Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex


Massachusetts General Hospital–Nuclear Magnetic Resonance Center, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114

Communicated by Bela Julesz, Rutgers University, Piscataway, NJ, May 22, 1995 (received for review November 10, 1994)

ABSTRACT The stages of integration leading from local feature analysis to object recognition were explored in human visual cortex by using the technique of functional magnetic resonance imaging. Here we report evidence for object-related activation. Such activation was located at the lateral-posterior aspect of the occipital lobe, just abutting the posterior aspect of the motion-sensitive area MT/V5, in a region termed the lateral occipital complex (LO). LO showed preferential activation to images of objects, compared to a wide range of texture patterns. This activation was not caused by a global difference in the Fourier spatial frequency content of objects versus texture images, since object images produced enhanced LO activation compared to textures matched in power spectra but randomized in phase. The preferential activation to objects also could not be explained by different patterns of eye movements: similar levels of activation were observed when subjects fixated on the objects and when they scanned the objects with their eyes. Additional manipulations such as spatial frequency filtering and a 4-fold change in visual size did not affect LO activation. These results suggest that the enhanced responses to objects were not a manifestation of low-level visual processing. A striking demonstration that activity in LO is uniquely correlated to object detectability was produced by the “Lincoln” illusion, in which blurring of objects digitized into large blocks paradoxically increases their recognizability. Such blurring led to significant enhancement of LO activation. Despite the preferential activation to objects, LO did not seem to be involved in the final, “semantic,” stages of the recognition process. Thus, objects varying widely in their recognizability (e.g., famous faces, common objects, and unfamiliar three-dimensional abstract sculptures) activated it to a similar degree. These results are thus evidence for an intermediate link in the chain of processing stages leading to object recognition in human visual cortex.

Extensive research involving single-unit recording, anatomical studies, and behavioral experiments in the monkey (for review see, e.g., ref. 1) have suggested that object recognition is a multistage process leading progressively from localized feature analysis in primary visual cortex, through a sequence of cortical areas, to more global object recognition in the inferotemporal cortex. More recently, studies in human visual cortex using positron emission tomography (PET) revealed selective activation of ventral and temporal regions associated with face recognition and attention to shapes (2–4). However, the exact location and nature of the areas contributing to object recognition in the human cerebral cortex remain unclear.

Using functional magnetic resonance imaging (fMRI) (5–8), which has relatively high spatial resolution and allows repeated testing, we looked for potential cortical “building blocks” participating in the object recognition process.

METHODS

This study is based on fMRI scans conducted in 16 normal adults. Functional brain activation was mapped in a 1.5-T General Electric scanner equipped with echo planar imaging (Advanced NMR Systems, Wilmington, MA). Informed consent was obtained from all subjects about the nature and possible consequences of the study.

MRI Parameters. Asymmetric spin echo sequences were used (echo time TE = 80 msec; repetition time TR = 2–3000 msec; offset = 25 msec; field of view = 40 cm; matrix = 128 × 64) to minimize the macrovascular signal contribution (9). In one case, the results were confirmed with an inversion recovery spin echo sequence (5) (inversion time TI = 1100 msec; TE = 43 msec) which provides improved large vessel suppression but lowered sensitivity. In most cases a 5-inch (127-mm) surface coil was placed posterolaterally to enhance signal detection. Voxel (volume element) resolution for the functional mapping was 3 × 3 × 5–6 mm. In addition, high-resolution anatomic [1.6 × 1.6 × 5–6 mm; T1 (longitudinal relaxation time) weighted] and flow-compensated sequences (with increased vascular sensitivity) were obtained for each scanning session so that they could be compared with the functional activation patterns.

Experimental Protocol. Computer-generated visual stimuli were projected onto a tangent screen mounted in front of the subject’s eyes and viewed via a tilted mirror, providing approximately 80 × 60 degrees of visual angle. A typical scanning session consisted of 5–12 scans; in each scan 960–1020 images were collected over 240–540 sec from six or seven contiguous brain slices. Each scan was divided into 11–12 epochs; in each epoch one set of 14 or 15 different pictures (either objects or textures, see below) were presented. Pictures were changed at a rate of 0.5 Hz, without intervening blank periods. All stimuli in these tests were stationary and achromatic. Epochs containing pictures alternated with epochs containing either a fixation point or a blank visual field.

Data were analyzed by using a voxel-by-voxel Kolmogorov–Smirnov statistic (10). Activation time courses from regions of interest defined as highly significant by the statistical analysis were normalized by setting the highest and lowest signal strength values in the time course to 100% and 0%, respectively, and averaging across subjects (10).

Picture Categories. To search for cortical areas involved in object detection we relied on tests that compared cortical activation in response to pictures of a variety of objects.

Abbreviations: MR, magnetic resonance; MRI, MR imaging; fMRI, functional MRI; LO, lateral occipital complex; 3D and 2D, three- and two-dimensional; SOR, selective object response index.

*To whom reprint requests should be sent at the present address: Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel 76100.

8135
contrasted with pictures of various texture patterns. Detailed description of these and related stimuli used in this study are as follows:

**Objects.** (i) Real-life objects. These objects were real-life, achromatic, stationary pictures subtending $40 \times 40 \pm 10$ degrees of visual angle. The pictures were of three main types: (a) faces of famous people, mainly well-known actors and politicians; (b) common recognizable objects such as tools, toys, clothing, animals, and plants; and (c) unfamiliar three-dimensional (3D) abstract sculptures (from the works of Henry Moore). (ii) Filtered pictures. In some experiments we presented pictures from categories $a$ and $b$ of the real-life objects but removed either the low spatial frequencies in the pictures (high-pass pictures, cut-off frequency 0.25 cycles/degree) or the high spatial frequencies (low-pass pictures, cut-off frequency 0.25 cycle/degree). (iii) Degraded pictures. In another experiment, increasing amounts of visual noise were added to pictures of famous faces. Visual noise was introduced in stepwise iterations. In each iteration the gray level of each pixel in the picture was changed from the previous iteration by an equally weighted randomly chosen value ranging from 0 to $\pm 50\%$ of the full black-to-white range. Three levels of noise were generated: (a) Moderate noise level, which consisted of three visual noise iterations as described above. Qualitatively, this amount of noise degraded the faces so that their individual "semantic" identity was lost. (b) High noise level, which consisted of five visual noise iterations, making the faces barely detectable as objects. (c) Pure noise, which consisted of a random distribution of gray levels across the pixels.

**Textures.** (i) Scrambled phase textures. These textures were designed as Fourier power-spectrum controls for the object pictures. They were constructed by performing a two-dimensional (2D) Fourier transformation on high-pass filtered pictures of the same objects. The resultant phase map was scrambled by adding to each phase angle a random value in the range $\pm 0.75\%$ radians. Using these randomized phase angles, we performed an inverse 2D Fourier transformation using the unchanged power spectrum derived from the objects. In this way a series of textures were generated whose power spectrum was precisely matched to that of the high-pass filtered objects. (ii) High-contrast scrambled phase textures. These patterns were identical to $i$, but their absolute contrasts (measured as the number of saturated white and black pixels in the picture) were matched to those of the real-life objects. (iii) "Wallpaper" textures. These high-contrast textures were obtained from a commercially available library (Aldus SUPERPAINT 3.0). They consisted of full-field patterns of repeated or randomly presented simple geometrical shapes (straight and curved lines, dots, gratings, squares, etc.).

**RESULTS**

Areas involved in object detection were explored by searching for preferential cortical activation in response to pictures of objects, compared with texture patterns. We chose to compare objects to textures because textures contain many visual features common to objects, without appreciable chance of producing significant object perception. In the occipital lobe, such objects-vs.-textures tests highlighted a localized, inferolateral region in most (15 of 16) subjects studied. At present we are uncertain whether this region is a single area or a cluster of several areas, so we refer to it as the lateral occipital complex (LO).

The level of LO selectivity across subjects varied to some extent, ranging from almost a complete lack of texture activation in some subjects to a moderate object bias in others. To obtain a quantitative measure of the object bias we calculated the "MR activation ratio." This index was defined as follows: activation ratio (epoch a/epoch b) = (averaged MR signal during epoch a minus baseline MR signal)/(averaged MR signal during epoch b minus baseline MR signal). For the object vs. texture MR activation ratio, epoch a contained real-life familiar and unfamiliar objects, while epoch b contained "wallpaper" textures (see Methods for details). The group average MR activation ratio for the object vs. textures test in LO measured in 14 subjects was $2.32 \pm 0.61$ (1 SD).

Anatomically, LO was located posteriorly on the lateral aspect of the fusiform gyrus, branching more anteriorly into ventral and dorsal branches reaching to the anterior border of Brodmann's areas 19 and 37, respectively. The Talairach coordinates for LO (11) were measured in six subjects. The coordinates of the point at which the two branches emerge were lateral, 42.8 $\pm$ 2.7 mm; posterior, $-72.7 \pm 8$ mm; and inferior, $-18.2 \pm 9.8$ mm (mean $\pm 1$ SD; $n = 6$). Most of the data reported here were obtained by using a posterolaterally placed surface coil that effectively limits measurements to one hemisphere. However, in five cases we used a head coil (two cases) or a bilateral surface coil (three cases), and in these cases some bilateral activation was observed in LO.

Since LO appeared to be located close to the motion-selective area MT (also known as V5) (12), we studied its topographical relationship to LO by testing within the same sessions for selective activation to moving vs. stationary random dot stimuli. This test has been used in several studies as a probe for the location of area MT/V5 (10, 12). A typical result, obtained in a single scanning session, is shown in Fig. 1. LO (red) was defined as the region with a highly significant preferential activation for objects vs. textures ($P < 0.01$, Kolmogorov–Smirnov nonparametric statistic), while area MT/V5 (green) was defined as the region with the highest selectivity among all the cortical areas showing preferential activation in response to moving vs. stationary stimuli (10). The object-biased LO can be seen to directly adjoin area MT/V5 (green). This topographic relationship of LO to area MT/V5 was maintained across all subjects studied ($n = 5$). Note the two branches of LO (Fig. 1A, arrows).

While the statistical maps provide sensitive localization of the different cortical areas, they present a very limited view of the functional characteristics of these areas. For more detailed analysis of the functional profiles of LO and other visual areas, we plotted the time courses of activation obtained from these areas during presentation of various categories of objects and textures (see Methods for detailed descriptions of the picture categories).

An important issue is whether the preferential object activation of LO is simply a consequence of differences in the spatial frequency content of objects and textures, since the Fourier power spectrum of textures is generally biased toward higher frequencies, compared with that of objects. Fig. 24 shows the results of an experiment in which this question was addressed. The experiment consisted of five epochs, containing the following picture categories: 1, pictures of objects and faces, high-pass filtered; 2, scrambled phase textures, having Fourier power spectra identical to those of the images in 1 but with randomized phase relationships; 3, same objects and faces as in 1, but unfiltered; 4, same textures as in 2, but with their absolute contrast increased to match the contrast of the objects and faces in 3; and 5, "wallpaper" textures made up of high-contrast repeating simple elements.

Fig. 24 compares the time course of activation in area V1 (blue) and LO (red) during the presentation of the five picture categories. The results were obtained in single experiments performed in four subjects. The graph shows normalized time courses averaged across the four subjects. The center of each error bar indicates the averaged level of activation during each epoch, while the size of the error bar indicates $\pm 1$ SEM for the averaged epoch activation across the four subjects. In contrast to area V1, LO showed markedly increased activation whenever pictures of objects (both real-life and filtered) were
presented. Activation to texture patterns, including those whose power spectra and contrast matched the pictures of objects, was substantially reduced in LO. These results indicate that the differential activation to objects vs. textures observed in LO is not due to a simple difference in power spectra or contrast of these stimuli.

This conclusion was further substantiated in an experiment in which we tested the effect of spatial frequency filtering of pictures of faces on LO activation. In the spatial frequency test, one epoch (epoch a) included presentation of high-pass filtered pictures of faces, while the other (epoch b) included presentation of low-pass filtered pictures of faces. Both filtered sets were still recognized as faces. Calculating the MR activation ratio for the two conditions showed that the average high-frequency/low-frequency activation ratio was 0.99 ± 0.05 \( (n = 2) \), indicating that the spatial frequency content of the images did not significantly affect LO activation, within the range tested (see Methods for details).

Another trivial explanation for the differential activation to objects and textures presumes a difference in the patterns of eye movements: subjects could use different scanning strategies when viewing the objects and textures, thus leading to the observed differential activation. Effects of eye movements were tested in three subjects by comparing LO activation in two epochs in which identical pictures (famous faces) including a fixation point were presented. The subjects were instructed to fixate during one epoch, and to vigorously scan the pictures with their eyes during the other epoch. The LO activation ratio for the eyes fixated vs. eyes moved epochs was 0.99 ± 0.02 (1 SEM), indicating that the patterns of eye movements do not significantly affect LO activation (see also ref. 13).

These results suggest that the object-biased LO activation cannot be attributed to some low-level aspect associated with object images. This activation may thus be related to the process of object recognition or identification. To test the possible role of recognition in LO activation, we performed an experiment in which pictures of objects having different degrees of object familiarity were presented. The experiment consisted of five picture categories that were presented in the following order (see Methods for detailed description of the categories): 1, common objects (animals, tools, etc.); 2, “wallpaper” textures; 3, unfamiliar objects (3D abstract sculptures); 4, “wallpaper” textures; and 5, famous faces. Fig. 2B shows the results of this experiment obtained in five subjects. The graph shows activation time courses obtained in area V1 (blue) and LO (red).

As in the previous experiment, LO (in contrast to V1) showed enhanced activation when objects were presented, compared to textures. When activation across the different picture categories is compared, it becomes clear that object familiarity did not affect the level of activation. Both similar objects and famous faces produced levels of activation that were similar to those produced by abstract, unfamiliar, sculptures. Thus, it appears that the process of “semantic” object identification does not affect LO activity.

To see how object detectability affected LO activation, increasing amounts of visual noise were added to one class of objects, in this case pictures of famous faces. Fig. 2C compares activation time courses for area V1 and LO. Four categories were shown in this experiment: 1, original faces without noise; 2, three visual noise iterations (see Methods), making the faces unrecognizable; 3, five visual noise iterations making the faces barely detectable as objects; and 4, pure noise (pixels with randomized gray levels). In area V1 the visual noise actually increased the activation, while LO showed a marked decrease. The LO response appeared surprisingly sensitive, so that even when subjects could barely discern any target, the response was still above pure noise activation.

A striking demonstration of the departure of LO from low-level feature-based activation was provided by the “Lincoln” illusion (14). In this illusion, blurring of objects digitized into large blocks restores their detectability. We studied brain activation associated with this illusion in five subjects. The experiment consisted of four epochs: 1, pictures of objects (mostly faces) broken into large blocks; 2, same pictures as in epoch 1, but blurred by low-pass filtering; 3, real-life pictures of these objects; and 4, meaningless blocked patterns, matched in number and gray levels to those in epoch 1.

To measure the activation across subjects in each epoch, we defined a selective object response index (SOR), which was calculated as follow: \( SOR = 100 \times \) (averaged MR signal during an epoch minus averaged MR signal during the meaningless blocks epoch)/(averaged MR signal during the meaningless blocks epoch minus baseline signal). Baseline MR signal was...
The signal recorded during blank or fixation point epochs. The SOR provides a measure of the pure object-related response as a fraction of the nonselective visual response. Error bars indicate ± SEM for the five subjects in which this study was conducted.

Fig. 3 shows the SOR for epochs 1–3. Note that blurring of the blocked images, which increases perceptual recognizability, also produced a marked increase in LO activation (filled bars), to a level similar to that produced by the real-life pictures. Neighboring cortex on the posteromedial side of LO (unfilled bars) as well as other occipital lobe areas (not shown) did not increase its activation in correlation with this illusion.

It could be argued that the enhanced LO activation produced during the Lincoln illusion may be due simply to the act of blurring of picture blocks, unrelated to the increased detectability of objects produced by this blurring. To test this possibility we conducted a separate experiment in which we measured, in four subjects, LO activation in response to random gray-level blocks similar in number and size to the ones used during the Lincoln illusion experiment, both unfiltered and blurred. The result showed that blurring of random blocks produced only a slight and highly variable change in LO activation [SOR = 3.7 ± 9.87 (SEM), n = 4].

They were alternated with epochs containing “wallpaper” texture patterns. In this experiment subjects were instructed to name the familiar objects, the faces, and the geometrical elements in the “wallpaper” texture patterns. Note that LO (red) activated equally well to images of familiar and unfamiliar objects, while showing reduced activation to textures. (C) Object degradation. In this experiment, pictures of faces were gradually degraded by adding visual noise (epoch 1 = no noise; epoch 4 = pure noise; see Methods). Note the contrasting time courses of V1 (which showed an increased activation to visual noise) and LO (which showed a clear reduction). Baseline activation (not shown) slightly increased in LO.
Finally, the effect of changing object size was tested by comparing LO activation to line drawings of common objects. Line-drawing size was 40 ± 5 degrees on a side in the "large-size" epoch and it was reduced by a factor of 4 in linear scale in the "small-size" epoch. The calculated large/small activation ratio for LO was 1.003 ± 0.025 (n = 3, 1 SEM), indicating that LO activation was size insensitive within the range tested here.

**DISCUSSION**

The main finding of the present work is the discovery of preferential and localized activation to pictures of objects relative to texture patterns, random visual noise, or phase-randomized control stimuli. This preferential activation is located at the lateral occipital corner of the hemisphere, in a region termed LO. Several control experiments indicate that this object-biased activation cannot be explained by factors such as the global Fourier power spectrum, absolute contrast, or eye movements. Furthermore, the enhanced activation can be elicited even by blurring images of blocked objects as demonstrated in the Lincoln illusion. This evidence suggests that LO activation is involved in the process of object detection. However, our results also indicate that LO activation does not differentiate familiar from unfamiliar objects, and so this region is apparently not involved in the final, more "semantic" or "cognitive" stages of the recognition process.

The present results do not necessarily imply that LO is activated only by pictures of complete objects. It could be that LO neurons are driven by certain image features embedded in the objects but not in the textures, or a collection of object "primitives." In that respect, the recent results of Tanaka and co-workers (for review, see ref. 1), which documented sensitivity to "moderately complex" visual features in neurons in the object-processing stream (areas V4, TEO, and the IT cortex), are very relevant. It could be that LO neurons have receptive field properties similar to those of the neurons described by Tanaka and co-workers, and that pictures of objects activate the LO more strongly since they contain a richer source of complex features, compared to the more repetitive or less naturalistic texture patterns.

The limited spatial resolution of the fMRI technique may mask further subdivisions within the functionally defined regions described here. Indeed, some heterogeneity was occasionally observed within LO. Studies using a higher spatial resolution are required to establish whether this heterogeneity reflects the existence of several related areas within LO. Functional subdivisions in cortical areas involved in object perception have been suggested recently by human cortical recording and stimulation studies (15).

Concerning the issue of blood vessel artifacts affecting the fMRI mapping, our control using the inversion recovery spin echo sequence, which is insensitive to macrovascular oxygenation (5) as well as the relatively large size of LO, suggests that, within the present spatial resolution, large vessel contribution did not significantly distort the basic results.

The possible homology of LO to mapped areas in the macaque is not yet clear. On the basis of its topography and relationship to MT/V5, one possibility is that human LO is homologous to monkey V4 and possibly PIT/TEO. In this case it may be part of a stretch of cortex extending ventromedially from LO to area V4v described recently (16). Alternatively, it may be that higher-order visual processing, performed in macaque area TE, is shifted to a more posterior cortical sites such as LO in humans (2).

The notion that the visual cortex is organized into two major "streams"—a ventral stream concerned with object vision and a dorsal stream concerned with spatial localization—has received substantial experimental support in recent years (e.g., ref. 17). The finding reported here, of adjoining areas in the occipital lobe, one specializing for visual motion (area MT/ V5) and one for object detection (LO), provides a striking demonstration of such specialized processing in the human visual system. Extrapolating from the macaque maps (e.g., ref. 1), LO may feed into areas more ventral and anterior, implicated in more elaborate aspects of object recognition.

We thank K. Gil-Spector, T. L. Davis, M. Vevea, and T. Campbell for software and technical support and Y. Citri, S. Edelman, C. R. Olson, and S. Ullman for helpful comments. This work was supported by National Institutes of Health Grants EY07980, RO1MH50054, PO1CA48729, HL39810, and CA40303 and by the Perera Fight for Sight grant.