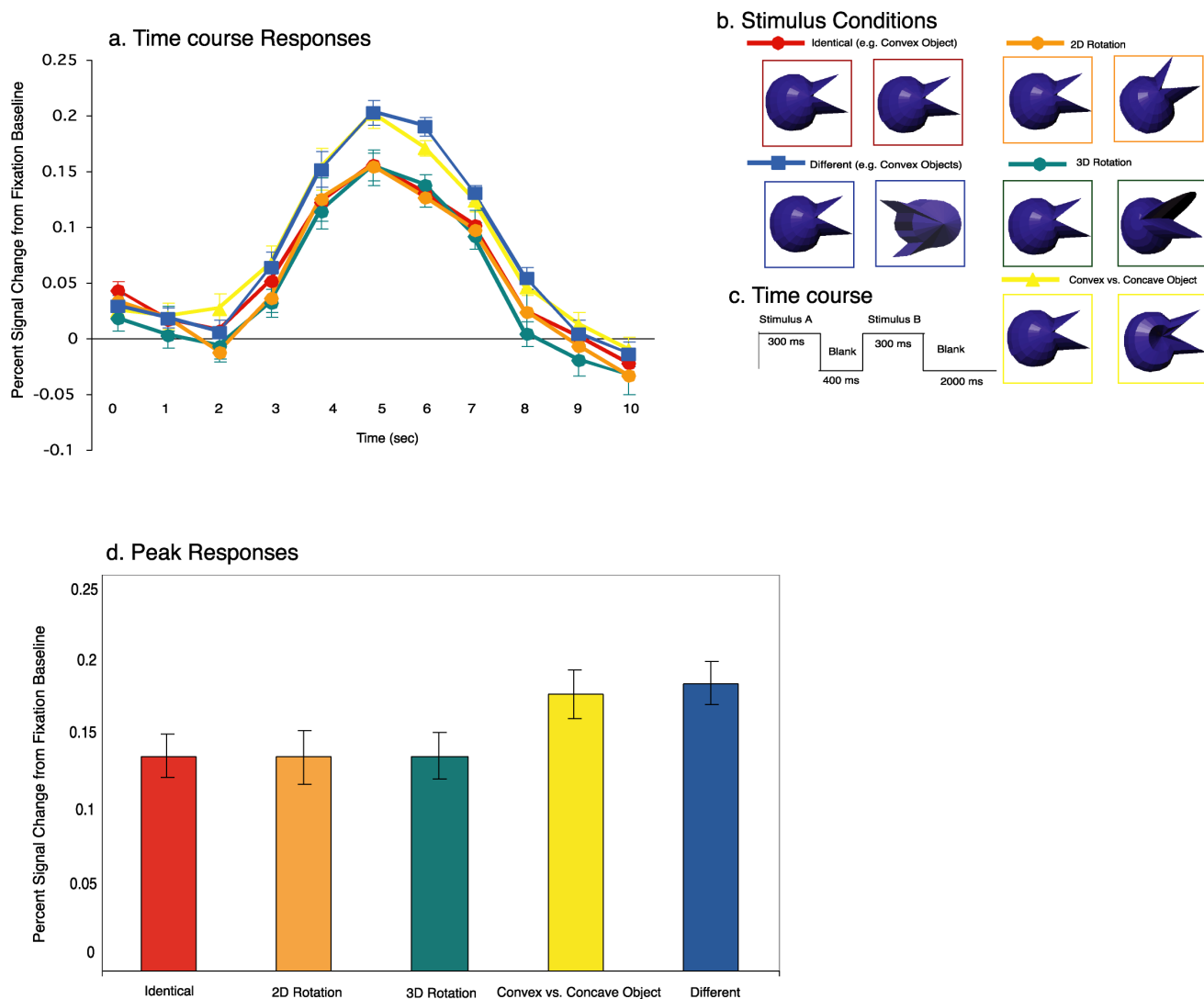
Figure 4. Results for experiment 2. Average percentage signal increases (from the fixation baseline trials) in the LOC across the stimuli (*b*) presented in the time course of a trial (*c*) in experiment 2 plotted for (*a*) 11 time points starting at stimulus onset (time = 0 s) and (*d*) the peak time points (time = 4–6 s). The error bars indicate standard errors on the percentage signal change averaged across scans and subjects.

versus shaded object conditions [$F(1,92) = 18.01$, $P < 0.001$]. Interestingly, the same analysis in pFs showed no adaptation for the 2-D versus stereo object [$F(1,92) = 5.81$, $P < 0.05$] or the 2-D versus shaded object and stereo versus shaded object conditions [$F(1,92) = 5.89$, $P < 0.001$]. These results suggest that in both subregions different neural populations may encode 2-D and 3-D images of objects. The adapted responses observed in the posterior but not the anterior subregion of the LOC for silhouettes of objects presented at different depth planes suggest differences across subregions in the LOC. Specifically, anterior rather than posterior subregions appear to encode information about the depth ordering of objects.

Similar analysis for experiment 2 showed adaptation for the 2-D and 3-D rotation conditions [$F(1,92) < 1$, $P = 0.38$] and the convex versus concave object condition [$F(1,92) < 1$, $P = 0.29$] in the LO. However, the same analysis in pFs showed adapted responses for the 2-D and 3-D rotation conditions [$F(1,92) < 1$, $P = 0.64$], but not the convex versus concave object condition [$F(1,92) = 4.64$, $P < 0.05$]. These results suggest that both anterior and posterior subregions in the LOC appear to encode

3-D images of objects independent of image changes (small 2-D or 3-D rotations) that change the objects' 2-D contours but preserve their 3-D shape. However, anterior rather than the posterior subregions in the LOC appear to encode the perceived 3-D shape of objects (convex versus concave).

Discussion

The LOC has been characterized as the main neural locus of visual shape processing. The current study asked whether neural populations in the LOC encode the perceived 3-D shape of objects or simply their 2-D contours. Using an fMRI adaptation paradigm, we showed that the sameness of the perceived 3-D shape of objects is necessary and sufficient for adaptation in the LOC.

These findings are consistent with recent human fMRI studies that have implicated the LOC in the analysis not only of 2-D but also of 3-D images of objects (Kourtzi and Kanwisher, 2000; Gilaie-Dotan *et al.*, 2001; Moore and Engel, 2001; Kourtzi *et al.*, 2002) independent of the cues that define the object shape (Sary *et al.*, 1993; Mendola *et al.*, 1999; Kourtzi and Kanwisher, 2000).
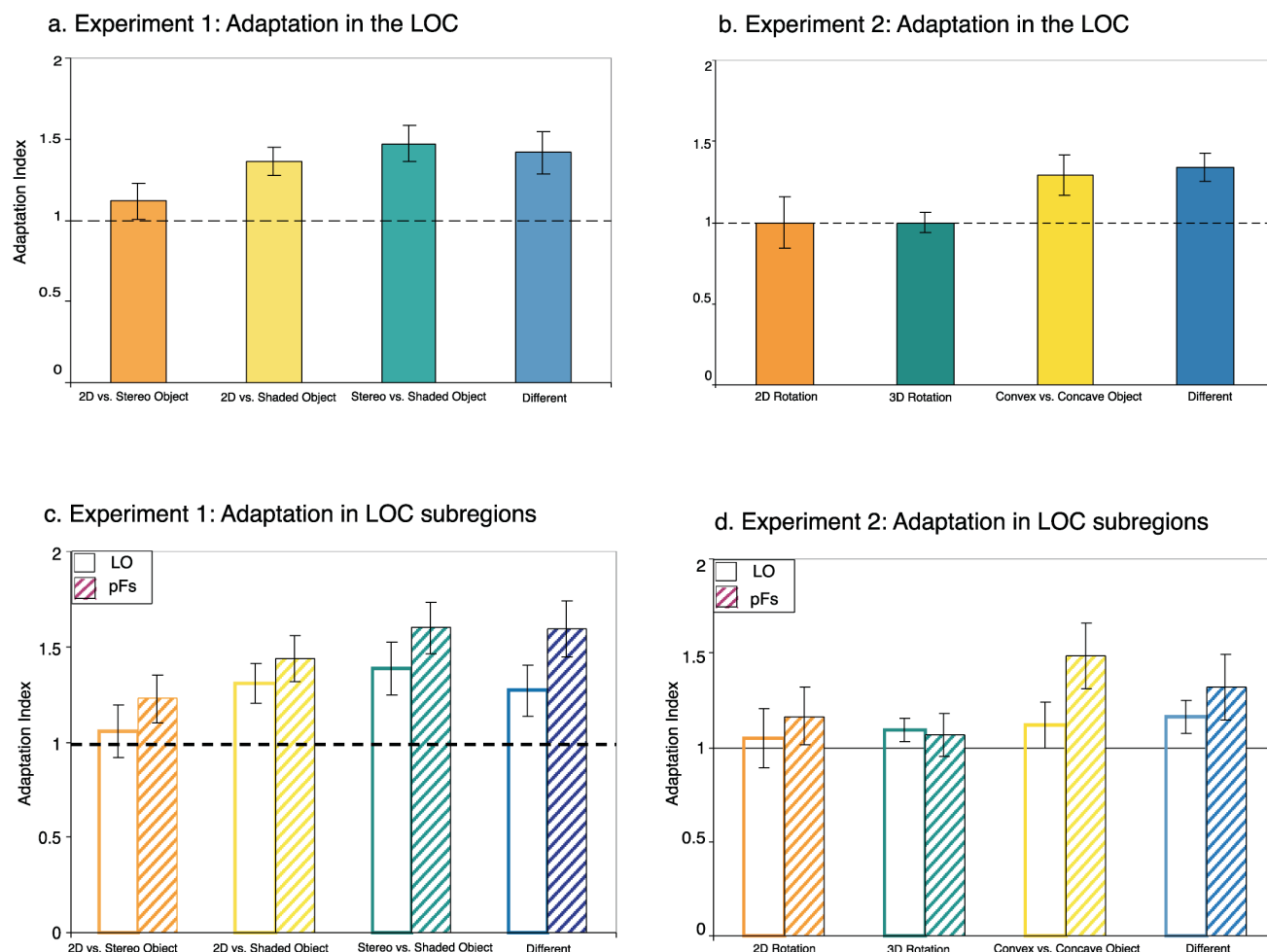
Figure 5. Responses in LOC subregions. Adaptation index plotted across conditions in the LOC for (*a*) experiment 1 and (*b*) experiment 2 and the posterior (LO) and anterior (pFs) subregions of the LOC (*c*) experiment 1 and (*d*) experiment 2. A ratio of 1 (dotted line) indicates adaptation, while responses significantly higher than 1 indicate no adaptation. The error bars indicate standard errors on the percentage signal change averaged across scans and subjects. The mean percentage signal change for each condition in the LOC is plotted in Figures 3 and 4 for experiments 1 and 2, respectively. For the LOC subregions, these means are as follows. For experiment 1: LO—identical = 0.171, different = 0.217, 2-D versus stereo object = 0.180, 2-D versus shaded object = 0.223, stereo versus shaded object = 0.236; pFs—identical = 0.104, different = 0.166, 2-D versus stereo object = 0.128, 2-D versus shaded object = 0.150, stereo versus shaded object = 0.167. For experiment 2: LO—identical = 0.146, different = 0.170, 3-D rotation = 0.163, 2-D rotation = 0.159, convex versus concave object = 0.153; pFs—identical = 0.108, different = 0.143, 3-D rotation = 0.161, 2-D rotation = 0.116, convex versus concave object = 0.126.

Furthermore, recent human (Orban *et al.*, 1999; Paradis *et al.*, 2000; Kourtzi *et al.*, 2002; Murray *et al.*, 2003) and monkey (Sereno *et al.*, 2002; Vanduffel *et al.*, 2002) fMRI studies have proposed a network of visual regions involved in the processing of 3-D objects defined by different cues (i.e. motion, texture, disparity).

Our study advances our understanding of shape processing in the LOC in the following respects. First, by using the adaptation technique we can discriminate between neural populations with different functional properties within the investigated regions. Secondly, we used a monocular (i.e. shading) rather than binocular (i.e. disparity) depth cue to define the 3-D shape structure of the objects. Shading has been shown to contribute to 3-D shape perception (Lekhy and Sejnowski, 1988, 1990; Langer and Bülthoff, 2001; Mamassian and Landy, 2001). Our experiments extend the results of previous studies that have investigated the role of binocular depth cues (i.e. disparity) in 3-D perception. For example, numerous neurophysiological studies have shown that 3-D information about objects defined by disparity is processed not only in early (Poggio and Fischer, 1977; Poggio *et al.*, 1988; Cumming and Parker, 1997, 2000;

Bakin *et al.*, 2000; von der Heydt *et al.*, 2000; Thomas *et al.*, 2002) and dorsal regions, i.e. MT, parietal regions (Maunsell and van Essen, 1983; Roy *et al.*, 1992; Shikata *et al.*, 1996; DeAngelis *et al.*, 1998; DeAngelis and Newsome, 1999; Taira *et al.*, 2000), but also in ventral visual areas, i.e. V4, inferotemporal regions (Janssen *et al.*, 1999, 2000a,b, 2001; Uka *et al.*, 2000; Tanaka *et al.*, 2001; Hinkle and Connor, 2002; Watanabe *et al.*, 2002).

However, a possible problem when studying the representation of shaded 3-D objects is that changes in the shading information result in changes in the luminance and the edges in the image (Fig. 1). For example, in experiment 1 the lack of adaptation between 2-D silhouettes and shaded images of the same objects could be due to differences in the luminance or the edge energy rather than differences in the perceived 3-D object structure. However, our previous studies have shown adaptation for shapes that have the same perceived 3-D structure independent of such low-level image differences (Kourtzi and Kanwisher, 2000). Specifically, we observed adaptation for 3-D line drawings and shaded photographs of the same objects in the LOC. Furthermore, in experiment 2 we observed adaptation effects for images of objects rotated in depth, despite the

luminance and edge differences across these images. These adaptation effects were similar to those observed for images of objects rotated in the frontal plane that did not result in any of these image changes. Finally, the adaptation effect observed for convex versus concave images of objects in the posterior (LO) but not the anterior (pFs) subregion of the LOC suggests differences in the processing of the perceived 3-D structure of objects in different regions rather than a non-specific effect of image luminance across regions.

The Role of Anterior Versus Posterior Subregions of the LOC in 3-D Object Processing

Recent human fMRI studies have characterized functionally posterior (LO) and anterior (pFs) subregions in the LOC. Specifically, anterior subregions have been shown to represent whole completed objects rather than their scrambled parts (Lerner *et al.*, 2001, 2002) and respond more invariantly than posterior ones to position, size and small orientation changes (Malach *et al.*, 1998; Grill-Spector *et al.*, 1999). Also, anterior subregions have been shown to correlate more strongly with performance in depth discrimination and object recognition tasks than posterior subregions (Grill-Spector *et al.*, 2000; Gilaie-Dotan *et al.*, 2001).

Consistent with these studies, our findings show adaptation across small orientation changes in both posterior and anterior subregions. Interestingly, anterior subregions show selective responses to the depth ordering of objects and their 3-D shape (convex versus concave), in contrast to posterior subregions that show adaptation across changes in these dimensions. These findings contradict a simplistic model of increasing neural invariance to image changes from posterior to anterior regions. In contrast, our results suggest that posterior subregions in the LOC may process 2-D features of objects independent of image transformations (small rotations or curvature changes), while anterior subregions may represent the perceived 3-D shape of objects and their depth position in visual scenes. These findings are consistent with neurophysiological studies in monkey inferior temporal cortex (IT) that have shown generalization of responses across orientation changes (Logothetis *et al.*, 1994, 1995; Booth and Rolls, 1998), but not across structural changes of the shape of parts (Vogels *et al.*, 2001), suggesting that neurons in inferotemporal regions may encode complex object features independent of small image transformations. More interestingly, recent studies (Janssen *et al.*, 1999, 2000a,b, 2001) have shown that neural populations in the anterior and medial monkey TE (the anterior part of IT) respond selectively to 3-D convex versus concave objects defined by disparity, compared to lateral TE where neurons are equally selective for 2-D and 3-D images of shapes.

Conclusions

The current experiments show that the posterior subregion of the human LOC may represent 2-D shape features invariantly across small image transformations, while the anterior subregion may encode the perceived 3-D shape of objects. These results provide insights in addressing the long-standing question of whether perception of 3-D objects is based on collections of 2-D images (Poggio and Edelman, 1990; Bülthoff *et al.*, 1995; Perrett *et al.*, 1998; Tarr and Bülthoff, 1998; Ullman, 1998; Riesenhuber and Poggio, 2000) or abstract models/descriptions of the 3-D object structure (Marr and Nishihara, 1978; Biederman, 1987; Hummel and Biederman, 1992). It is possible that neural populations in the posterior LOC mediate shape analysis based on image properties, while neural populations in the anterior

LOC mediate object and scene recognition based on rather abstract 3-D representations. These mechanisms may mediate different tasks that are important for successful interactions in the world. That is, abstract 3-D representations may play an important role when we need quickly to interpret complex scenes and recognize objects independent of image changes. However, image-based processing of shapes may mediate the continuous updating of visual representations that is necessary for guiding our actions promptly and successfully towards objects that may change their position and/or appearance in space and time. The precise functional connections between the LOC subregions involved primarily in visual shape analysis and areas implicated in higher cognitive functions (i.e. guidance of actions) remain to be determined. Addressing this issue is important for understanding the mechanisms mediating functional connectivity in the human visual brain in action.

References

Avidan G, Harel M, Hendler T, Ben-Bashat D, Zohary E, Malach R (2002a) Contrast sensitivity in human visual areas and its relationship to object recognition. J Neurophysiol 87:3102–3116.

Avidan G, Hasson U, Hendler T, Zohary E, Malach R (2002b) Analysis of the neuronal selectivity underlying low fMRI signals. Curr Biol 12:964–972.

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 RBH, 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.

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147.

Booth MC, Rolls ET (1998) View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. Cereb Cortex 8:510–523.

Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 16:4207–4221.

Buckner RL, Goodman J, Burock M, Rotte M, Koutstaal W, Schacter D, Rosen B, Dale AM (1998) Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. Neuron 20:285–296.

Bülthoff HH, Edelman S (1992) Psychophysical support for a two-dimensional view interpolation theory of object recognition. Proc Natl Acad Sci USA 89:60–64.

Bülthoff HH, Edelman S, Tarr MJ (1995) How are three-dimensional objects represented in the brain? Cereb Cortex 5:247–260.

Cohen MS (1997) Parametric analysis of fMRI data using linear systems methods. Neuroimage 6:93–103.

Cumming BG, Parker AJ (1997) Responses of primary visual cortical neurons to binocular disparity without depth perception. Nature 389:280–283.

Cumming BG, Parker AJ (2000) Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. J Neurosci 20:4758–4767.

Dale AM, Buckner RL (1997) Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 5:329–340.

DeAngelis GC, Newsome WT (1999) Organization of disparity-selective neurons in macaque area MT. J Neurosci 19:1398–1415.

DeAngelis GC, Cumming BG, Newsome WT (1998) Cortical area MT and the perception of stereoscopic depth. Nature 394:677–680.

Edelman S, Bülthoff HH (1992) Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. Vision Res 32:2385–2400.

Gauthier I, Hayward WG, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (2002) BOLD activity during mental rotation and viewpoint-dependent object recognition. Neuron 34:161–71.

Gilaie-Dotan S, Ullman S, Kushnir T, Malach R (2001) Shape-selective stereo processing in human object-related visual areas. Hum Brain Mapp 15:67–79.

Grill-Spector K, Malach R (2001) fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol 107:293–321.

Grill-Spector K, Kushnir T, Edelman S, Itzchak Y, Malach R (1998) Cue-invariant activation in object-related areas of the human occipital lobe. Neuron 21:191–202.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex Neuron 24:187–203.

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.

Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vision Res 41:1409–1422.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2330.

Hinkle DA, Connor CE (2002) Three-dimensional orientation tuning in macaque area V4. Nat Neurosci 5:665–670.

Hummel JE, Biederman I (1992) Dynamic binding in a neural network for shape recognition. Psychol Rev 99:480–517.

James TW, Humphrey GK, Gati JS, Menon RS, Goodale MA (1999) Repetition priming and the time course of object recognition: an fMRI study. Neuroreport 10:1019–1023.

Janssen P, Vogels R, Orban GA (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. Proc Natl Acad Sci USA 96:8217–8222.

Janssen P, Vogels R, Orban GA (2000a) Selectivity for 3-D shape that reveals distinct areas within macaque inferior temporal cortex. Science 288:2054–2056.

Janssen P, Vogels R, Orban GA (2000b) Three-dimensional shape coding in inferior temporal cortex. Neuron 27:385–397.

Janssen P, Vogels R, Liu Y, Orban GA (2001) Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. J Neurosci 21:9419–9429.

Kanwisher N, Chun MM, McDermott J, Ledden PJ (1996) Functional imaging of human visual recognition. Cogn Brain Res 5:55–67.

Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. J Neurosci 20:3310–3318.

Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. Science 293:1506–1509.

Kourtzi Z, Bülthoff HH, Erb M, Grodd W (2002) Object-selective responses in the human motion area MT/MST. Nat Neurosci 5:17–18.

Langer MS, Bülthoff HH (2001) A prior for global convexity in local shape-from-shading. Perception 30:403–410.

Lekhy SR, Sejnowski TJ (1988) Network model of shape-from-shading: neural function arises from both receptive and projective fields. Nature 333:452–454.

Lekhy SR, Sejnowski TJ (1990) Neural network model of visual cortex determining surface curvature from images of shaded surfaces. Proc R Soc Lond B 240:251–278

Lerner Y, Hendler T, Ben-Bashat D, Harel M, Malach R (2001) A hierarchical axis of object processing stages in the human visual cortex. Cereb Cortex 11:287–297.

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

Lisberger SG, Movshon JA (1999) Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. J Neurosci 19:2224–2246.

Logothetis NK, Pauls J, Bülthoff HH, Poggio T (1994) View-dependent object recognition by monkeys. Curr Biol 4:401–414.

Logothetis NK, Pauls J, Poggio T (1995) Shape representation in the inferior temporal cortex of monkeys. Curr Biol 5:552–563.

Malach R, Reppas JB, Benson RB, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RBH (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci USA, 92:8135–8138.

Malach R, Grill-Spector K, Kushnir T, Edelman S, Itzchak Y (1998) Rapid shape adaptation reveals position and size invariance. Neuroimage 7:S43.

Mamassian P, Landy MS (2001) Interaction of visual prior constraints. Vision Res 41:2653–2668

Marr D, Nishihara HK (1978) Representation and recognition of the spatial organization of three-dimensional shapes. Proc R Soc Lond B Biol Sci, 200:269–294.

Maunsell JH, Van Essen DC (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey II. Binocular interactions and sensitivity to binocular disparity. J Neurophysiol 49:1148–1167.

Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RBH (1999) The representation of real and illusory contours in human cortical visual areas revealed by fMRI. J Neurosci 19:8560–8572.

Miller EK, Li L, Desimone R (1991) A neural mechanism for working memory and recognition memory in inferior temporal cortex. Science 254:1377–1379.

Miller EK, Erikson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque J Neurosci 16:5154–5167.

Moore C, Engel SA (2001) Neural response to perception of volume in the lateral occipital complex. Neuron 29:277–286.

Mueller JR, Metha AB, Krauskopf J, Lennie P (1999) Rapid adaptation in visual cortex to the structure of images. Science 285:1405–1408.

Murray SO, Olshausen BA, Woods DL (2003) Processing shape, motion and three-dimensional shape-from-motion in the human cortex. Cereb Cortex 13:508–516.

Orban GA, Sunaert S, Todd JT, Van Hecke P, Marchal G (1999) Human cortical regions involved in extracting depth from motion. Neuron 24:929–940.

Paradis AL, Cornilleau-Peres V, Droulez J, van de Moortele PF, Lobel E, Berthoz A, Le Bihan D, Poline KB (2000) Visual perception of motion and 3-D structure from motion: an fMRI study. Cereb Cortex 10:772–783.

Perrett DI, Oram MW, Ashbridge E (1998) Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations. Cognition 67:111–145.

Poggio T, Edelman S (1990) A network that learns to recognize three-dimensional objects. Nature 343:263–266.

Poggio GF, Fischer B (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. J Neurophysiol 40:1392–1405.

Poggio GF, Gonzalez F, Krause F (1988) Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. J Neurosci 8:4531–4550.

Riesenhuber M, Poggio T (2000) Models of object recognition. Nat Neurosci 3:1199–1204.

Rosen BR, Buckner RL, Dale AM (1998) Event-related functional MRI: past, present, and future. Proc Natl Acad Sci USA 95:773–780.

Roy JP, Komatsu H, Wurtz RH (1992) Disparity selectivity of neurons in monkey extrastriate area MST. J Neurosci 12:2478–2492.

Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. Science 260:995–997.

Sereno ME, Trinath T, Augath M, Logothetis NK (2002) Three-dimensional shape representation in monkey cortex. Neuron 33:635–652.

Sinha P, Poggio T (1996) Role of learning in three-dimensional form perception. Nature 384:460–463.

Shikata E, Tanaka Y, Nakamura H, Taira M, Sakata H (1996) Selectivity of the parietal visual neurons in 3-D orientation of surface of stereoscopic stimuli. Neuroreport 7:2389–2394.

Taira M, Tsutsui KI, Jiang M, Yara K, Sakata H (2000) Parietal neurons represent surface orientation from the gradient of binocular disparity. J Neurophysiol 83:3140–3146.

Tanaka H, Uka T, Yoshiyama K, Kato M, Fujita I (2001) Processing of shape defined by disparity in monkey inferior temporal cortex. J Neurophysiol 85:735–744.

Tarr MJ, Bülthoff HH (1998) Image-based object recognition in man, monkey and machine. Cognition 67:1–20.

Thomas OM, Cumming BG, Parker AJ (2002) A specialization for relative disparity in V2. Nat Neurosci 5:472–478.

Uka T, Tanaka H, Yoshiyama K, Kato M, Fujita I (2000) Disparity

selectivity of neurons in monkey inferior temporal cortex. J Neurophysiol 84:120–132.

Ullman S (1998) Three-dimensional object recognition based on the combination of views. Cognition 67:21–44.

Vanduffel W, Fize D, Peuskens H, Denys K, Sunaert S, Todd JT, Orban GA (2002) Extracting 3-D from motion: differences in human and monkey intraparietal cortex. Science 298:413–415.

von der Heydt R, Zhou H, Friedman H (2000) Representation of stereoscopic edges in monkey visual cortex. Vision Res 40:1955–1967.

Vogels R, Biederman I, Bar M, Lorincz A (2001) Inferior temporal neurons show greater sensitivity to nonaccidental than to metric shape differences. J Cogn Neurosci 13:444–453.

Vuilleumier P, Henson RN, Driver J, Dolan RJ (2002) Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nat Neurosci 5:491–499.

Watanabe M, Tanaka H, Uka T, Fujita I (2002) Disparity-selective neurons in area V4 of macaque monkeys. J Neurophysiol 87:1960–1973.

Wiggs CL, Martin A (1998) Properties and mechanisms of perceptual priming. Curr Opin Neurobiol 8:227–233.