

Cortical Regions Involved in Perceiving Object Shape

Zoe Kourtzi and Nancy Kanwisher

Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, and Massachusetts General Hospital-Nuclear Magnetic Resonance Center, Charlestown, Massachusetts 02129

The studies described here use functional magnetic resonance imaging to test whether common or distinct cognitive and/or neural mechanisms are involved in extracting object structure from the different image cues defining an object's shape, such as contours, shading, and monocular depth cues. We found overlapping activations in the lateral and ventral occipital cortex [known as the lateral occipital complex (LOC)] for objects defined by different visual cues (e.g., grayscale photographs and line drawings) when each was compared with its own scrambled-object control. In a second experiment we found a reduced response when objects were repeated, independent of whether they appeared in the same or a different format (i.e.,

grayscale images vs line drawings). A third experiment showed that activation in the LOC was no stronger for three-dimensional shapes defined by contours or monocular depth cues, such as occlusion, than for two-dimensional shapes, suggesting that these regions are not selectively involved in processing three-dimensional shape information. These results suggest that common regions in the LOC are involved in extracting and/or representing information about object structure from different image cues.

Key words: shape perception; object recognition; neuroimaging; visual cortex; lateral and ventral occipital; adaptation

Within a few hundred milliseconds of casting your eyes on a novel object, you have extracted a representation of its shape. To do this the visual system must use a variety of different cues present in the image that provide information about the structure of the object. For example, line drawings portray object structure using contours, whereas grayscale photographs provide surface cues such as shading information. Monocular or binocular depth cues can also convey information about object shape. Each of these kinds of cues implicates very different image features that carry the critical shape information and very different computations likely to be involved in extracting that information. The present study uses functional magnetic resonance imaging (fMRI) to ask whether common mechanisms are involved in the extraction and representation of object shape, independent of the image cues defining that shape.

Several fMRI studies have implicated a region of lateral and ventral occipital cortex [the lateral occipital complex (LOC)] in the analysis of object shape. First, this area has been shown to respond more strongly when subjects view images of objects than when they view scrambled versions of these images in which object structure is not present (Malach et al., 1995; Kanwisher et al., 1996). Second, the response in this region does not depend on whether the depicted objects are familiar or not, suggesting that

the LOC is involved in the stimulus-driven analysis of object shape rather than in a later stage of processing that depends on stored knowledge of specific objects (Malach et al., 1995; Kanwisher et al., 1996). Third, the LOC responds to object silhouettes independent of whether their contours are defined by texture, motion, or luminance (Grill-Spector et al., 1998) and independent of changes in the position or size of the object (Malach et al., 1998; Grill-Spector et al., 1999). These findings suggest that the LOC is engaged in the analysis not of low-level image features but of some higher-level aspect of object shape.

But what precise aspect of visual shape analysis is performed in the LOC? Neuropsychological patients with deficits in shape perception are often more impaired in recognizing objects from line drawings than from grayscale photographs (Farah, 1990), suggesting that different mechanisms may be involved in extracting shape information from contours and from shading cues. Consistent with this hypothesis, the brain areas activated by photographs (Malach et al., 1995) appear to be more lateral and posterior to the areas activated by line drawings (Kanwisher et al., 1996). A further question is whether depth information plays a critical role in the shape analysis performed in the LOC.

The following experiments address these questions using fMRI. Experiments 1 and 2 examine whether common regions are involved in processing line drawings of objects and grayscale photographs of the same objects that include information about shading and surfaces. Experiment 3 tests whether any cortical regions are selectively involved in processing information about the three-dimensional structure of objects.

MATERIALS AND METHODS

Subjects

Eleven right-handed Massachusetts Institute of Technology students participated in Experiments 1 and 3. Each subject participated in these two experiments in the same scan. One subject was excluded from the analysis because of excessive head motion. Another nine subjects participated in Experiment 2.

Received Dec. 22, 1999; revised Feb. 10, 2000; accepted Feb. 11, 2000.

This research was supported by the National Institute of Mental Health Grant 56037 and a Human Frontiers grant to N.K. We would like to thank Bart Anderson, Patrick Cavanagh, Paul Downing, Russell Epstein, Kalanit Grill-Spector, Ken Nakayama, and Pawan Sinha for their helpful comments and suggestions on this manuscript. Thanks to Carol Yin for helping with the stimulus generation. We would also like to thank Bruce Rosen and many people at the Massachusetts General Hospital-Nuclear Magnetic Resonance Center for technical assistance and support.

Parts of this manuscript have been presented previously at the 1999 meeting of the Society for Neuroscience in Miami.

Correspondence should be addressed to Dr. Zoe Kourtzi, Department of Brain and Cognitive Science, Massachusetts Institute of Technology, NE20-4043, 3 Cambridge Center, 77 Massachusetts Avenue, Cambridge, MA 02139-4307. E-mail: zoe@psyche.mit.edu.

Copyright © 2000 Society for Neuroscience 0270-6474/00/203310-09\$15.00/0

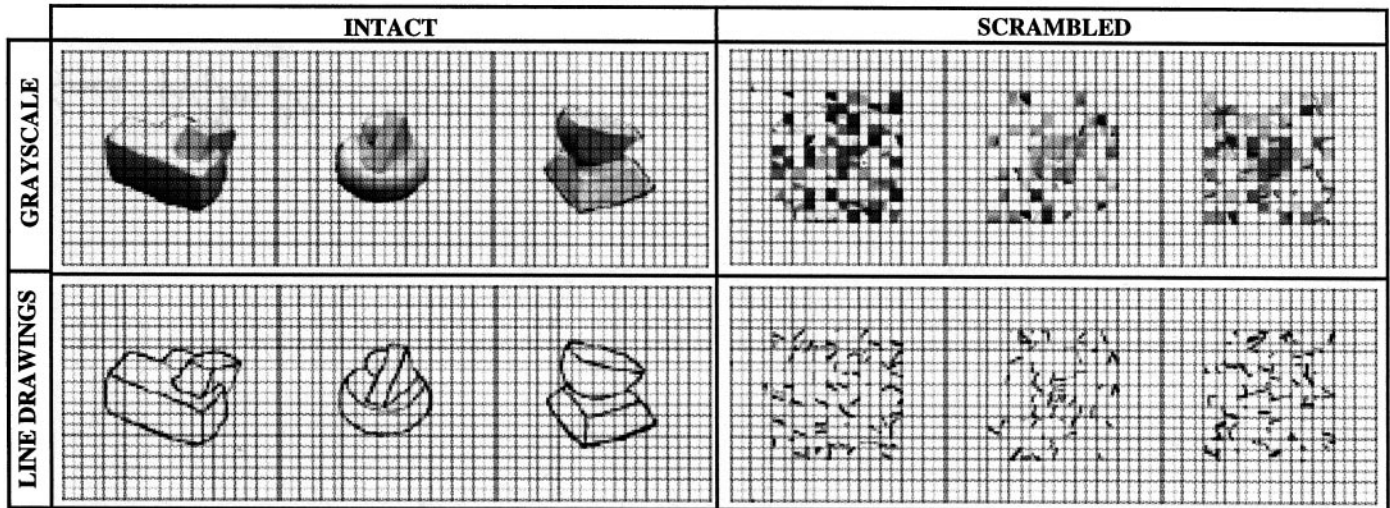


Figure 1. Three example stimuli from each of the four stimulus conditions in Experiment 1.

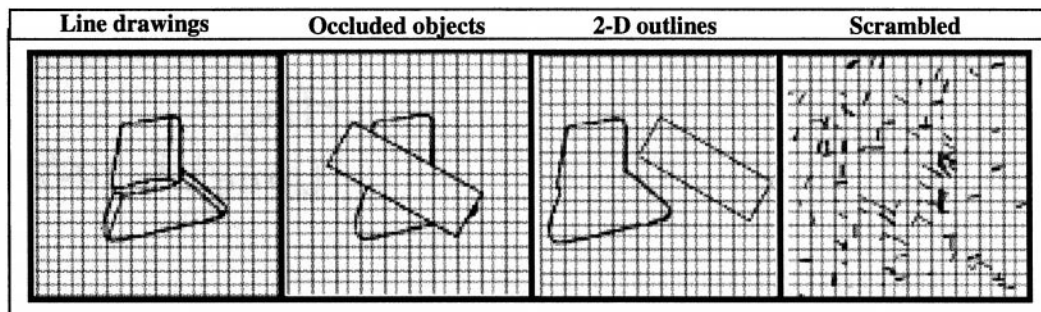


Figure 2. An example stimulus from each condition in Experiment 3. 2-D, Two-dimensional.

Visual stimuli

The stimuli used in all the experiments were 300×300 pixel images of objects. In Experiment 1 the stimuli were grayscale photographs and line drawings of the same novel objects made of clay as well as scrambled versions of each set (Fig. 1). The line drawings were generated by tracing the external outline and the internal contours of the novel objects depicted in the grayscale photographs. 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 both in the intact and in the scrambled images.

In Experiment 2, we used grayscale photographs and line drawings of familiar objects. In Experiment 3, four different stimulus conditions (see Fig. 2) were used to contrast the response to images of objects containing monocular depth cues with the response to images not containing depth cues. In particular, the images containing depth cues were (1) line drawings of the novel three-dimensional objects used in Experiment 1 and (2) outlines of the same objects in which all internal contours were removed and the shape was occluded by a rectangular contour. It has been shown that occluded objects are perceived and represented as whole objects in a depth plane behind the plane of the occluder (Shimojo et al., 1988; Nakayama et al., 1989; Kovacs et al., 1995). These conditions with depth cues were compared with two different kinds of control stimuli: (1) two-dimensional outlines of the same objects described above (i.e., images of the same outline shapes and the rectangular occluder next to each other) and (2) scrambled versions of the two-dimensional outlines.

Procedure and design

Blocked presentation designs (Experiments 1 and 3). In Experiments 1 and 3, a blocked presentation design was used. Each scan consisted of 16 16 sec stimulus epochs with fixation periods interleaved (see Fig. 3d). Twenty different images of the same type were presented in each epoch. Each image was presented for 200 msec with a blank interval of 600 msec

between items. Each of the four stimulus types was presented in four different epochs within each scan, in a design that balanced for the order of conditions.

Each subject was run on four scans, each of which lasted 5 min and 36 sec. Seven subjects were run on two scans in a passive-viewing task and on two scans in a one-back-matching task. Three subjects were run on the one-back-matching task for all the runs. For the passive-viewing conditions, the subjects were asked to observe the images carefully while fixating on a dot in the center of the image. For the one-back-matching condition, subjects 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 the observer's attention on all the stimulus conditions used.

Event-related adaptation (Experiment 2): logic and design. Experiment 2 used adaptation effects (Miller et al., 1991, 1996; Buckner et al., 1998; Malach et al., 1998; Wiggs and Martin, 1998; James et al., 1999) to test whether common or distinct neural populations respond to the shape of an object depicted in a line drawing versus a grayscale photograph. In fMRI adaptation, decreased activation is found when a stimulus is repeated (Buckner et al., 1998; Malach et al., 1998). This adaptation effect has been observed in the LOC even when the location or size of the object changes across repetitions (Malach et al., 1998; Grill-Spector et al., 1999), demonstrating that a common neural population responds to an object across changes in location and size. If distinct neural populations responded to different locations and sizes of an object, then no adaptation would be observed across such changes.

Previous fMRI studies testing for adaptation across image changes used blocked designs in which the observers were presented with a number of images repeatedly in a block of trials. However, the event-related design (Buckner et al., 1998; Rosen et al., 1998) used in the present study enabled us to test the degree of adaptation across two specific stimuli unconfounded from the effects of other stimuli in the experiment. In each trial, two images were presented in rapid succession.

The two images were either of the same object or of completely different objects, and in each case the two images were in either the same format (i.e., line drawing vs grayscale photograph) or in different formats. The line drawings were generated by tracing the external outline and the internal contours of the objects depicted in the grayscale photographs. That is, a line drawing and a grayscale photograph of the same object were identical in all other visual features except for their format. This design allowed us to compare the signal reductions of the same compared with different objects as a function of whether the two objects were presented in the same or different formats. Decreased activations for repeated images of the same object independent of changes in object format would suggest that the same neural population responds to the shape of the object independent of the cues defining that shape.

These adaptation effects were measured within the LOC using an independent data set to define the LOC region of interest. Specifically, subjects were scanned on four “localizer” runs of an experiment identical to Experiment 1 involving intact and scrambled images of grayscale images and line drawings of objects. The region of interest (ROI) for each subject was defined as the set of all contiguous voxels in the ventral occipitotemporal cortex that were activated more strongly ($p < 10^{-3}$) by the intact than by the scrambled images.

After the four localizer scans, each subject was run on four scans in the event-related experiment. Each scan lasted 6 min and 4 sec and consisted of one 5 min and 32 sec epoch of experimental trials and two 16 sec fixation epochs, one at the beginning and one at the end of the scan. The experimental trials consisted of a pair of images presented sequentially. Each image was presented for 300 msec with a blank interval of 400 msec between images. After the two images were presented in each trial, a blank interval of 1 sec followed before the beginning of a new trial. Thus, each experimental trial lasted 2 sec. Recent fMRI event-related studies have demonstrated that responses to trials presented every 2 sec add approximately linearly, and thus the effects of interleaved trials of different conditions can be separated by event-related signal averaging (Boynton et al., 1996; Cohen, 1997; Dale and Buckner, 1997).

Four types of trials were presented, in which (1) the identities of the objects were the same and so were the formats of the images (in half of these trials the two images were identical grayscale photographs, and in the other half they were identical line drawings), (2) the identities of the objects were different but the formats were the same, (3) the identities of the objects were the same but the formats were different (in half of the trials a grayscale photograph appeared first, followed by a line drawing; the order of the two stimuli was reversed for the other half of the trials), or (4) both the identities of the objects and their formats were different. There were 36 trials of each type in each of the four scans. Sixteen fixation-only trials were interleaved in the experiment to serve as the baseline for the data analysis (see below). Thus, a total of 162 trials were presented in each scan. Images were not repeated across trials.

The order of trials was such that the composition of the immediately preceding trial was identical across conditions. That is, for each of the four conditions, the same number of trials of each condition occurred in the immediately preceding trial. Although it was impossible to match the history of preceding trials farther back than the immediately preceding trial, this history was derived by random assignment. Furthermore, to counterbalance for any imbalances, two versions of the experiment were made in which the serial positions of trials were exchanged between each pair of critical conditions. Thus, for the same format trials, the set of serial positions occupied by the trials in the same object condition and those in the different object condition were exchanged; the same was true for the different format trials.

Subjects were instructed to view the images passively while fixating.

MRI acquisition

For all the experiments, scanning was done on the 3 T GE scanner (modified for echo planar imaging) at the Massachusetts General Hospital-Nuclear Magnetic Resonance Center in Charlestown, MA. A custom bilateral surface coil was used to provide a high signal-to-noise ratio in posterior brain regions. A bite bar was used to minimize head motion. Standard imaging procedures (Gradient Echo pulse sequence, TE, 30 msec; flip angle, 90°; 180° offset, 25 msec) were used as described previously (Tong et al., 1998).

For the blocked design experiments (Experiments 1 and 3), data were collected from 12 6-mm-thick near-coronal slices oriented parallel to the brainstem and covering the occipital lobe as well as the posterior portions of the temporal and the parietal lobes. One hundred sixty-eight functional images were collected for each slice in each scan.

For the event-related experiment in Experiment 2, data were collected every second (TR = 1) from six 7-mm-thick near-coronal slices. One hundred sixty-eight functional images were collected for each slice in each localizer scan, whereas 356 images were collected for each slice in each event-related scan.

Data analysis

Experiment 1. To identify regions in the brain that show significantly stronger activation to intact than to scrambled grayscale images and to intact than to scrambled line drawings, *t* tests at the level of $p < 0.001$ were applied on the data averaged over the four scans. To identify regions in the LOC involved in the processing of both grayscale images and line drawings (the “overlap ROI”), we identified voxels responding significantly to both grayscale images (intact vs scrambled) and line drawings (intact vs scrambled). We also identified regions that reached significance (in the comparison for intact vs scrambled images) for grayscale photographs but not for line drawings (“grayscale-only ROI”) and for line drawings but not for grayscale photographs (“line drawing-only ROI”). This analysis was conducted on the data for each subject individually as well as on the group data generated by averaging across subjects coregistered into Talairach space (Talairach and Tournoux, 1988).

To derive the MR signal intensity over the period of the scan in each of these regions while using an independent data set to derive the ROIs, we split the group data into two sets. That is, the last two scans (one passive-viewing scan and one one-back-matching scan) from each subject were used to identify the ROIs described above (overlap ROI, grayscale-only ROI, and line drawing-only ROI). Using these independently defined ROIs, we then derived the time course of the MR signal intensity in each of these voxels for the data from the first two scans (one passive-viewing scan and one one-back-matching scan). The average percent signal change (PSC) was then calculated for each stimulus type, using the average signal intensity during fixation epochs as a baseline. Because the fMRI response typically lags 4–6 sec after the neural response, our data analysis procedure treated the first image in each epoch as belonging to the condition of the preceding epoch and omitted the next two images (during the transition between epochs) from the analysis. A two-way ANOVA with shape cue (grayscale images vs line drawings) and shape structure (intact vs scrambled) as within-measure variables (repeated measurements from the four epochs averaged across the first two scans) was run on the average percent signal for each one of the ROIs. Because the data were analyzed within independently defined ROIs, no correction for multiple voxelwise comparisons was required.

Experiment 2. For each individual subject, regions in the LOC that showed significantly stronger activation to intact than to scrambled images were identified by applying *t* tests at the level of $p < 0.0001$ on the average of the four localizer scans. These regions served as the ROI for analyzing the data in each event-related scan for each subject.

For each event-related scan, the time course of MR signal intensity was extracted by averaging the data from all the voxels within the independently defined ROI. The average time course of signal intensity was then calculated for each condition, using the average signal intensity during the fixation trials as a baseline. Specifically, in each scan we averaged the signal intensity across the 36 trials in each condition at each of 12 corresponding time points (seconds). These event-related time courses of signal intensity were then converted to time courses of percent signal change for each of the four conditions by subtracting the corresponding value for the fixation condition and then dividing by that value. The resulting time course for each condition was then averaged across scans for each subject and then across subjects (see Fig. 5). These time courses allowed us to observe the activation signal throughout each trial, because the fMRI response typically lags 4–6 sec after the neural response. A stronger percent signal change for trials with objects of different identity than for trials with objects of the same identity independent of the format would suggest an adaptation effect on the identity of the object.

ANOVAs on the average percent signal change across individual subject data were conducted with stimulus identity (with two levels, same object pairs and different object pairs independent of the object format) and stimulus format (with two levels, same format and different format independent of object identity) as within-measure variables. Adaptation effects should be observed at the peak (time point 5) but not at the beginning (time point 0) of the trials (see Fig. 5). As a result, we tested the effects of the above variables at the beginning and the peak of the trials.

Experiment 3. For the group analysis, the overlap ROI based on the group data from Experiment 1 was used as a localizer for this experiment. The average percent signal change was calculated for each stimulus type within the overlap ROI. An ANOVA tested for differences across stimulus conditions and tasks within the overlap ROI. A similar analysis was conducted on the data for each subject. The overlap ROI was calculated for each subject on the basis of the data for that subject from Experiment 1. The average percent signal change was calculated for each stimulus type in Experiment 3 within the overlap ROI for each subject.

RESULTS

Experiment 1: activations for grayscale images and line drawings of novel objects

Figure 3 shows the regions in the lateral occipital complex and the parietal cortex in one representative subject that showed significantly stronger activation for intact than for scrambled grayscale images (Fig. 3*a*) and for intact than for scrambled line drawings of novel objects (Fig. 3*b*) used in Experiment 1. Note the highly similar pattern of activation for the two different kinds of stimuli. Similar results were observed in each subject.

Figure 3*c* shows the regions responding significantly more strongly to intact objects (grayscale photographs or line drawings) than to scrambled objects on the basis of the group analysis. Figure 3*d* shows the time course of the percent signal change over the period of the scan on the data from the first two scans in an independently derived ROI of all the voxels in the ventral pathway that responded significantly more strongly to intact than to scrambled images of objects in the last two scans.

Analysis of data from the LOC

Figure 4*a* shows the LOC regions that were activated significantly more strongly ($p < 10^{-3}$) by intact than by scrambled images for both grayscale images and line drawings (overlap ROI, shown in *yellow*), for grayscale images but not line drawings (grayscale-only ROI, shown in *red*), and for line drawings but not grayscale images (line drawing-only ROI, shown in *blue*) on the group data across all scans. To determine whether these different sets of voxels have reliably different functional properties, we analyzed the responses in each of the resulting regions independently. To do this analysis in an unbiased way, we used the data from two of the scans to define each of these regions of interest; data from the other two scans were then used to measure the response to each of the four stimulus categories in each of these regions. The responses in these regions were quantified as the PSC from the fixation baseline condition. Figure 4*b–d* illustrates the PSC for each condition for each one of the above ROIs for the data from the first two scans.

A two-way ANOVA on the PSC data over the four epochs in the overlap ROI with shape cue (grayscale images vs line drawings) and shape structure (intact vs scrambled) as repeated measures variables (responses from four epochs in two averaged runs) showed a significant main effect of shape structure [$F_{(1,3)} = 57.7$; $p < 0.01$]. There was no main effect of shape cue [$F_{(1,3)} = 1.1$; $p > 0.3$], and there was no interaction of shape cue and shape structure [$F_{(1,3)} < 1$]. The same analysis on the PSC in the grayscale-only ROI showed a significant main effect of shape structure [$F_{(1,3)} = 39$; $p < 0.01$]. There was no main effect of shape cue [$F_{(1,3)} = 5.3$; $p > 0.1$], and there was no interaction of shape cue and shape structure [$F_{(1,3)} < 1$]. However, the same analysis on the PSC in the line drawing-only ROI did not show significant main effects of shape structure [$F_{(1,3)} = 4$; $p > 0.1$] or shape cue [$F_{(1,3)} = 3$; $p > 0.1$]. No interaction of shape cue and shape structure [$F_{(1,3)} = 1.5$; $p > 0.1$] was observed.

These statistical results, in combination with the response profiles shown in Figure 4, show that most of the regions in the LOC (overlap regions and Grayscale-only regions) are selective for both grayscale photographs and line drawings of objects. Only a small number of voxels that were significantly activated only by line drawings did not show selectivity for both types of images. Thus, these results show that most regions in the LOC are involved selectively in the analysis of object shape from both grayscale photographs and line drawings. Furthermore, a direct comparison of the response to intact grayscale photographs versus intact line drawings in the group data did not reveal any voxels that responded significantly more strongly ($p < 10^{-3}$) to one stimulus type than to the other. Thus, these data suggest that all regions (within our scanning range and resolution) that respond to object shape show a stimulus sensitivity to shape that is independent of the cues defining that shape.

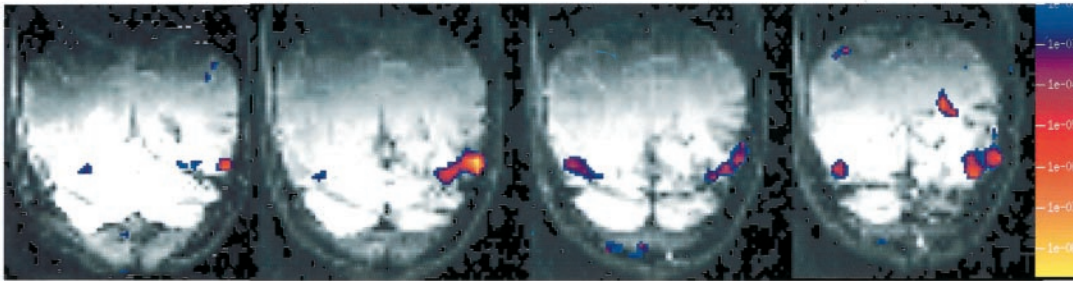
It is possible that the stronger activation observed for intact than for scrambled images reflects stronger attentional engagement by the intact images. However, in the one-back-matching task the subjects had to attend at least as strongly to the scrambled images as to the intact ones because the one-back-matching task is considerably more demanding for scrambled images. We tested this attentional account by analyzing the PSC in the one-back-matching condition alone. To this end, we selected the regions that showed stronger activation for intact than for scrambled images in the passive runs and used them to localize the ROI for analyzing the one-back-matching runs. This analysis was conducted on the group data. For regions in the LOC, a two-way ANOVA on the PSC with shape cue (grayscale images vs line drawings) and shape structure (intact vs scrambled) as repeated measures variables showed a significant main effect of shape structure [$F_{(1,3)} = 15.7$; $p < 0.05$]. There was no main effect of shape cue [$F_{(1,3)} = 1.9$; $p > 0.26$], and there was no interaction of shape cue and shape structure [$F_{(1,3)} < 1$]. Thus, the same pattern of selectivity for shape independent of visual cues is observed in the LOC even when the one-back data were analyzed alone. Thus, this response is unlikely to reflect differences in task difficulty or attentional allocation across conditions.

Taken together these results show a considerable overlap between regions in the LOC that respond to grayscale images and line drawings. Indeed, it seems that most or all of the regions in the LOC process object structure independent of the cues defining the object's shape.

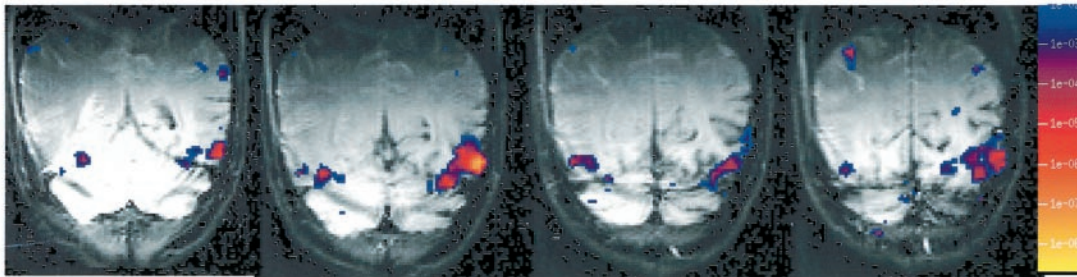
Analysis of data from the parietal cortex

Interestingly, activations for intact images were also observed in regions of the parietal cortex (see Fig. 3). Recent neurophysiological studies suggest a role of the parietal cortex in shape processing (Sakata et al., 1997; Sereno and Maunsell, 1998). Several imaging studies also report activations in the parietal regions for shape processing (Price et al., 1996; Dolan et al., 1997; Faillenot et al., 1997; Kraut et al., 1997; Vanni et al., 1997). However, numerous studies suggest an important role of the parietal cortex in attention-related processes (Colby et al., 1995; Wojciulik and Kanwisher, 1999). Thus, to test whether regions in the parietal cortex are involved in shape processing, it is important that we control for possible attentional confounds, namely, the possibility that stronger activations in regions of the parietal cortex for intact images of objects than for scrambled images may result because intact images engage a subject's attention more

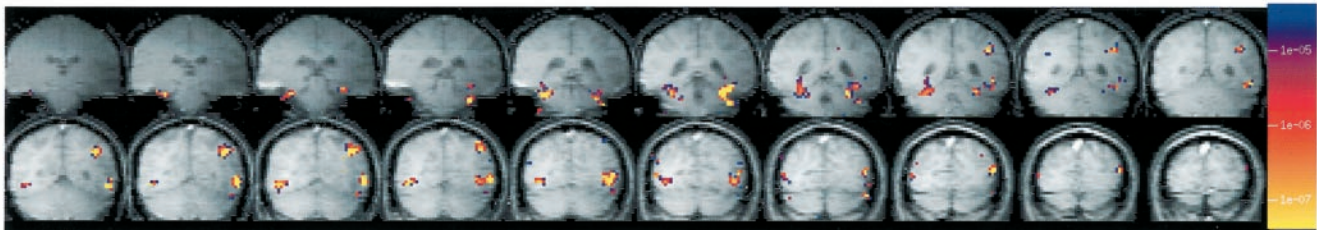
a. Grayscale Photographs: Intact objects > Scrambled objects.



b. Line drawings: Intact objects > Scrambled objects.



c. Group data: Intact objects > Scrambled objects.



d. Time course of signal intensity in the LOC.

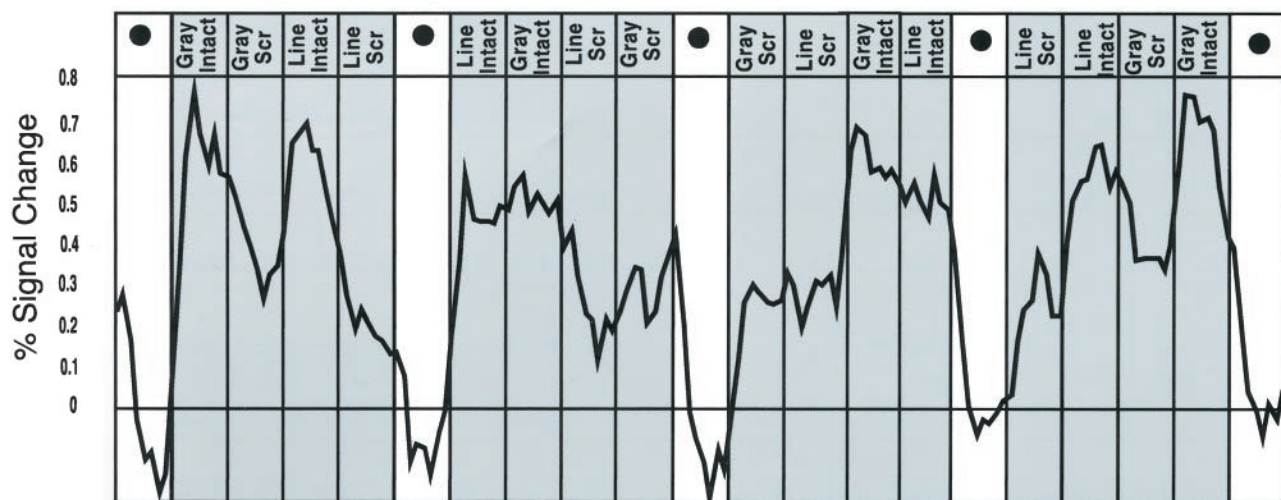


Figure 3. *a, b*, Four consecutive slices from one subject showing regions responding significantly more strongly to intact than to scrambled objects represented as grayscale photographs (*a*) and line drawings (*b*). Statistical results are overlaid on a T1-weighted anatomical image of each slice; significance levels reflect the results of (uncorrected) *t* tests on a comparison of the MR signal intensity during the viewing of intact objects versus scrambled objects. The right hemisphere appears on the *left*. *c, d*, Group analysis on functional data from 10 subjects (coregistered in Talairach space) showing regions responding significantly to intact versus scrambled images (*c*; $p < 10^{-7}$, equivalent to $p < 10^{-1}$ after Bonferroni correction) and the time course of the percent signal change (from the fixation baseline) for the group data over the period of the scan (*d*). The colors in the scales at the *right* (*a–c*) indicate the significance levels (*p* values) for the *t* tests applied to the data. The group data from the first two scans were used to derive the ROI; the time course shown is for the independent data of the third and fourth scans, averaged over all voxels in the ROI. A black dot indicates fixation. Gray Intact, Grayscale intact images; Gray Scr, grayscale scrambled images; Line Intact, intact line drawings; Line Scr, scrambled line drawings.

a. Significant responses to grayscale photographs (red), line drawings (blue), and both (yellow).

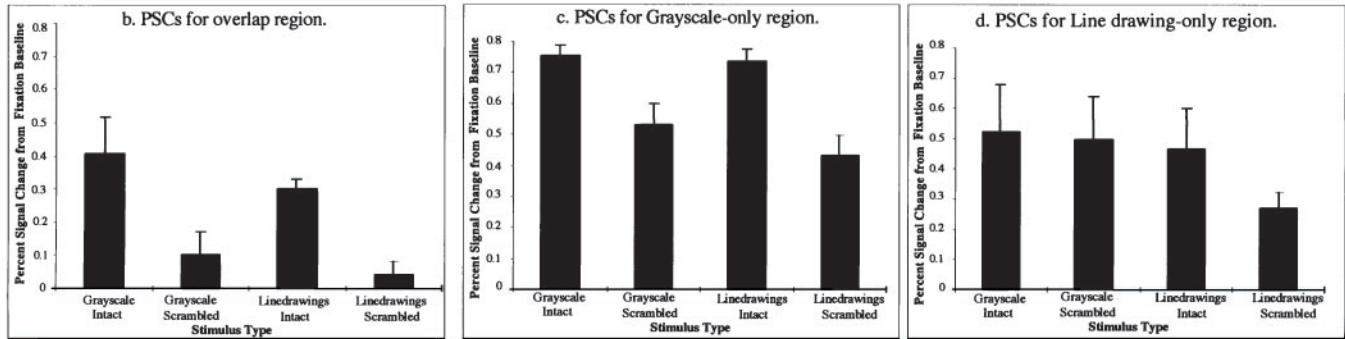
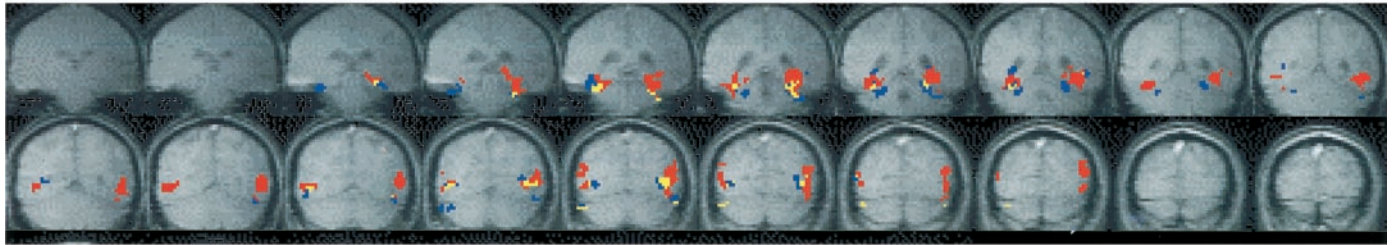


Figure 4. Results of Experiment 1. *a*, Voxels in the ventral pathway in which the statistical comparison of intact versus scrambled objects reached significance ($p < 0.001$, uncorrected; group data including all scans) for (1) both grayscale images and line drawings (overlap ROI in yellow), (2) grayscale images but not line drawings (grayscale-only ROI in red), and (3) line drawings only (line drawing-only ROI in blue). *b–d*, The average percent signal change (from the fixation baseline) for each stimulus type based on the group data from the first two scans in the overlap ROI (*b*), the grayscale-only ROI (*c*), and the line drawing-only ROI (*d*) as defined in the last two scans. Data are reported only from regions in the ventral pathway and not from other regions in the brain (i.e., the parietal cortex) as discussed in Results (see Experiment 1). The error bars indicate SEs on the PSC measured across the four repetitions of each stimulus condition averaged across scans.

strongly than do scrambled images. Most past studies showing shape responses in the parietal lobe do not control for possible attentional confounds. In the current experiment we controlled for these confounds by running the subjects on the one-back-matching task.

To analyze independently the PSC in the one-back-matching condition on the group data, we selected the regions that showed stronger activation for intact than for scrambled images in the passive runs and used them as a localizer for the one-back-matching runs. A two-way ANOVA on the PSC with shape cue (grayscale images vs line drawings) and shape structure (intact vs scrambled) as repeated measures variables showed a significant main effect of shape cue [$F_{(1,3)} = 10.1$; $p = 0.05$], with stronger activation for grayscale images than for line drawings, but no main effect of shape structure [$F_{(1,3)} < 1$] and no interaction of shape cue and shape structure [$F_{(1,3)} < 1$].

Thus, the analysis of the activations in the parietal regions did not show any selectivity for intact images of objects. It is possible that the observed activations in the parietal cortex are caused by the fact that intact images of objects engage the observer's attention more strongly than do scrambled images of the same objects. Interestingly, our data showed a main effect of the shape cue with stronger activations in the parietal cortex for grayscale images than for line drawings. This effect seems consistent with studies suggesting that temporal regions are involved mainly in extracting shape from luminance edges, whereas parietal regions may extract shape from other depth cues, such as shading (Humphrey et al., 1996) or surface slant (Sakata et al., 1997). Further imaging studies are needed to investigate more systematically the role of the parietal lobes in shape processing.

Experiment 2: adaptation for grayscale images and line drawings of familiar objects

The regions in the LOC that responded significantly more strongly to intact than to scrambled images (grayscale images or line drawings) in the localizer scans for each subject were used as the ROI for analyzing the data in the event-related scans. The average time course of percent signal change in this ROI (from the corresponding fixation baseline) was calculated for all trial types (same identity/same format, same identity/different format, different identity/same format, and different identity/different format trials) on the data from the event-related scans for each subject.

Consistent with previous event-related data, the peak response for all conditions occurred at a latency of 5 sec after the beginning of the trial. The magnitude of this peak response was analyzed in a two-way repeated ANOVA across subjects with stimulus identity (same identity, different identity trials independent of format) and stimulus format (same format, different format trials independent of object identity) as repeated measures variables. As shown in Figure 5 a main effect of stimulus identity [$F_{(1,8)} = 41.1$; $p < 0.001$] but no effect of stimulus format [$F_{(1,8)} < 1$] and no interaction of stimulus identity and stimulus format [$F_{(1,8)} < 1$] were observed. A similar analysis at the beginning of the trial revealed no significant differences between trial conditions.

Finally, a similar analysis was conducted in early retinotopic regions bordering the calcarine sulcus that were activated significantly more strongly by all types of visual patterns (intact and scrambled images) than by fixation. These regions did not show any adaptation effects. Specifically, no effect of stimulus identity [$F_{(1,8)} < 1$] or stimulus format [$F_{(1,8)} < 1$] was observed (same

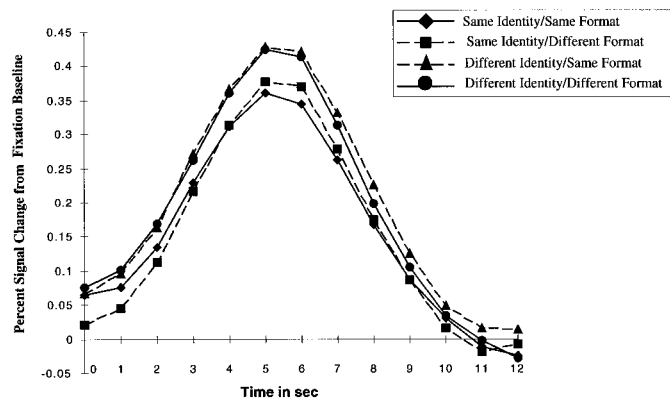


Figure 5. Results of Experiment 2. Average percent signal increase (from the fixation baseline trials) for each trial condition. Trials start at time = 0 sec and peak at time = 5 sec.

identity/same format trials, PSC = 0.35; same identity/different format trials, PSC = 0.26; different identity/same format trials, PSC = 0.33; and different identity/different format trials, PSC = 0.34). These results suggest that the adaptation observed for the same objects in the LOC does not arise in early retinotopic cortex.

Taken together, the results of this experiment show decreased activation within LOC regions when the same object is repeated twice in a trial. Importantly this adaptation effect was not significantly smaller when the same object was defined by different shape cues (i.e., lines and shading) than by the same cue. These results show that neural populations within the LOC are involved in the analysis and/or representation of object structure independent of the cues that define the shape.

One could suggest that the decreased signal observed for the same images indicates reduced attention to repeated stimuli compared with different images. However, the percent signal change for the same objects of the same format was not significantly lower from that for the same objects of different format (as the attentional hypothesis should predict). Moreover, 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.

Experiment 3: activations for three-dimensional contours, occluded shapes, and two-dimensional outlines

The overlap ROI as defined by the group data of Experiment 1 was used to analyze the PSC in the group data for this experiment. A two-way repeated ANOVA (task \times stimulus type) with task (passive vs one-back matching) and stimulus type (line drawings, occluded objects, two-dimensional outlines, and scrambled) as repeated measures variables was used. As shown in Figure 6*a*, a main effect of stimulus type [$F_{(1,3)} = 22.7$; $p < 0.01$] was observed. The main effect of task did not reach significance [$F_{(1,3)} = 6.8$; $p = 0.079$] nor did the interaction of stimulus type and task [$F_{(1,3)} < 1$]. The PSC within the overlap ROI was significantly greater for line drawings [$t(3) = 3.2$; $p < 0.05$], occluded objects [$t(3) = 2.8$; $p < 0.05$], and two-dimensional outlines [$t(3) = 3.5$; $p < 0.05$] than for scrambled images. No significant differences ($p > 0.05$) were observed in the PSC in any of the comparisons between line drawings, occluded objects, and two-dimensional outlines.

As shown in Figure 6*b*, a similar analysis on the PSC extracted

for each subject individually and averaged across subjects within the overlap ROI showed a main effect of stimulus type [$F_{(7,21)} = 22.3$; $p < 0.001$]. The PSC within the overlap ROI was significantly greater for line drawings [$t(7) = 3.2$; $p < 0.05$], occluded objects [$t(7) = 2.8$; $p < 0.05$], and two-dimensional outlines [$t(7) = 3.5$; $p < 0.05$] than for scrambled images. No significant differences ($p > 0.05$) were observed in the PSC in any of the comparisons between line drawings, occluded objects, and two-dimensional outlines.

Finally, t tests at the level of $p < 0.001$ on both the group data and the data from each individual subject did not reveal any voxels that were activated significantly more strongly for objects defined by monocular depth cues (i.e., line drawings and occluded objects) than for two-dimensional outlines.

In summary, although Experiment 3 again found evidence that the LOC is involved in processing object structure, it provides no evidence that this region is selectively involved in the extraction of three-dimensional information in particular. Nonetheless, it remains possible that future experiments will find a stronger response to three-dimensional than to two-dimensional shapes in the LOC or elsewhere. For one thing, the two-dimensional shapes used in Experiment 3 were outlines of the three-dimensional line drawings presented in the same scans, so it is possible that top-down influences caused the two-dimensional outlines to be seen as three-dimensional. For another, we cannot exclude the existence of neural mechanisms selective for three-dimensional information either beyond the cortical regions we scanned or at a finer spatial scale than our techniques could resolve. In any event, the results of this experiment show that the LOC responds strongly to both two- and three-dimensional shapes compared with scrambled control stimuli, further implicating it in the analysis of object structure independent of the cues defining the object's shape.

DISCUSSION

Activations for grayscale images and line drawings of objects

The shape of an object can be conveyed by a variety of different visual cues. Line drawings portray object structure with contours alone, whereas grayscale photographs contain surface information such as shading and texture. Experiments 1 and 2 asked whether the same or different cognitive and/or neural mechanisms are involved in processing object structure from grayscale images and line drawings of objects.

The results of Experiment 1 showed a considerable overlap between regions in the ventral visual pathway that are significantly activated by grayscale images (vs their scrambled controls) and those activated by line drawings (vs their scrambled controls). Furthermore, Experiment 2 showed decreased activations for repeated objects across changes in their format (from grayscale photograph to line drawing or vice versa), showing that common neural populations in the LOC respond to both grayscale photographs and line drawings of the same object. If the overlapping activations observed in Experiment 1 were caused by averaging the MR signal across different neural populations, one responsive to grayscale photographs and one responsive to line drawings, then no decrease in the activation should have been observed when images of the same objects were presented once as grayscale photographs and once as line drawings. Taken together, the results of Experiments 1 and 2 show that a large region of ventral occipitotemporal cortex is involved in the analysis of object structure independent of the cues defining the object's shape. Although

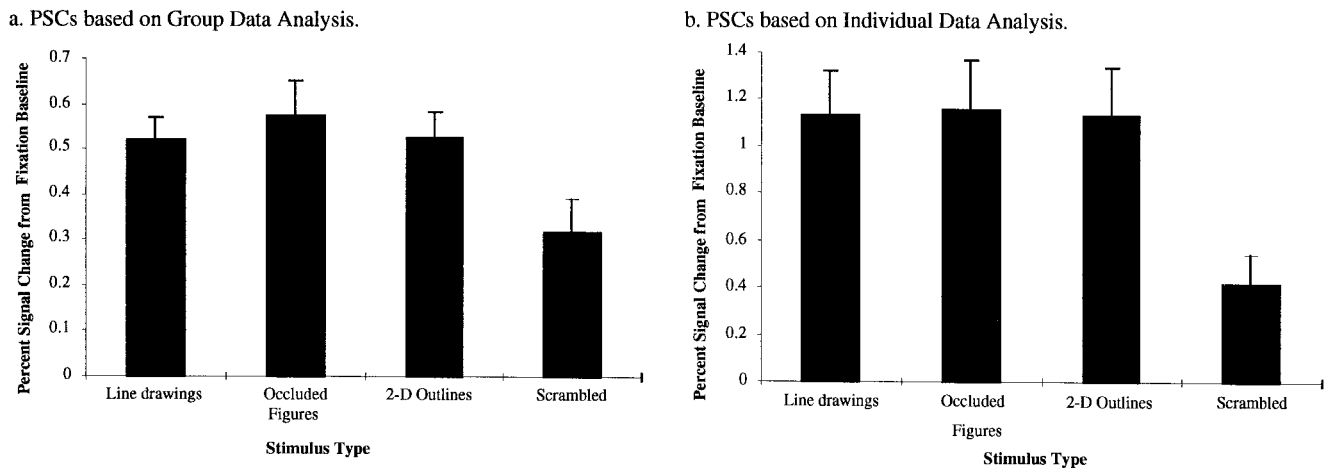


Figure 6. Results of Experiment 3. Average percent signal increase (from the fixation baseline) for each stimulus type in the overlap ROI observed in the group data (*a*) and in the averaged data across subjects (*b*). In the group data analysis, the error bars indicate SEs on the PSC measured across the four repetitions of each stimulus condition averaged across scans. In the analysis across subjects, the error bars indicate SEs on the PSC measured across subjects.

we refer to this region as the “lateral occipital complex,” following Malach et al. (1995), it is not restricted to lateral or occipital regions but extends ventrally into the posterior temporal lobe.

Our results are consistent with numerous neurophysiological studies in monkeys showing selectivity for shape features in cortical areas in the ventral visual pathway, such as V4 and the inferotemporal cortex (Gallant et al., 1993; Logothetis and Sheinberg, 1996; Tanaka, 1996; Pasupathy and Connor, 1999; Vogels, 1999). Moreover, recent neurophysiological (Sary et al., 1993) and imaging (Grill-Spector et al., 1998) studies provide evidence of neural mechanisms involved in analyzing objects independent of the visual cues defining their shape.

The present studies advance our understanding of shape processing in several important respects. First, although previous studies have reported cue-invariant responses to two-dimensional shapes, these studies have not addressed the neural coding of three-dimensional shape information. Our experiments show that common areas respond to three-dimensional shapes defined by lines and shading information. Second, by obligating subjects to attend at least as strongly to the scrambled as to the intact objects via the use of a one-back task (in Experiment 1), we have controlled for differences in attentional engagement that may have contributed to previously reported results. Most importantly, our finding of format-invariant adaptation (in Experiment 2) shows that the common response of the LOC to shapes defined by different cues cannot be accounted for in terms of distinct but interleaved neural populations, each responsive to a different cue. Instead, our data show that a common neural population responds to a given shape whether it is depicted in a line drawing or a grayscale photograph. Finally, the results of Experiments 1 and 2 relate to two further questions, which we discuss briefly next.

Researchers in computational vision have described algorithms for extracting shape from shading (Lekhy and Sejnowski, 1988, 1990) and shape from texture (Landy and Bergen, 1991). If special cortical areas exist in the human brain to perform these functions, one would expect them to be more activated during the viewing of grayscale photographs than of line drawings. For the posterior brain regions we scanned, we found no convincing evidence of such a selective activation for grayscale photographs compared with line drawings. However it remains possible that cortical mechanisms specialized for either shape from shading or

shape from texture exist either beyond the brain regions we scanned or at a finer spatial scale than we could resolve.

Our results may also provide a partial answer to the long-standing question of why lines “work” to produce percepts of object structure. This question arises because the borders of objects and surfaces in real images (and photographs) are marked not by lines but by discontinuities in brightness, texture, color, et cetera (Cavanagh, 1995). The ability to understand line drawings visually may not result from a learned cultural convention, because some evidence suggests that infants and tribesmen with little previous experience instantly recognize objects depicted in line drawings (Hochberg and Brooks, 1962; Kennedy and Ross, 1975). Cavanagh (1995) has suggested that line drawings work for us because they directly activate the brain’s own internal code for object structure. This conjecture is supported by our finding that a line drawing of an object activates a neural population very similar to that activated by a photograph of the same object.

What then is the nature of this internal code for object structure? In neural terms, what is the nature of the processes and representations that occur in the LOC? The fact that this region responds more to intact than to scrambled objects is consistent with a role in any aspect of shape perception, including image segmentation, contour extraction, representing information about depth relations between objects or parts, and the extraction of a high-level description of the shape of the object. The present evidence that this region responds to object shape independent of image format (see also Grill-Spector et al., 1998) and further evidence that representations in this region are invariant to changes in the position and size of an object (Malach et al., 1998; Grill-Spector et al., 1999) show that the LOC does not merely encode low-level features of objects. However the precise function of this region remains to be determined. Our final experiment tested one specific hypothesis.

Activations for three-dimensional and occluded objects versus two-dimensional objects

Experiment 3 tested whether the LOC is selectively involved in extracting object structure from monocular depth cues. The results showed that the LOC is not activated significantly more strongly when monocular depth information is present in the stimulus (e.g., line drawings of three-dimensional objects or oc-

cluded shapes) compared with when there are no depth cues in the image (e.g., two-dimensional outlines). These results show that this region is involved in extracting shape information from both two- and three-dimensional objects.

Recent human fMRI studies have shown stronger activation in the LOC than in earlier retinotopic regions for shapes defined by stereoscopic depth cues (Mendola et al., 1999), suggesting that these regions are involved in the processing of depth information. In these studies red–green random dot stereograms defining a shape were compared with random dot fields that had no shape information. It is possible that the stronger activation observed in the LOC for images with stereoscopic cues was because these images defined shapes while their controls did not. In our studies we compared images with or without monocular depth information, with information about the object shape present in each. Our results provide no evidence of a selective involvement of the LOC in the extraction of monocular depth information.

Conclusions

The current experiments show that common regions in the human LOC are involved in the extraction and/or representation of object structure independent of the image cues that define that structure (e.g., lines, shading, texture, or monocular depth cues). However, the precise nature of the representations and processes that occur in this region remains to be determined. The event-related adaptation procedure introduced here will provide a powerful new tool for tackling this question. The answers will be informative for our understanding of both the functional organization of human visual cortex and the mechanisms underlying human object recognition.

REFERENCES

- 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.
- Cavanagh P (1995) Vision is getting easier every day. *Perception* 24:1227–1232.
- Cohen MS (1997) Parametric analysis of fMRI data using linear systems methods. *NeuroImage* 6:93–103.
- Colby CL, Duhamel JR, Goldberg ME (1995) Oculocentric spatial representation in parietal cortex. *Cereb Cortex* 5:470–481.
- Dale AM, Buckner RL (1997) Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp* 5:329–340.
- Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RSJ, Friston KJ (1997) How the brain learns to see objects and faces in an impoverished context. *Nature* 389:596–599.
- Faillenot I, Toni I, Decety J, Grégoire M, Jeannerod M (1997) Visual pathways for object-oriented action and object recognition: functional anatomy with PET. *Cereb Cortex* 7:77–85.
- Farah MJ (1990) *Visual agnosia: disorders of object vision and what they tell us about normal vision*. Cambridge, MA: MIT.
- Gallant JL, Braun J, Van Essen DC (1993) Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259:100–103.
- Grill-Spector K, Kushnir T, Edelman S, Itzhak 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, Itzhak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203.
- Hochberg JE, Brooks V (1962) Pictorial recognition as an unlearned ability. *Am J Psychol* 75:624–628.
- Humphrey GK, Symons LA, Herbert AM, Goodale MA (1996) A neurological dissociation between shape from shading and shape from edges. *Behav Brain Res* 76:117–125.
- 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.
- Kanwisher N, Chun MM, McDermott J, Ledden PJ (1996) Functional imaging of human visual recognition. *Brain Res Cogn Brain Res* 5:55–67.
- Kennedy JM, Ross AS (1975) Outline picture perception by the Songe of Papua. *Perception* 23:391–406.
- Kovacs G, Vogels R, Orban GA (1995) Selectivity of macaque inferior temporal neurons for partially occluded shapes. *J Neurosci* 15:1984–1997.
- Kraut M, Hart Jr J, Soher BJ, Gordon B (1997) Object shape processing in the visual system evaluated using functional MRI. *Neurology* 48:1416–1420.
- Landy MS, Bergen JR (1991) Texture segregation and orientation gradient. *Vision Res* 31:679–691.
- 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 [Biol]* 240:251–278.
- Logothetis NK, Sheinberg DL (1996) Visual object recognition. *Annu Rev Neurosci* 19:577–621.
- 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, Itzhak Y (1998) Rapid shape adaptation reveals position and size invariance. *NeuroImage* 7:S43.
- 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.
- Nakayama K, Shimojo S, Silverman GH (1989) Stereoscopic depth: its relation to image segmentation, grouping and the recognition of occluded objects. *Perception* 18:55–68.
- Pasupathy A, Connor CE (1999) Responses to contour features in macaque area V4. *J Neurophysiol* 82:2490–2502.
- Price CJ, Moore CJ, Humphreys GW, Frackowiak RSJ, Friston KJ (1996) The neural regions sustaining object recognition and naming. *Proc R Soc Lond [Biol]* 263:1501–1507.
- Rosen BR, Buckner RL, Dale AM (1998) Event-related functional MRI: past, present, and future. *Proc Natl Acad Sci USA* 95:773–780.
- Sakata H, Taira M, Kusunoki A, Tanaka Y (1997) The parietal association cortex in depth perception and visual control of hand action. *Trends Neurosci* 20:350–357.
- Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260:995–997.
- Sereno AB, Maunsell JH (1998) Shape selectivity in primate lateral intraparietal cortex. *Nature* 395:500–503.
- Shimojo S, Silverman GH, Nakayama K (1988) An occlusion-related mechanism of depth perception based on motion and interocular sequence. *Nature* 333:265–268.
- Talairach J, Tournoux P (1988) *Co-planar stereotaxic atlas of the human brain*. New York: Thieme.
- Tanaka K (1996) Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109–139.
- Tong F, Nakayama K, Vaughan JT, Kanwisher N (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21:753–759.
- Vanni S, Revonsuo A, Hari R (1997) Modulation of the parieto-occipital alpha rhythm during object detection. *J Neurosci* 17:7141–7147.
- Vogels R (1999) Effect of image scrambling on inferior temporal cortical responses. *NeuroReport* 10:1811–1816.
- Wiggs CL, Martin A (1998) Properties and mechanisms of perceptual priming. *Curr Opin Neurobiol* 8:227–233.
- Wojculik E, Kanwisher N (1999) The generality of parietal involvement in visual attention. *Neuron* 23:747–764.