

Available online at www.sciencedirect.com
 ScienceDirect

Vision Research 46 (2006) 4024–4031

**Vision
Research**

www.elsevier.com/locate/visres

Neural evidence for intermediate representations in object recognition [☆]

Kenneth J. Hayworth ^{a,*}, Irving Biederman ^{a,b}^a *Neuroscience Program, University of Southern California, Hedco Neurosciences Bldg., 3641 Watt Way, Los Angeles, CA 90089-2520, USA*^b *Department of Psychology, University of Southern California, USA*

Received 23 February 2006; received in revised form 18 July 2006

Abstract

The lateral occipital complex (LOC), a cortical region critical for human object recognition, has been shown to primarily code the shape, rather than the surface properties, of an object. But what aspects of shape? Using an fMRI-adaptation (fMRI-a) paradigm in which subjects judged whether two contour-deleted images of objects were the same or different exemplars, virtually all the adaptation in LOC [especially in LOC's most anterior portion (pFs)] could be attributed to repetition of the parts, almost none to the repetition of local image features, such as lines or vertices, templates, or basic- or subordinate-level concepts of the object. These results support the hypothesis that the neural representation of shape in LOC is an intermediate one, encoding the parts of an object.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Object recognition; fMRI; Priming; Adaptation; Contour-deleted Images; Lateral occipital complex

1. Introduction

The lateral occipital complex (LOC) (Grill-Spector, Kourtzi, & Kanwisher, 2001; Malach et al., 1995), a cortical region critical for object recognition (James, Culham, Humphrey, Milner, & Goodale, 2003) has been shown to primarily code the shape of an object rather than the object's surface properties, in that it treats a photograph and a line drawing of an object equivalently (Kourtzi & Kanwisher, 2000). There are many ways in which shape can be described, with recent theoretical treatments converging on two alternatives: (a) a hierarchy built up of nested arrangements of local features, such as lines and vertices (Riesenhuber & Poggio, 2002), or (b) a structural description, activated by a hierarchy of local features, as in (a), but specifying simple parts and their relations (Biederman, 1987; Hummel & Biederman, 1992; Humphreys & Riddoch, 1987).

Briefly, parts are the constituent shapes (typically convex or singly concave and either 2D or 3D) of a complex object

that are formed by decomposing the object at points of matched concavities. This procedure, which minimizes concavities in the extracted parts, tends to make them "simple." (See Biederman, 1987, for a more extensive discussion.) It is important to realize that these parts do not, in themselves, carry a specific semantic interpretation. For example, the semantic interpretation of a part, e.g., that it is a *leg* (of a table) or the *handle* (of a frying pan), may be denoted only *after an object is identified*. The cylindrical part comprising a leg of a table may be the same visual primitive (geon) as the cylindrical part comprising the handle of a pan. Complex objects, even novel ones, can thus be described as an arrangement of simple parts.

Here, we use an fMRI adaptation paradigm employing contour-deleted line drawings of objects to determine the sensitivity of LOC to local features and to parts. We report that LOC (especially its most anterior region pFs) is insensitive (little release from adaptation) to a change in the local features. In contrast, we found that LOC shows robust release from adaptation when the part-composition (depicting the same object exemplar) is changed. These results provide a neural basis for previous behavioral priming studies (Biederman & Cooper, 1991) employing a similar logic and stimulus set and provide evidence that LOC

[☆] The authors declare that they have no competing financial interests.

* Corresponding author. Fax: +1 213 740 5687.

E-mail address: khaywort@usc.edu (K.J. Hayworth).

represents an object's parts rather than its local features or subordinate- or basic-level concept.

1.1. Previous related work

Briefly presented, masked pictures of common objects are named more quickly and accurately on their second presentation than their first (Bartram, 1974). A sizable part of this facilitation, or priming, is visual and not just due to activation of a name or basic-level concept of the object because different exemplars with the same name, e.g., a grand piano if preceded by an upright piano, show less facilitation. This visual priming is believed to reflect the remarkably fast and facile processes by which objects are identified (Cooper, Biederman, & Hummel, 1992). By varying, between the first and second presentations, what image features are used in the depiction of an object, one can determine what features are necessary for this facilitation, and in this way probe the types of internal representations underlying object recognition.

Here, we show, with an fMRI event-related adaptation paradigm (which, like behavioral priming studies, has the capacity to probe internal—in this case neural—representations underlying object recognition), that the representation of shape in LOC can be best described in terms of an object's parts (in their specified relations) rather than the local image features or activation of a representation of the subordinate-level concept of the object. The local features are required to activate a representation of the parts but once parts are activated, there is almost no contribution of these features to fMRI adaptation. Since the current fMRI adaptation experiment shares a common logic (and stimulus set) with the behavioral priming study of Biederman and Cooper (1991) we detail that common logic below.

The behavioral priming study of Biederman and Cooper (1991) used two sets of contour-deleted line drawings of objects. In the *Local Feature Deleted* (LFD) set of these

stimuli, every other vertex and line was deleted from each part (so that half the contour was removed), as illustrated in Fig. 1a. One member of a complementary pair was composed of the deleted contour from the other member. Either member of a complementary pair could activate the same parts. Scrutiny is required to determine that members of a complementary pair are not identical. The behavioral experiments of Biederman and Cooper (1991; replicated by Boucart et al., 2002) demonstrated that such images primed their complements, shown approximately 7 min later, composed of the deleted features, as well as they primed themselves. There was a clear advantage of the Identical condition over a Different Exemplar condition. That difference was interpreted as a (lower bound) estimate of visual priming. (“Lower bound” because exemplars of the same basic-level class tend to be more similar in shape than instances from different basic-level classes.) The equivalence of Identical and Complementary conditions indicated that the priming, assessed by the speed and accuracy of naming, was not mediated by the local features (lines and vertices). Because the Identical (and Complementary) conditions were associated with lower error rates and shorter naming times than the Different Exemplar condition, lexical (name) or basic-level concept priming could not account for the facilitation. These results suggested that the priming was mediated by a representation of the simple parts.

However, before that conclusion could be accepted, an alternative explanation had to be ruled out: It was possible that the priming was mediated by activation of a *subordinate-level* concept. That is, even though subjects said “piano” (the basic-level term) in the naming task, when they saw either an upright or a grand, their memory representation could have been that of the subordinate-level concept, rather than an assemblage of the parts. To address this explanation, a *Part Deleted* (PD) set of stimuli was prepared in which each member of a complementary pair had half of the *parts* (Fig. 1b). The subordinate-level priming

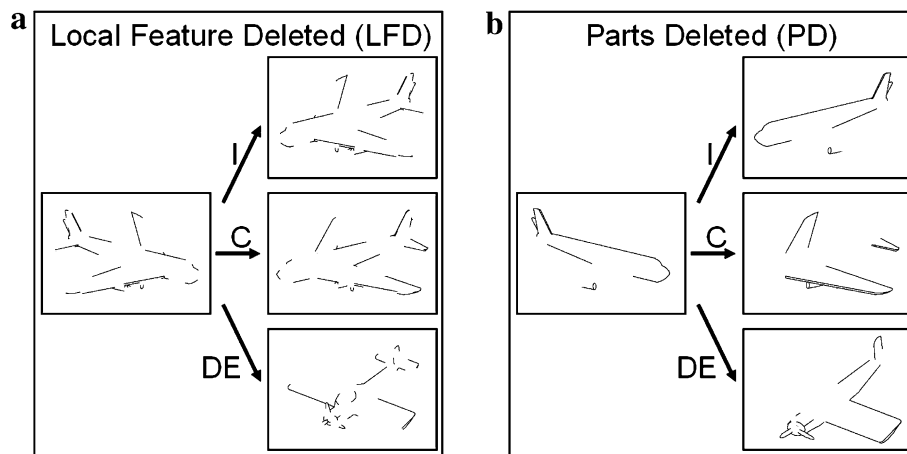


Fig. 1. Examples of local feature deleted (LFD) and part deleted (PD) trial types in the fMRI event-related experiment. (a) Shown is a stimulus S1 (a LFD line drawing of a jetliner) and the stimuli for S2 that would constitute Identical (I), Complement (C), or Different Exemplar (DE) trial types. (b) Shown is a stimulus S1 (a PD line drawing of a jetliner) and the stimuli for S2 that would constitute the three trial types. Note in all cases that S2 is mirror-reversed with respect to S1. Subjects had to judge whether S2 was the same exemplar as S1 so I and C trials required a “same” response.

hypothesis would suggest the same pattern of results as for the LFD experiment. However, with these PD stimuli there was *no* visual priming from the complements, in that second block naming speed and accuracy of the complements were now equivalent to those measures for the different exemplars and both were named more slowly and less accurately than the identical stimuli. The equivalence of the Complementary and Different Exemplar conditions also served to rule out a template account, in which the representation would be that of the original, intact exemplar, which could be partially activated by the PD image. If that were the case, then the complement would have, minimally, produced greater activation of the template than the Different Exemplar. Given the pattern of results from both behavioral experiments, *all* the visual priming could be attributed to activation of a representation of the parts (in their specified relations); none to the local features (lines and vertices), a template of the complete object, the lexical term (the name), or basic- or subordinate-level activation of a concept.

The goal of the present study was to assess the nature of the neural representation of objects in LOC through the use of an fMRI-adaptation experiment utilizing these LFD and PD stimuli. Just as the behavioral priming between LFD complements and lack of priming between PD complements (relative to their respective different exemplars) implied an intermediate level of representation in visual cognition, a similar pattern of results showing fMRI-adaptation between LFD complements and lack of adaptation between PD complements in LOC would imply that the neural coding of objects in LOC was of an intermediate level of representation.

2. Methods

2.1. Stimuli

The stimuli (Fig. 1) were the local feature deleted (LFD) and part deleted (PD) line drawings of objects from Biederman and Cooper's (1991) priming experiments. All stimuli used in this study are available for download from <ftp://geon.usc.edu/contdelstim>. Stimuli were presented to the subject via a video projector directed onto a screen in the bore of the magnet. Objects subtended approximately 10° of visual angle.

2.2. Fast event related fMRI-adaptation protocol

Using an event-related, fMRI-adaptation design, we assessed whether area LOC reveals the same pattern of results as the priming experiments. In a given run, subjects ($n = 8$) viewed either LFD or PD stimuli, presented as a sequential pair (S1, S2) on each trial. The images on a trial were always from the same basic-level category. Each trial lasted 2 s; S1 was presented for 300 ms followed by a 400 ms blank, and then S2 for 300 ms (Fig. 2). These display presentation parameters are identical to those used by Kourtzi and Kanwisher (2000) in their fMRI-adaptation study. The subject then had the remaining one second in the trial to respond same/different exemplar using a button box. No feedback was provided. Responses were obtained from five of the subjects; the other three were run without a response requirement. (The relative ordering of BOLD adaptation for the various conditions was the same with or without the motor response.) The relation between S1 and S2 defined Identical (I), Complementary (C), and Different Exemplar (DE) conditions. Subjects (who had the response requirement) judged whether S1 and S2 depicted

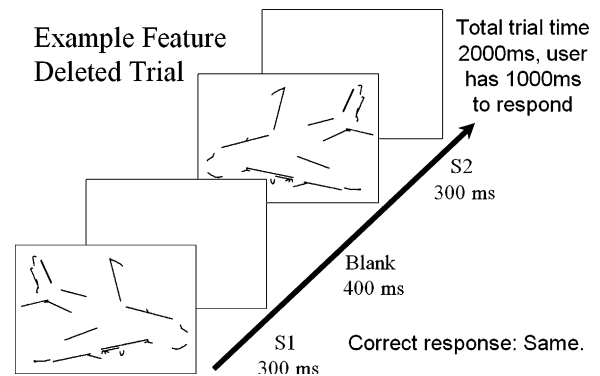


Fig. 2. Example of an I trial for the fast event related fMRI experiment.

the same or different exemplar. For example, if S1 and S2 were both grand pianos (as on I and C trials), whether they differed or not in the specific contours that were deleted, they were to respond "same"; if S1 was grand and S2 upright (as on DE trials), then "different." If the pair of images to be judged on a trial were just shown at the same position and orientation, the task would be too easy, particularly on the I trials. Instead, the stimuli were always reflected around the vertical axis. In the original Biederman and Cooper (1991) priming study with these stimuli, half the second block images were similarly reflected with no effect on priming.

This fast event related experiment was designed such that the BOLD response in the Identical (I) and Different Exemplar (DE) conditions together would provide a scale against which the magnitude of adaptation of the complement (C) condition could be measured. This scale was assessed independently for the LFD and PD experiments. Thus, it is never the case that an LFD BOLD response is compared directly to a PD BOLD response. Within each experiment (LFD and PD) stimuli were balanced across all of the conditions (I, C, and DE) such that each image appeared an equal number of times in all three conditions. This is common practice in fMRI-adaptation experiments to avoid any BOLD difference between the conditions not attributable to adaptation. Thus, the only difference between the (I, C, and DE) conditions was the ordering of the presentation of the images on individual trials.

Runs consisted of an equal number of the three conditions, Identical (I), Complementary (C), and Different Exemplar (DE), along with blank trials in a random jittered design with trial history balanced over all conditions (including a blank condition) for the preceding 3 trials. Total number of trials per run was 259, and each subject participated in 4–8 of these runs, half of which were LFD while the other half were PD. In addition to these main experimental runs, each subject also participated in at least two LOC localizer runs.

2.3. ROI defined by a specially designed LOC localizer

The level of fMRI adaptation was assessed for the LFD and PD tasks by measuring the event-related BOLD response in a region of interest (ROI) defining each subject's LOC. Rather than using a standard LOC localizer (e.g., Malach et al., 1995), which compares the BOLD activation for images of objects to that for a scrambling of the objects (resembling

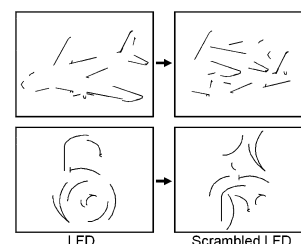


Fig. 3. Examples of LFD and scrambled LFD line drawings used as a LOC localizer.

texture), we defined our LOC by a specially designed localizer that compares activation for LFD line drawings vs. their scrambled versions (Fig. 3). The scrambling for this new localizer was done by translating the local features of the LFD images within the same approximate envelope, thus preserving the local features and their orientations in the scrambled versions. Each 16-s block of the localizer consisted of 32 images, each shown for 500 ms. In each run, blocks of fixation, LFD, and scrambled LFD (sLFD) images were presented.

Whereas the regions of significant differential activation (Fig. 4) for the standard localizer and this LFD minus sLFD localizer agreed in general (and thus ROI's defined by both are equivalent), the locations of their *maximum* differential activities did not. The standard localizer produced maximum differential response in the lateral occipital area (LO); the LFD minus sLFD localizer produced its maximum response in the posterior fusiform area (pFs).

2.4. Functional imaging details

2.4.1. Subjects

The subjects were six males and two females, all students at the University of Southern California, with a mean age of 25 years. Six of the subjects were right handed. All subjects were screened for safety and gave informed consent in accordance with the USC Institutional Review Board Guidelines.

2.4.2. Data acquisition

Scanning was performed at USC's Dana and David Dornsife Cognitive Neuroscience Imaging Center using its Siemens Trio 3T scanner. A standard single channel head coil was used for all acquisition.

2.4.3. Anatomical imaging

High-resolution T1-weighted structural scans were performed on each subject using an MPRAGE sequence (TR = 2070 ms, 192 sagittal slices, 256 × 256 matrix size, 1 × 1 × 1 mm voxels).

2.4.4. Functional imaging

Full brain functional images were acquired using an echo planar imaging (EPI) pulse sequence (TR = 1000 ms, TE = 30 ms, flip angle = 65°, 64 × 64 matrix size, in plane resolution 3 × 3, 3 mm thick slices, 15 roughly axial slices centered on the ventral aspects of the occipital and temporal lobes).

2.4.5. Data analysis (preprocessing)

All functional imaging runs were preprocessed using the BrainVoyager software package (Brain Innovation BV, Maastricht, The Netherlands). This preprocessing included 3D motion correcting (using Trilinear interpolation) each volume to the first run's first volume (which was always acquired immediately after the anatomical MPRAGE data to ease functional to anatomical coregistration). Slice scan time correction was then

performed using sinc interpolation to make every slice within a volume seem like it was acquired simultaneously. A space-domain 3D spatial smoothing was performed using a 4 mm full-width at half-max Gaussian filter on the volumes. Finally, each volume sequence was filtered in the time domain for linear trend removal, and using a high pass filter set to 3 cycles over the run's length.

2.4.6. Statistical analysis

Statistical analysis of all functional imaging runs was performed primarily using BrainVoyager. First, each subject's motion corrected functional images were coregistered with their same-session high-resolution anatomical scan. Then each subject's anatomical scan was transformed into Talairach coordinates. Finally, using the above transformation parameters, the functional scans were transformed into Talairach coordinates as well. All statistical tests reported were performed on this transformed data.

Analysis of the block design LOC localizer runs was performed for each subject individually using a general linear model to fit the functional data across the Talairach normalized brain. A comparison of (LFD—Scrambled LFD) was performed on this fit data to define each subject's LOC. The threshold for significance (*t*-test) was varied on a per-subject basis to give a robust set of significant voxels indicating the location of this subject's LOC. Bilateral ROI's were then defined for the LO and pFs regions of LOC using BrainVoyager's ROI growth algorithm by picking a seed voxel within each of the anatomical areas and limiting the growth of the ROI to prevent extension beyond the rough anatomical area. In this way the ROI's for the LO and pFs regions covered the entire LOC activation region and overlapped only slightly with each other at the border of LO and pFs. (LO ROI's: mean size = 7083 mm³, standard deviation = 3223 mm³; pFs ROI's: mean size = 6686 mm³, standard deviation = 4878 mm³).

For the fast event related runs, a deconvolution analysis was performed on data averaged over all voxels within each subject's localizer-defined ROI's. The fast event related data's BOLD response over this ROI was deconvolved using a 20-point fitting function using the BrainVoyager software. The β values for this deconvolution were used to calculate the %BOLD change (as a function of time).

In order to quantify the significance of the differences between the deconvolved hemodynamic responses for the three conditions, the peak (average of the β values for time points 5 and 6) of this response was computed for each condition for each subject. A paired *t*-test (one tail) was then performed over all ($n = 8$) subjects to determine if the Complementary condition was significantly different from the Identical or Different Exemplar condition. The significance values reported in the text are from this comparison.

3. Results

3.1. Subject behavioral responses

Subjects achieved high levels of accuracy, although the task was more difficult for the PD task than the LFD task (96% correct on the LFD; 83% on the PD). Fig. 5 shows a breakdown of responses for each condition.

A d' analysis was performed on the responses of each subject during the Part Deleted runs to determine the discriminability of the Complementary vs. Different Exemplar conditions (ignoring the responses for identical trials which were close to perfect). Even for this most difficult discrimination, all subjects showed d' values greater than 1, with mean $d' = 1.62$. In order to assess potential bias due to the greater difficulty of the PD task, BOLD analysis was performed both with and without the incorrect responses (as discussed below). Results were identical.

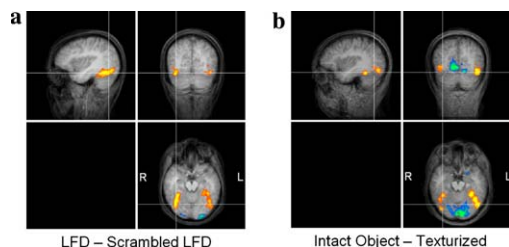


Fig. 4. (a) Average of Talairach normalized brains ($n = 6$) showing the activation pattern produced by the comparison (LFD minus Scrambled LFD). (b) Results of a standard LOC localizer for comparison. Average of Talairach normalized brains ($n = 6$) [separate subjects from (a)] showing the activation pattern produced by the comparison (Intact Object minus Texturized). The localized sites of LOC defined by the two localizers are in good agreement. (Right is left in the fMRI images, threshold for t -map $p < .05$, Bonferroni corrected.)

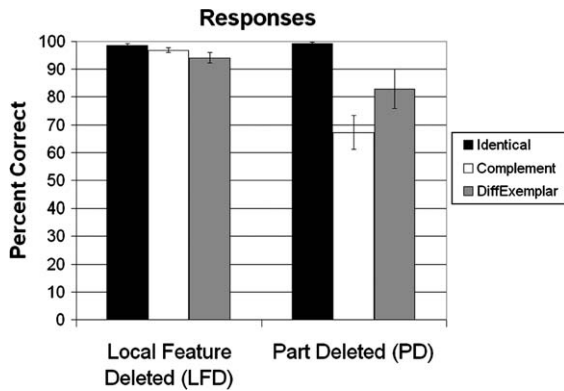


Fig. 5. Accuracy in performing the LFD and PD tasks for each of the three trial types. (Error bars show the standard error of the mean.) The mean correct response times (RTs) showed the same pattern as the error rates: For the LFD stimuli, the RTs for the I, C, and DE conditions were 629, 660, and 745 ms, respectively. For the PD condition, the corresponding RTs were 631, 775, and 758 ms.

3.2. fMRI adaptation results

The hemodynamic response function (in the pFs region of LOC), averaged across subjects ($n = 8$), for the LFD task is shown in Fig. 6a. The reduced activation of the Identical stimuli compared to the Different Exemplar condition, $t(7) = 6.008$, $p = .0003$, is evidence of BOLD adaptation that is visual and not attributable to lexical or basic-level concept activation. The response in the Complementary condition is almost equivalent to that in the Identical condition, again significantly reduced in comparison to the Different Exemplar condition, $t(7) = 3.813$,

$p = .003$, indicating that virtually none of the adaptation is released by completely changing the local features in the image.

Fig. 6b shows the results for the PD trials. Again, the Identical images showed significantly greater adaptation than the Different Exemplar trials, $t(7) = 3.078$, $p = .009$, providing evidence for visual rather than basic-level priming. In contrast to the Complementary condition with the LFD stimuli, here the Complementary condition's BOLD response is equivalent to the Different Exemplar condition, being significantly greater than the Identical condition, $t(7) = 3.203$, $p = .008$. These results indicate that when there is a change in the part composition of a specific exemplar, virtually all adaptation in pFs is lost, indicating that the representation of an object in pFs is neither a template of the complete object (that could be activated by the parts from the complement to a greater extent than the different exemplar) nor a subordinate-level concept. These fMRI results provide a strong parallel to the original behavioral priming study.

Fig. 7a and b show the hemodynamic response, averaged across subjects, for both the LFD and PD experiments over the LO area of LOC ($n = 7$, one subject's LO could not be unambiguously defined and was thus excluded). The overall pattern of results is similar to that seen in pFs but with a smaller difference between the I and DE conditions suggesting less sensitivity to the specific shape of the object. This difference between the LO and pFs subregions of LOC is consistent with the findings of other researchers showing a possible increase in feature complexity in pFs (Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001).

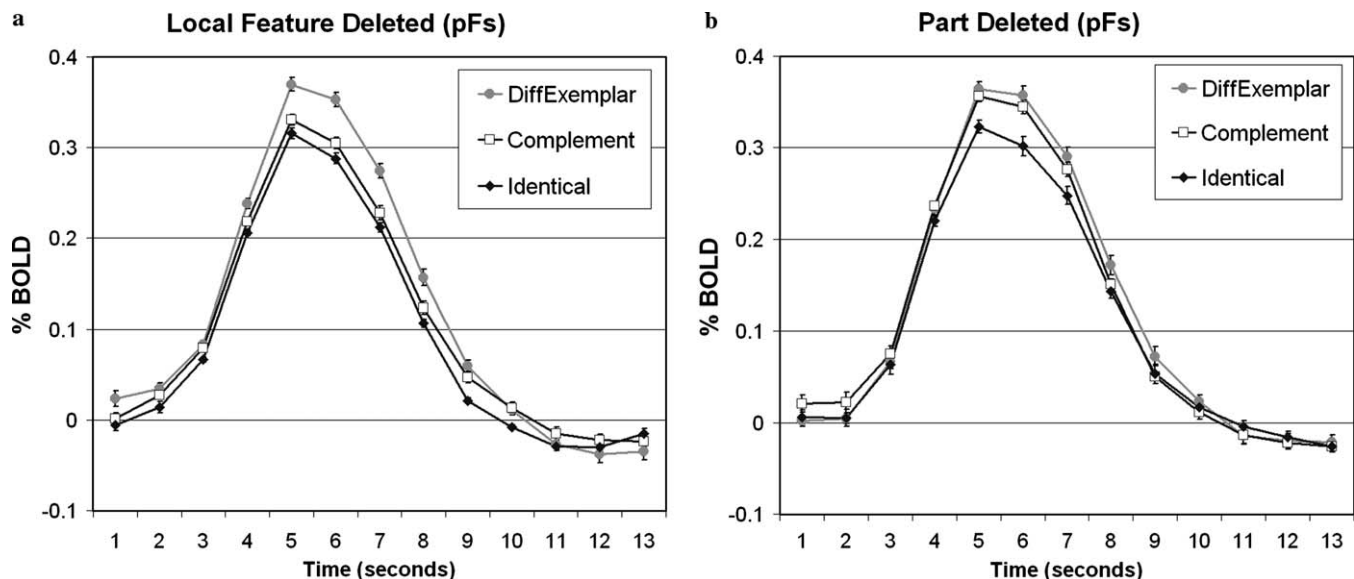


Fig. 6. Event-related BOLD response (% change over fixation baseline) averaged over all subjects ($n = 8$) in the pFs region of LOC. (a) LFD hemodynamic response as a function of time following the second stimulus presentation. The Identical and Complementary conditions were almost equivalent, with both showing significant adaptation with respect to the Different Exemplar condition. (b) PD hemodynamic response. Complementary and Different Exemplar conditions were almost equivalent with both showing a significant release from adaptation relative to the Identical condition. Statistical significance of these differences is given in the text. (Error bars in this and subsequent figures reflect the standard error of the BOLD difference scores, between I, C, and DE conditions, for each time point for each subject.)

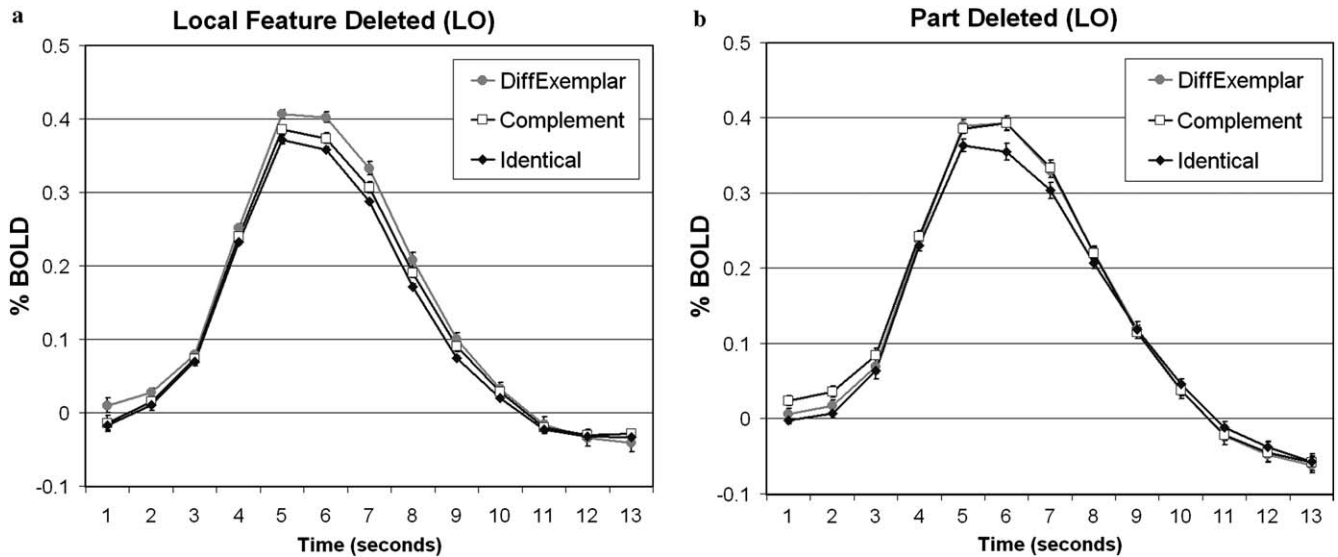


Fig. 7. Event-related BOLD response (% change over fixation baseline) averaged across subjects ($n = 7$) in the LO region of LOC. These results are similar to those from the pFs region of LOC shown in Fig. 6. (a) LFD hemodynamic response. In the LFD task, Identical, and Complementary conditions were not significantly different, $t(6) = 1.500$, $p = .09$, with both showing significant adaptation with respect to the Different Exemplar condition, $t(6) = 4.348$, $p = .002$, and $t(6) = 2.207$, $p = .03$. (b) PD hemodynamic response. In the PD task Complementary and Different Exemplar conditions were almost equivalent with both showing a release from adaptation relative to the Identical condition, $t(6) = 1.880$, $p = .05$, and $t(6) = 2.015$, $p = .05$.

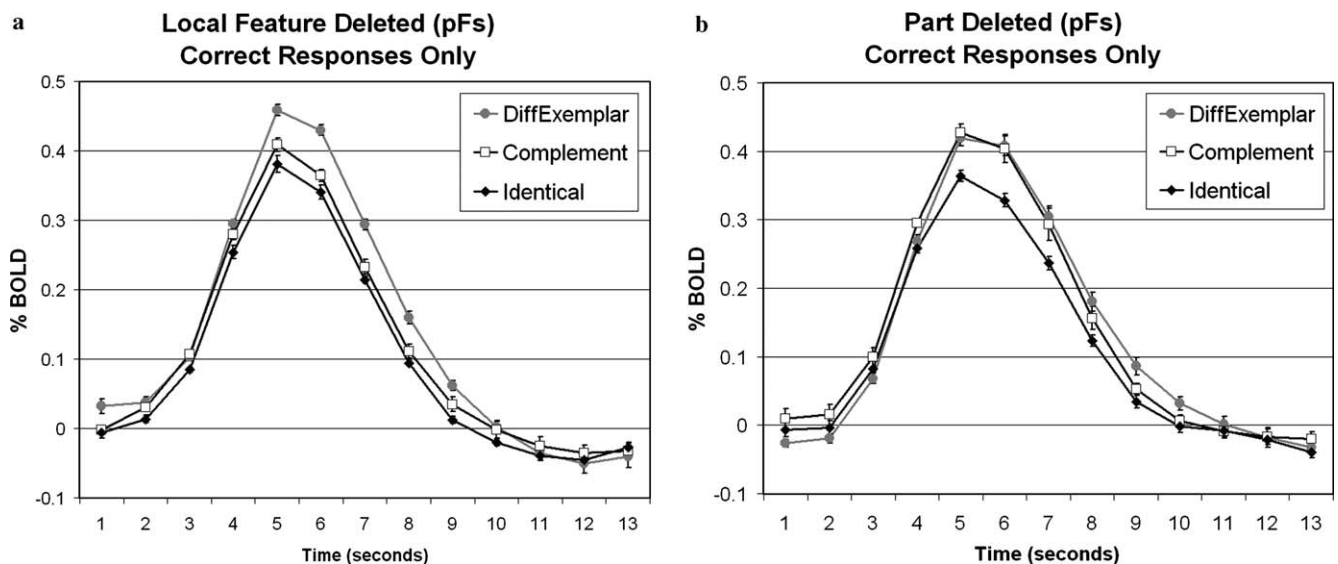


Fig. 8. Event-related BOLD response (% change over fixation baseline) in the pFs region of LOC for *correct responses only* averaged over all subjects for which responses were obtained ($n = 5$). Results are virtually identical to those presented in Fig. 6. (a) LFD hemodynamic response as a function of time following the second stimulus presentation. The Identical and Complementary conditions were almost equivalent, with both showing significant adaptation with respect to the Different Exemplar condition, $t(4) = 6.028$, $p = .002$, and $t(4) = 4.439$, $p = .006$, respectively. (b) PD hemodynamic response. Complementary and Different Exemplar conditions were almost equivalent with both showing a significant release from adaptation relative to the Identical condition, $t(4) = 3.251$, $p = .02$, and $t(4) = 4.932$, $p = .004$ respectively. (Note: A similar analysis of correct trials for region LO did not reach significance due to LO's lower differential responses (as noted above) and insufficient number of subjects for which both response data and LO ROI were available ($n = 4$)).

Fig. 8a and b again show the hemodynamic response in pFs for five subjects, excluding all error trials. This ensures identical motor responses for corresponding conditions, and assesses a possible source of bias due the increased difficulty of the PD task. Results were identical to that for all the trials (statistical significances provided in figure caption.)

4. Discussion

Taken together, these results strongly imply that LOC (especially its most anterior aspect pFs) is representing the parts of an object, rather than local features, templates, or object concepts. These results do not rule out complex arrangements of local features as proposed by some feature

hierarchy models of vision (e.g., Riesenhuber & Poggio, 2002); however, they do place strong limits on the types of feature arrangements LOC could be using. For example, such feature arrangements are unlikely to be as simple as individual vertices since the LFD complements were designed to share few of these yet they gave nearly equivalent adaptation to the LFD identical condition. This suggests that any feature arrangement used by LOC may be equivalent to simple parts (geons).

4.1. Alternative explanations

These results, suggesting a level of representation in LOC more molar than simple vertices, are in line with many previous studies demonstrating LOC's relative insensitivity to low-level stimulus properties. For example, Kourtzi and Kanwisher's (2000) reported that objects repeated in either the same format or different formats (line drawing vs. gray scale) showed equivalent BOLD signal adaptation. Images in different formats present dramatically different low-level image properties (e.g. power spectra) to the visual system yet LOC seems insensitive to these differences. In light of this previous research, our entire set of stimuli was very well matched in all low-level properties (all being contour deleted line drawings), thus any residual low-level differences between our conditions are unlikely to explain the adaptation results.

One could still argue that LOC encodes objects in terms of a list of simple vertices and that LFD complements adapt each other more than PD complements because, to the extent that objects contain some repeated and/or symmetric structure, the process of LFD complementation generates some overlap in the types of vertices shared by both complements. For example, an object composed of a circle next to a square would produce LFD complements both having right angle vertices (in different spatial relations), but the PD complements would not share right angle vertices. In this experiment, we explicitly avoided symmetry in our object set, yet some overlap at what is typically referred to as the "feature-list" level (Hummel & Biederman, 1992) may have remained. Several lines of evidence argue against this interpretation. First, the original priming experiments showed no effect of mirror reversal on priming efficiency, and the absolute BOLD levels of the LFD and PD Identical (I) conditions exactly match, $t(7)=0.7201$, $p=.25$. This argues against any differential effects between the LFD and PD experiments being caused by mirror reversing S2. Second, the degree of BOLD adaptation seen in the LFD complement condition is not simply intermediate between the identical and different exemplar conditions, it is virtually equivalent to the identical condition (especially in pFs). LFD complements may overlap more in their feature-list descriptions than PD complements, but show nowhere near a complete overlap. Thirdly, LOC's lack of sensitivity to local features is in line with previous experiments showing that variation in whether an object is shown in front or in back of partially occluding bars does not yield a release

from adaptation (Kourtzi & Kanwisher, 2001). Last, the LFD vs. sLFD localizer experiment used to generate ROIs in this experiment made use of scrambled stimuli (sLFD) that had *exactly* the same feature-list description as the LFD stimuli they were compared against. LOC gave a robust drop in BOLD signal to the sLFD stimuli, demonstrating again that LOC is relatively insensitive to local features such as simple vertices.

4.2. Object representation in LOC

In summary, many previously published experiments have demonstrated that LOC is insensitive to low-level image properties, and some experiments have implied that LOC may be insensitive to local features like simple vertices. The failure of the LFD complement condition to show significant release from adaptation relative to the identical LFD condition can be viewed as yet additional evidence of this lack of sensitivity to "image format."

The more surprising result is the PD complement condition's complete release from adaptation (equivalent to the different exemplar condition). The current state of knowledge about representation in LOC could be reasonably summarized as follows: LOC seems to represent an object's identity relatively independent of the format of presentation. What we have shown is that this is, in fact, *not* the case: if one presents the same exemplar object using different parts then the result is a full release from adaptation in LOC. We have found an image manipulation preserving object identity (assessed behaviorally) that still shows complete release from adaptation. That manipulation is complementation of parts, exactly as implied by the behavioral priming experiments and our theory.

The finding of preferential coding of parts in LOC is consistent with a recent fMRI study (Hayworth & Biederman, 2005) in which subjects viewed brief movies in which a part of a two-part object could either change shape, change its relation to the other part, or the whole object could move. The three conditions were equated with respect to pixel changes. Changes in part shape produced greater activation in LOC compared to the other conditions, which were equivalent.

Several investigations (e.g., Grill-Spector et al., 2001) have shown a release from adaptation in LOC when a different basic-level class of an object (Dog → Grand Piano) is shown compared to a presentation of the identical shape of the object. What about sensitivity to subordinate level variations? Grill-Spector et al. (1999) showed that 32 s runs of 32 identical image of cars (with a 125 ms blank after each image) yielded a reduced BOLD response compared to runs in which varied images of cars were shown. To our knowledge, the present study is the first, using an event-related design (with high trial-to-trial variability in basic level classes), to demonstrate release from adaptation in LOC not just to a change in basic-level class, but also to a change in subordinate-level class (Upright Piano → Grand Piano). However, our results go further and indicate that it

is not the concept of the object, at any level, that is adapted, but a representation of the object's parts. More generally, the strong parallel of the fMRI-a results in the present study with the behavioral priming version of the study (Biederman & Cooper, 1991), lends support to the interpretation that fMRI-adaptation is an indicant of visual priming, and that the site of this visual priming may coincide anatomically with LOC.

Acknowledgments

Supported by NSF 0420794, 0426415, and 0531177, and the Dornsife Endowment. We thank Xiaomin Yue, Bosco Tjan, Zhong-Lin Lu, Michael C. Mangini, Anne Treisman, and Jiancheng Zhuang for their assistance.

References

- Bartram, D. J. (1974). The role of visual and semantic codes in object naming. *Cognitive Psychology*, *6*, 325–356.
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*, 115–147.
- Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: evidence for intermediate representations in visual object recognition. *Cognitive Psychology*, *23*, 393–419.
- Boucart, M., Biederman, I., Cuervo, C., Danion, J.-M., & Wagemans, J. (2002). A study of the effect of benzodiazepines on structural and conceptual priming. *Psychopharmacology*, *165*, 43–50.
- Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: a review and further evidence. *Canadian Journal of Psychology*, *46*, 191–214.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422.
- 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.
- Hayworth, K. J., & Biederman, I. (2005). Differential fMRI activity produced by variation in parts and relations during object perception. *Journal of Vision*, *5*, 740a.
- Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480–517.
- Humphreys, G. W., & Riddoch, M. J. (1987). *Visual object processing: a cognitive neuropsychological approach*. London: Lawrence Erlbaum Associates Ltd.
- James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain*, *126*, 2463–2475.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*, 3310–3318.
- Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science*, *299*, 1506–1509.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex*, *11*, 287–297.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jlang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Science United States of America*, *92*, 8135–8139.
- Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current Opinion in Neurobiology*, *12*, 162–168.