Coding of Object Shape in the Lateral Occipital Complex (LOC)

Eshed Margalit

Neuroscience Program Honors Thesis

Advisor: Professor Irving Biederman

Keywords

LOC, object familiarity, object representation, between-part relations, part shape, Lateral Occipital Cortex, Posterior Fusiform, novel objects, fMRI.
1. ABSTRACT

The Lateral Occipital Complex (LOC), an area of human cortex critical for shape perception (James et al. 2003), is comprised of the Lateral Occipital Cortex (LO) and the Posterior Fusiform gyrus (pFs). It is localized as the region that shows greater BOLD activity when viewing intact objects compared to their pixel-scrambled (texturized) versions. But image scrambling a) eliminates object familiarity, b) destroys the integrity of the parts, and c) leaves the relations among the parts undefined. To assess the role of these factors in LOC activation while maintaining strict control for low-level features, we employed computer-modeled objects. By varying the relations among the parts we could depict either familiar or novel intact objects. We also created images with scattered parts by disconnecting contiguous parts, allowing us to assess whether the effects of scrambling an image can be attributed to the loss of parts, relations, or both. The activation to intact objects was greater than the activation to scattered objects, which in turn was greater than to the scrambled objects. The difference in BOLD response between intact and scattered was equivalent to the difference between scattered and scrambled, showing that the loss of part shape and the loss of interpart relations contribute equally to the greater activation to intact objects in LOC.

2. INTRODUCTION

2.1 Background

2.1.1 LOC Function

The lateral occipital complex (LOC), a cortical area critical for object recognition (James et al., 2003), was shown by Malach et al. (1995) to have a greater BOLD response to visual presentation of intact objects than to phase-scrambled objects which resemble patterned texture. To discount the potential influence of familiarity (intact objects are recognizable while texture is not), Malach demonstrated that abstract sculptures, which were unfamiliar to subjects, also produced greater activation in LOC than the scrambled images. Malach concluded that there was no effect of familiarity in LOC, although the low-level visual differences between stimulus categories (familiar objects and unfamiliar sculptures) were not controlled. Margalit et al. (in press) replicated Malach’s finding by comparing activation from computer-modeled familiar objects to activation from the same objects with the parts rearranged into an unfamiliar configuration, thus preserving low-level properties (such as luminance) and part composition across conditions of familiarity. With these stricter controls, no net result of familiarity was seen in LOC, although the possibility of response differences in the spatial pattern of voxel activation within the region was not ruled out. While this result suggests that object familiarity does not impact BOLD response magnitude in LOC, it remains unclear whether the preferential response to intact objects (both familiar and novel) compared to scrambling is due to the shape of the parts or the relations between parts. That it is shape that is critical for LOC activation—rather than surface properties—is documented by the equivalence in the maintenance of adaptation when line drawings depicting only the orientation and depth discontinuities of photographed objects are viewed compared to a re-presentation of the original photograph (Grill-Spector, Kourtzi, & Kanwisher, 2001).

2.1.2 Response of LOC to object parts and relations
The neural representation of objects in LOC has been shown to be based primarily on the objects’ parts and relations (Hayworth & Biederman, 2006; Hayworth et al., 2009; Lescroart & Biederman, 2012), particularly in fMRI adaptation (fMRIa) paradigms. fMRIa takes advantage of neuronal adaptation—the tendency of neuronal populations to reduce their activity level when the same stimulus is repeatedly presented. When a new stimulus is presented after the onset of adaptation, however, neural activity levels rapidly recover—a “release” from adaptation that is captured in the BOLD signal. When images are presented as line drawings, LOC activity demonstrates a release from adaptation when lines corresponding to an object’s parts (e.g. the wing of a plane) are reorganized or removed, but not when smaller deletions of line segments and vertices (e.g. one edge of the plane’s wing) were made. This result is consistent with earlier behavioral studies (Biederman & Cooper 1991) showing that visual priming, as measured by naming speed and accuracy of line drawings, was facilitated by pairs of images which differed in local edges and vertices. Similar priming was not seen when pairs of images differed in parts, suggesting the sensitivity of LOC to object parts, instead of local features.

Sensitivity to variation between object parts has been shown by Lescroart and Biederman (2012) who showed that a Multivoxel Pattern Analysis (MVPA) performed on voxels in LOC could reliably distinguish different arrangements, defined by different medial-axes configurations of the same geons, despite variations in the overall orientations of the objects and the participants judging the shapes of the parts rather than their configuration.

2.1.3 Functional Subdivisions

LOC is composed of two distinct cortical areas: the lateral occipital cortex (LO) and the posterior fusiform gyrus (pFs), as shown in Fig 1.
While they are sometimes considered homogenous in function, Hayworth & Biederman’s (2006) finding that pFs shows smaller releases from adaptation than LO to a change in local features (line segment deletions) suggests that pFs may encode higher-level features than LO. The role of LO and pFs in the preferential response to intact, three-dimensional objects, however, has not been explicitly delineated.

2.2 Approach

To determine whether the preferential response of LOC to intact than to scrambled objects was due to the disruption of the shape of the individual parts or to the disruption of the relations between parts, we created a stimulus set which “scattered” an object’s parts by separating, translating, and rotating the parts (Fig. 2C). If activation to the scattered objects is equivalent to that of the corresponding intact images (familiar or novel), then part shape is the critical modification in scrambling. If, however, activation to scattered objects is equivalent to that of scrambled objects, then part relations are the critical variable. Furthermore, these potential outcomes may spatially differ between LO and pFs. To investigate the selectivity of LOC to shape and part relations, we conducted an fMRI study in which subjects were shown 3D
rendered images of familiar objects, unfamiliar objects, scattered objects, and scrambled objects (see Methods for stimulus details).

![Figure 2: Sample stimuli from the experiment. A) Two examples of Familiar stimuli, a paint roller and a house. B) Novel, or unfamiliar, version of the Familiar stimuli. C) Scattered stimuli. D) Scrambled stimuli.](image)

3. METHODS

3.1 Subjects

Thirteen right-handed university students (twelve from the University of Southern California, 6 females, mean age 20.7 years, age range 18 to 27) participated in the study. All subjects were screened for safety and provided informed consent in accordance with the University of Southern California’s Institutional Review Board Guidelines.

3.2 Stimuli

In the “Familiar” condition, stimuli were simple 3D-rendered objects composed of geons which subjects reliably identified as familiar and nameable (Fig 2A). In the “Novel” condition, the parts composing the Familiar objects were rearranged such that they formed an unfamiliar object (Fig 2B). In the “Scattered” condition, the object parts were separated in space, breaking the contiguity of the object (Fig 2C); that is, object parts appeared dissociated from one another. Stimuli in the “Scrambled” condition were the result of permuting random blocks of 3x3 pixels from the familiar object images (Fig 2D). All non-scrambled stimuli were created in Blender (Stichting Blender Foundation) and the scrambled stimuli were created in MATLAB. All rendered stimuli had no apparent texture, but were designed to display realistic lighting and shadows.
3.3 Procedure and Experimental Design

3.3.1 Structural Scans

A high-resolution (0.8mm x 0.8mm x 0.8mm) T1-weighted structural scan was conducted for each subject for the purpose of within- and between-subjects registration during fMRI analysis. In accordance with the University of Southern California’s Institutional Review Board Guidelines, T2-weighted scans were also collected for the purposes of radiological evaluation. Magnetic field maps were collected at the end of each scan session to map magnetic field inhomogeneity, allowing us to mitigate distortion artifacts in the functional scans.

3.3.2 Functional Scans

In addition to the structural scans, each subject participated in four functional runs: the first two runs followed a block design (340 seconds each) and the last two runs followed a rapid event-related design (560 seconds each). Subjects were encouraged to take one-minute breaks between each run but had the option to begin each run at their convenience. Each block design run consisted of 17 blocks, counter-balanced across the four conditions—familiar, novel, scattered, and scrambled—with a balancing look-back of 1 block (Fig 2). The first block of each run was excluded from the statistical analyses. Within each block, 10 stimuli were presented for 1667ms each and followed by a fixation cross presented for 333ms (inter-stimulus interval; ISI), resulting in 20 second blocks.

For the two event-related runs, stimuli from five conditions—familiar, novel, scattered, scrambled, and null—were presented in counterbalanced order with a balancing look-back of 2 trials. Each stimulus was shown for 500ms and followed by a fixation cross which remained on screen for 2, 3, or 4 seconds (ISI order was pseudorandomized and counterbalanced). In the null condition, only the fixation cross was presented, so null trials effectively extended the ISIs. For instance, if the ISI for a null trial was 3 seconds, then the fixation cross would appear for a total of 3.5 seconds for that trial.

For all functional runs, subjects were instructed to maintain focus on the stimuli, which were presented in the center of the screen surrounded by a rectangular colored border. Subjects were instructed to respond as quickly and accurately as possible by button press when the surrounding
border changed color, without moving their eyes away from central fixation. In the block design
runs, the border changed color once per block, with the change occurring at a random point
within each block. In the event-related runs, the border changed color with 25% probability on
any given trial. This orthogonal task was designed to maintain subject attentiveness while
viewing the stimuli without any explicit stimulus-dependent task thus relying on the automaticity
of shape processing (Smith & McGee, 1980). The experimenter monitored eye movements to
ensure stability and centrality of gaze. Stimuli were presented in MATLAB using the
PsychToolbox package (Brainard 1997; Pelli 1997).

3.3.3 Data Acquisition and Imaging Parameters

Data were collected at the Dana and David Dornsife Cognitive Neuroscience Imaging Center at
the University of Southern California using a SIEMENS 3T MAGNETOM Prisma with a 32-
channel head coil. Responses were collected using an MRI-compatible button box. High-
resolution (0.8 x 0.8 x 0.8 mm) T1-weighted images were collected using an MPRAGE
sequence. Functional T2*-weighted images (2.0 x 2.0 x 2.0 mm) were collected using an echo
planar imaging (EPI) sequence with TR = 1000ms, TE = 35ms, and a multi-band acceleration
factor of 8.

3.4 Data Analysis

3.4.1 Whole Brain Voxel-wise Analysis

FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00,
part of FSL (FMRIB’s Software Library, www.fmrib.ox.ac.uk/fsl). Slice interpolation was used
for slice timing correction, MCFLIRT was used to generate 3D motion parameters for
regression, and spatial smoothing was set to a 5mm full-width at half-maximum. Data was high-
pass filtered at a cutoff of 80 seconds (the duration of four blocks). Z (Gaussianised T/F) statistic
images were thresholded using clusters determined by Z>2.3 and a (corrected) cluster
significance threshold of P=0.05 (Worsley 2001). FILM (FMRIB’s Improved Linear Model)
prewhitening was used to provide a robust and accurate nonparametric estimation of time series
autocorrelation on each voxel’s time series (Woolrich, 2001). Registration to high resolution
structural and MNI standard space images was carried out using FLIRT (Jenkinson, 2002).

3.4.2 ROI Analysis

Functional regions of interest (ROIs) were defined separately for each subject. Four bilateral
ROIs were defined: LO as defined by the Familiar minus Scrambled contrast, LO as defined by
the Novel minus Scrambled contrast, pFs as defined by the Familiar minus Scrambled contrast,
and pFs as determined by the Novel minus Scrambled contrast. For all contrasts, clusters often
extended into retinotopic areas V2, V3, and V4, as defined by the PALS visuotopic annotations
(Van Essen, 2005) in Freesurfer. Thus, to isolate the LOC ROI, clusters were re-thresholded
beyond the initial Z = 2.3 cutoff, such that clusters extending into areas V2, V3, and V4 could be
distinct from clusters near areas LO and pFs. This process was conducted iteratively by
inspection for each subject. The resulting maps contained a group of clusters that included areas
LO and pFs but did not extend into other retinotopic areas, as well as a separate group of clusters
that included retinotopic areas but did not include LO or pFs. For each contrast, all clusters
belonging to an ROI (LO or pFs) but not any retinotopic area were combined into a single ROI.
ROIs defined by each contrast (Familiar minus Scrambled and Novel minus Scrambled) were
combined disjunctively to create a single LO ROI and a single pFs ROI for each subject, such that voxels in the final ROIs were members of the Familiar minus Scrambled ROI, the Novel minus Scrambled ROI, or both. For ROI statistical analyses, mean activation values in terms of % BOLD signal change within the ROI for each contrast were compared. The ROI analyses are restricted to the 8 subjects for whom we were able to define both LO and pFs. The definition of areas LO and pFs was consistent across subjects as shown in Fig 4, where brighter areas correspond to overlap between individual subject ROIs.

Figure 4: Individual subject ROIs projected into the MNI152 standard space. LO is represented in blue, and pFs is represented in red. Voxel brightness corresponds to the number of subjects who had that voxel included in their individual ROI, such that brighter areas indicate higher overlap.

Because there is no net effect of familiarity in LOC (Margalit et al., In press; Fig 5), we collapsed Familiar and Novel stimuli into an “Intact” category in ROI-based analyses for ease of interpretation.
Figure 5: ROI analysis results from Margalit et al. (In press). There was no significant difference in % BOLD signal change between the Familiar and Novel conditions.

3.4.3 Peri-stimulus time course extraction

The hemodynamic response function (% BOLD signal change as a function of time) was approximated at 26 time points (2 before stimulus onset and 24 following stimulus onset) in 1-second intervals using a finite-impulse-response (FIR) model. Activation values were extracted from the ROIs defined above, with variance weighted by the effective spatial resolution induced by spatial smoothing. To isolate the peri-stimulus time course for each stimulus condition, we compared LOC activity in each stimulus condition to baseline activity while viewing a fixation cross (e.g. Scattered minus Fixation).

4. RESULTS

4.1 Whole-Brain Voxelwise Analyses

4.1.1 Familiar, Novel, and Scattered minus Scrambled contrasts
The LOC was reliably localized by the Familiar minus Scrambled contrast (Fig 6), the Novel minus Scrambled contrast (Fig 7), and the Scattered minus Scrambled contrast (Fig 8).

Figure 6: Statistical activation maps (horizontal slices) for the Familiar > Scrambled contrast. Activation is apparent in areas LO and pFs.
Figure 7: Statistical activation maps for the Novel > Scrambled contrast. Activation appears similar in location but more robust than in the Familiar > Scrambled contrast shown in Fig 6.
Figure 8: Statistical activation maps for the Scattered > Scrambled contrast. Activation appears similar in location to the Familiar > Scrambled contrast (Fig 6) and Novel > Scrambled contrast (Fig 7).

4.1.2 Effects of Familiarity

No voxel clusters in the entire acquisition volume showed greater activation to Familiar stimuli than to Novel stimuli, suggesting no preferential effect of familiarity on BOLD responses. The left occipital pole and the right parietal lobe did, however, demonstrate larger BOLD responses for Novel stimuli than for Familiar stimuli (Fig 9), suggesting that these areas respond more strongly to Novel stimuli than to Familiar stimuli.
Figure 9: Statistical activation maps for the Novel > Familiar contrast. Activation is apparent in the left occipital pole and the right parietal lobe.

4.1.3 Effects of Scattering Parts

The Novel > Scattered contrast yielded significant clusters of activation bilaterally in area LO (Fig 10).
The Scattered > Familiar and Scattered > Novel contrasts did not yield any significant clusters of activation anywhere in the brain for the voxel-wise analysis, indicating that BOLD responses to Intact objects were higher than or equivalent to BOLD responses to Scattered objects.

No voxel clusters had higher BOLD responses for Scattered stimuli than to Familiar or Novel stimuli, suggesting that, without a more sensitive analysis, we cannot conclude that Intact stimulus categories (Familiar and Novel) evoked larger or smaller BOLD responses than the Scattered stimuli.

4.2 Region-of-Interest (ROI) Analyses

In both ROIs (LO and pFs), a one-sample t-test relative to 0% revealed that the % BOLD signal change was significantly greater than 0 for the Intact minus Scrambled contrast in LO, \( t(15) = 8.24, p < 10^{-7}, \) Cohen’s \( d = 4.25, \) and pFs, \( t(15) = 6.07, p < 10^{-5}, \) Cohen’s \( d = 3.13, \) indicating preferential activation to shapes compared to texture in both regions (Fig 11, green line).

Response magnitude to Scattered stimuli was significantly higher than to Scrambled stimuli in both LO, \( t(7) = 5.03, p = 0.002, \) Cohen’s \( d = 3.80, \) and pFs, \( t(7) = 3.60, p = .01, \) Cohen’s \( d = 2.72, \) indicating preferential activation to intact part shape than to pixel-scrambled texture in both regions (Fig 11, red line). % BOLD signal change was also significantly greater than 0 for the Intact minus Scattered contrast in LO, \( t(15) = 4.52, p = 0.0004, \) Cohen’s \( d = 2.33, \) and pFs, \( t(15) = 4.12, p = .0009, \) Cohen’s \( d = 2.13, \) indicating preferential activation to interpart relations than to isolated object parts in both regions (Fig 11, blue line).

The signal change for the Scattered minus Scrambled contrast was not significantly different from that of the Intact minus Scattered contrast in LO, \( t(22) = 1.99, p = 0.06, \) or in pFs, \( t(22) = 0.33, p = 0.75, \) indicating that loss of part shape and loss of interpart relations yield approximately equivalent decreases in the BOLD signal (Fig 11).
Figure 11: % BOLD signal change for the Intact minus Scrambled contrast (green, top), the Scattered minus Scrambled contrast (red, middle), and the Intact minus Scattered (blue, bottom) contrast for the two ROIs, LO and pFs.

In LO, % BOLD signal change was significantly lower than 0 for the Familiar minus Novel contrast, \( t(7) = 3.49, p = 0.01, \) Cohen’s \( d = 2.63 \), but not in pFs, \( t(7) = 0.16, p = 0.88 \), suggesting the LO responds more strongly to Novel images than to Familiar images (Fig 12).

![Graph](image)

Figure 12: % BOLD signal change for the Familiar minus Novel contrast for the two ROIs, LO and pFs. LO responded more strongly to Novel images than to Familiar images.

4.3 Peri-stimulus Time Course Evaluation

The relationship observed in the block-design ROI analysis (Intact > Scattered > Scrambled) was replicated in the data from the event-related design runs (Fig 13). The peak % BOLD signal change of the peri-stimulus time course was highest for the Intact minus Fixation contrast (0.40% increase), followed closely by the Scattered minus Fixation contrast (0.35% increase). The BOLD response to the Scrambled images was the lowest of the three (0.14% increase).
4.4 Behavioral Results

4.4.1 Error Rates and Reaction Times

A one-sample t-test relative to a chance level of 5.9% showed that subjects reliably identified changes in border color above chance (mean proportion of border changes identified = 99.3%, SD = .02, t(12) = 46.8, p < 10^-25). Responses were made well within the allotted window of 2000 ms (M = 549.6 ms, SD = 113.7 ms), indicating that subjects remained alert and attentive during the scan.

4.4.2 Familiarity Judgments

Because the object images in some cases lacked the textural characteristics and small irregularities of the familiar objects, we assessed whether subjects could accurately judge whether an object image was familiar or novel. Fifteen subjects who did not participate in the main experiment viewed the stimuli and judged whether they were familiar or not. The response (a key press) terminated the stimulus presentation. The mean accuracy in judging the images was 91.5% ± 2.47%, significantly greater than chance. A one-sample t-test relative to the chance level of 50%, t(14) = 23.7, p < 10^-19, Cohen’s d = 1.40, indicates that the subjects could reliably distinguish familiar objects from novel objects. These judgments were made in a mean response time of 1416 ± 242 ms, less than the 1670 ms allowed in the fMRI task. 13 of the 15 subjects had mean response times less than 1670 ms.
5. DISCUSSION

For the past 20 years, the LOC has been localized in fMRI research by contrasting activation to intact objects and scrambled objects. While there have been many different methods of scrambling, all of them share two characteristics: scrambling disrupts the shape of the parts which comprise the object, but it also disrupts the spatial relationships between the constituent parts. By introducing modified stimuli, we were able to separately evaluate the relative costs of the disruption to part shape and interpart part relations in LOC activation.

Our results show that the loss of part shape and the loss of interpart relations contribute approximately equally to the greater activity in LOC attributable to intact as compared to scrambled objects. We also replicated Margalit et al. and Malach’s finding that there is no net effect of object familiarity on BOLD responses within LOC. We did, however, find a preference for Novel images in LO which is consistent with our whole-brain voxelwise finding that the Novel minus Familiar contrast yielded clusters of activation in or near early visual cortex, also apparent in Margalit et al., In Press. The lateral occipital activation for the Novel minus Familiar contrast may be a result of the novel objects inviting greater inspection, whereas the parietal activation apparent in the whole-brain results may reflect greater implicit (i.e., imagined) motor interactions with the novel 3D objects.

That the Familiar minus Novel contrast did not yield any significant clusters of activation indicates that, although semantic information about familiarity is encoded at some level in the visual processing hierarchy, these effects did not emerge in a whole-brain univariate analysis. We expect that a more sensitive analysis, such as a multi-voxel pattern analysis, would reveal differences in response patterns to familiar and unfamiliar objects in the anterior temporal lobes (associative areas implicated in object memory), if not in LOC itself. Indeed, a 2015 study by Iordan et al. found that a classifier trained on patterns of activation within LOC was able to decode basic, subordinate, and superordinate category membership of images of dogs, flowers, airplanes, and shoes. Further analyses, including those restricted to the LOC regions of interest and multi-voxel analyses, will need to be conducted before stronger statements regarding the representation of familiarity in the visual system can be made. Nevertheless, we believe this study provides strong evidence that LOC represents objects both by part shape and part relations.

Our interpretation of the results with the Scattered condition rests, in part, on the assumption that the scattered images were each perceived as several separated parts, each one a single geon, from a single object. This inference, if confirmed, would likely have been invited from the intact objects, whether novel or familiar, which were similarly composed of several geons. The images of scattered geons could, alternatively, have been viewed as scenes of separate objects. Interestingly, the coding of between-part relations within LOC, as evidenced both in the effects of scattering in the present study as well as Lescroart and Biederman’s (2012) MVPA of medial axis sensitivity in LOC, bears a parallel to the coding of between–object relations in LOC. Kim and Biederman (2011) and Kim, Biederman, & Juan (2011) compared LOC activation to pairs of objects that either were interacting to form a single scene, e.g., a watering can watering a plant, or non-interacting, e.g. the watering can mirror reversed to be oriented away from the plant so that it was interpreted as two separate objects. The interacting objects could be familiar, e.g., a bird perched on a birdhouse, or novel, the bird perched on an ear. LOC was more strongly activated by the interacting objects compared to the non-interacting objects with an additional boost from the novelty of the interaction. None of these effects were observed in PPA.
(Parahippocampal Place Area). The enhanced BOLD signal from the interaction appeared to arise in LO in that TMS (Trans Magnetic Stimulation) applied to LO abolished the advantage of the between-object interactions but TMS applied to the intraparietal sulcus had no effect. Thus, it appears that LOC codes not only object or part shape, but also inter-part or inter-object relations.

6. REFERENCES


