Vision Research 49 (2009) 2297–2305

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics?

Jiye G. Kim [a,*], Irving Biederman [a,b], Mark D. Lescroart [b], Kenneth J. Hayworth [b]

[a] University of Southern California, Department of Psychology, 3641 Watt Way, Los Angeles, CA 90089, USA
[b] University of Southern California, Neuroscience Program, 3641 Watt Way, Los Angeles, CA 90089, USA

ARTICLE INFO

Article history:
Received 9 December 2008
Received in revised form 19 June 2009

Keywords:
Object recognition
fMRI
Adaptation
Lateral occipital complex
Gabor-jet model

ABSTRACT

A change in the basic-level class when viewing a sequence of two objects produces a large release from adaptation in LOC compared to when the images are identical. Is this due to a change in semantics or shape? In an fMRI-adaptation experiment, subjects viewed a sequence of two objects and judged whether the stimuli were identical in shape. Different-shaped stimuli could be from the same or different basic-level classes, where the physical similarities of the pairs in the two conditions were equated by a model of simple cell similarity. BOLD responses in LOC for the two conditions were equivalent, and higher than that of the identical condition, indicating that LOC is sensitive to shape rather than to basic-level semantics.

1. Introduction

The lateral occipital complex (LOC), composed of the lateral occipital cortex (LO) and the posterior fusiform gyrus (pFs), has been implicated as critical for object perception (James, Culham, Humphrey, Milner, & Goodale, 2003). It is operationally defined as the area showing greater activation from the perception of intact objects compared to their scrambled counterparts (Malach et al., 1995). When viewing a sequence of two objects, a change in object category from a cup to a phone, for example, produces a large release from adaptation in LOC compared to when the same images are repeated (Grill-Spector, Kourtzi, & Kanwisher, 2001). Is this sensitivity to a basic-level category change a function of the change in shape, per se, or the semantic changes associated with the change of shape?

That the release from adaptation could be a consequence of a change in shape rather than a change in surface features was suggested by Grill-Spector et al.'s (2001) finding that the adaptation was maintained (so as to be equivalent to identical images) when a photograph of a cup was presented following a line drawing of that same cup. That at least some of the adaptation is tuned to shape – specifically, to the object's parts – and not semantics was shown by Hayworth and Biederman (2006). Using an event-related design, they found that viewing a sequence of a pair of complementary images of line drawings of objects, each composed of half of the object's parts, showed a complete release from adaptation in LOC, in that the BOLD response was equivalent to a same name, different shaped exemplar (e.g., grand piano → upright piano). In contrast, when the complements were produced by deleting every other vertex and line from each part so that each member of a complementary pair had the same parts, there was no release from adaptation relative to the identical image, indicating that the sensitivity was not due to the change in local features, but produced by the change in parts.

Given that at least some of the coding in LOC, as measured by adaptation, is specific to shape, is there still a component that is specific to (basic-level) semantics? To address this question the present study used a fast event-related fMRI-adaptation (fMRI-a) design (Grill-Spector & Malach, 2001) to compare the release from adaptation when an image sequence was composed of different exemplars from the same basic-level category (e.g., a Labrador and an Irish Setter) with images from different basic-level categories (e.g., a Labrador and a monkey). Semantic sensitivity would be expected to produce a smaller BOLD response with objects from the same category.

Because of the demonstrated sensitivity of LOC adaptation to shape, a critical methodological step was to equate the physical similarities of image pairs that were within vs. across basic-level categories. The scaling was performed with the Gabor-jet model developed by von der Malsburg and his colleagues (Lades et al., 1993). The Gabor-jet system simulates the local multi-scale, multi-orientation filtering of the visual field that is characteristic of V1. Unlike other models that have been used to quantify the visual similarity of images (Grill-Spector et al., 1999; Humphreys, Riddoch, & Quinlan, 1988; Kurbat, 1997; Laws & Gale, 2002), the Gabor-jet model is sensitive to an image's spatial layout and internal details. Moreover, there is strong behavioral validation of the

* Corresponding author.
E-mail address: jiyekim@usc.edu (J.G. Kim).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.06.020

Gabor-jet model in match-to-sample experiments with metrically varying faces and highly irregular blobs. The physical similarity of the distractor to the matching stimulus (which was identical to the sample), as assessed by Gabor-jet calculations, was strongly correlated (mid .90s,) with error rates and reaction times, indicating excellent predictability of that measure of similarity and psychophysical discriminability (Yue, Subramaniam, & Biederman, 2007).

The evidence for semantic influences in LOC is somewhat equivocal. Koutstaal et al. (2001) and Simons, Koustaal, Prince, Wagner, and Schacter (2003) observed a reduction in BOLD activation to images belonging to the same compared to different basic-level classes in left pFs. In fMRI event-related experiments, subjects were shown a series of images during a study phase while they performed a size-judgment task (bigger than a 13 in. square box?). They continued performing the task in the subsequent test phase with images that they had either seen before (repeated same), different exemplars of the study objects (repeated different), or new objects of different semantic categories as the study objects (novel). In the left pFs, repeated different images elicited smaller BOLD responses compared to novel images, indicating sensitivity to the same semantic category during the study phase. A shortcoming of these studies, however, was that the physical similarity of the stimuli was not controlled. Arbitrarily selected object images with the same name tend to be more physically similar than images with different names (Biederman & Cooper, 1991b), perhaps producing the smaller BOLD responses in the repeated different condition.

Similarly, a recent study using support vector classification showed that voxel activity patterns in LOC can better discriminate objects between categories, e.g., a chair and a teapot, than exemplars within a category, e.g., two different types of teapots (Eger, Ashburner, Haynes, Dolan, & Rees, 2008). However, as their stimuli were not scaled for physical similarity, their results may simply reflect a greater sensitivity of LOC for distinguishing between a chair and a teapot because they differ more in shape than two different types of teapots.

Other studies (Grill-Spector et al., 1999; Vuilleumier, Henson, Driver, & Dolan, 2002) have reported little or no effect of semantics when comparing neural responses to different exemplars of the same basic-level class (e.g., an open umbrella and a different folded umbrella) vs. images belonging to different classes (e.g., an umbrella and a hammer). They found no significant difference in the overall adaptation in LOC to different images belonging in the same vs. different basic-level classes, although Grill-Spector et al. found a small region (less than 5%) of LOC that adapted to different objects of the same semantic category by using a voxel-by-voxel analysis. As the physical similarity of the stimuli, again, were not controlled, it is possible that the voxels may have adapted to similar physical features shared by the different exemplars. Chouinard, Morrissey, Kohler, and Goodale (2008) utilized the mean pixel-wise Euclidian-distance in an attempt to control for physical similarities between image pairs of the same and different semantic categories. Consistent with the previous studies, they reported that LOC is not sensitive to semantic information, as it did not show greater release from adaptation to objects from different semantic classes compared to different exemplars from the same semantic class. Some limitations to the shape-scaling model used by Chouinard et al. are its insensitivity to orientation and the absence of any behavioral validation of the model. The Gabor-jet system, employed in the current study, overcomes this shortcoming by utilizing multiple scales and orientations. Additionally, there is strong behavioral verification for its scaling of perceptual similarity as described above (with additional evidence described in a pilot study and in Section 2).

The goal of the current experiment was to assess whether LOC exhibits sensitivity to basic-level semantics. Subjects viewed a sequence of two objects that could either be identical in shape and orientation (Ident), identical in shape but mirror reversed (Mirror), different in shape but of the same basic-level class (Within), and different in shape and basic-level class (Across). Subjects were to judge whether the images were physically identical, ignoring the reversal (Fig. 1). The reversal was included to increase the difficulty of the physically identical trials. This meant that both the within- and across-category pairs had the same response, as those trials would all be physically different. A comparison of the two identical shape conditions also provided a test of LOC's sensitivity to mirror reversal.

2. Methods

2.1. Subjects

Sixteen University of Southern California students (all right-handed, 14 males and two females, mean age of 26.4 years, range: 19–36) participated in the experiment. All subjects were screened for safety, and gave informed consent in accordance with the procedures approved by University of Southern California's Institutional Review Board Guidelines.

2.2. Data acquisition

Scanning was performed at the Dana and David Dornsife Cognitive Neuroscience Imaging Center at the University of Southern

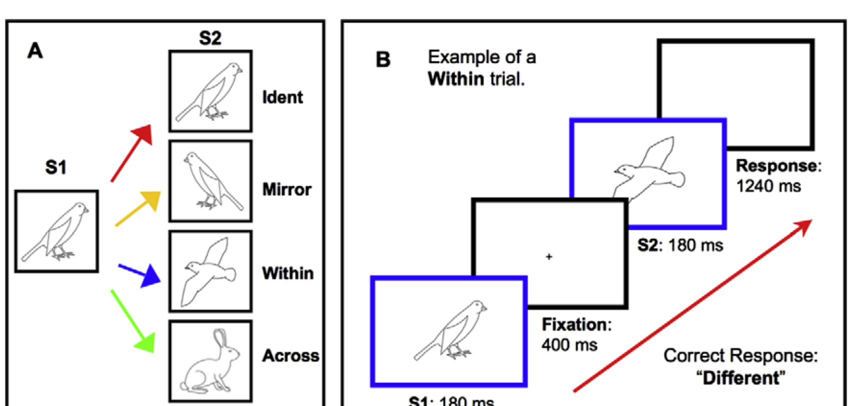

Fig. 1. Illustration of stimuli and presentation sequence. (A) Sample image pairs for the four conditions: Ident, Mirror, Within, and Across. The mean physical similarity values for S1 and S2 for the Within and Across conditions were equal. (B) Presentation sequence for a Within condition trial. For this sample trial, the correct response is "different".

California using a Siemens 3T MAGNETOM Trio with a standard 12-channel coil. Responses were recorded with an MRI compatible button box.

2.3. Stimuli

The stimulus set consisted of 64 line drawings composed of 32 object pairs belonging to the same basic-level category (see Fig. 1A for examples). For the Within condition, exemplars of the same basic-level category that were perceptually very distinct in shape were selected. Pairs of images in the Across condition were of different but closely related basic-level classes and within the same superordinate-level category (e.g., a dog and a monkey, a chair and a cabinet, but not a dog and a cabinet, or a monkey and a chair). This was done to reduce the physical variations that tend to occur when objects belong to distant semantic categories (Dixon, Bub, & Arguin, 1997; Humphreys & Forde, 2001).

A post study questionnaire of eleven subjects, eight of them from the same population as the fMRI participants showed that 91.1% of the exemplar pairs in the Within conditions were given exactly the same basic-level name. Those that were given different names were semantically equivalent terms such as "telephone" and "phone" or "rabbit" and "bunny".

2.4. Gabor-jet scaling of similarity

The physical similarity values were computed for all image pairs using the Gabor-jet model (Fiser, Biederman, & Cooper, 1996; Lades et al., 1993). Each image was sized to fit into a 256×256 pixel array. For the present scaling, a 10×10 square grid, with a horizontal separation of 20 pixels between adjacent horizontal and vertical nodes, was centered on the image. A "jet," consisting of 40 Gabor filters at five scales and eight orientations was centered at each of the 100 nodes of the grid. (Our scaling did not engage the lattice deformation stage by which the jets move to obtain a minimum similarity value.) For each filter, a magnitude of activation vector for that grid point was computed by multiplying the image value by the sine and cosine coefficients of the filter value. This was done with every filter in the jet at every point on the grid for both images yielding a matrix of magnitude values for each filter for each image. A single similarity value was calculated by taking the dot product of the two matrices (one for each image) and dividing by the norm of the two matrices, producing a value between 0 and 1. This computation essentially gives the correlation of the differences of feature information and is equal to 1.00 between two identical images. It is important to note that the similarity values primarily reflect metric variations, such as the 2D positions of the object's contours and the orientation and scale of the whole object as well as its subregions (i.e., the object's parts), without assigning greater weight to whether differences are of nonaccidental properties (NAPs). Examples of NAP differences are whether a contour is straight or curved, a pair of contours is parallel or not, the particular type of vertex produced by the cotermination of contours, or whether a part is present or absent. Subjects are more sensitive to NAPs than metric properties (e.g., Biederman & Bar, 1999). For this reason, the predictability of the similarity measure for our stimuli, which did differ in NAPs, would not be expected to be as high as with the stimuli of Yue et al. (2007), which only varied metrically.

Our design required that the mean similarity values for the image pairs in the Within and Across conditions be equivalent. For the Within condition the value was 0.275 (SD: 0.11) and in the Across condition, 0.281 (SD: 0.10), a difference that was far from significance, $t(190) = 0.37$, $p = .71$. The mean similarity value for image pairs in the Mirror condition was 0.439 (SD: 0.21), which was significantly greater than that of the Within, $t(126) = 5.46$, $p < .001$, and Across, $t(190) = 7.06$, $p < .001$, conditions. The range of the pairs in the Mirror condition was wider, possibly due to a few stimuli that were nearly symmetrical. The Gabor-jet similarity values correlated moderately (and positively) with subjective judgments of physical similarity, $r = .35$, $p < .001$, ($df = 190$). These judgments were made by 10 subjects on a seven-point scale.

2.5. Behavioral procedure

Subjects performed the shape verification task during fMRI scanning. As in the behavioral pilot study (described below), the subjects performed the task by pressing a key if the shapes of S1 and S2 were physically identical (as they were in conditions Ident and Mirror) and a different key if the shapes were different (conditions Within and Across). The total duration of a single trial was 2000 ms: S1 was presented for 180 ms followed by a fixation for 400 ms, and S2 for 180 ms (Fig. 1B). Subjects had the rest of the trial to respond by button box. The stimuli were presented via a video projector and mirrors onto a screen in the bore of the magnet. Each object subtended a visual angle of approximately $3^\circ \times 3^\circ$. S1 was always presented at the center of the screen. For the last eight of the 16 subjects, S2 was translated 2° diagonally to one of four random positions with respect to S1. This was done in an attempt to reduce the near ceiling effect (95% correct) observed for the first half of the subjects, although it had only a minimal effect.

Thirty practice trials were given prior to the scanning, which were composed of different images than those of the experimental trials. Subjects were instructed to maintain fixation throughout the task.

2.6. fMRI-a design

There were a total of four runs, each consisting of an equal number of five trial types in a pseudo-randomly jittered design: Ident, Mirror, Within, Across, and Fixation. The Fixation condition consisted of a blank screen with a black dot presented centrally. The trial history of two look-backs was balanced across all trial types. This resulted in 252 trials per run, where the first two trials were excluded from the analyses and a total of 200 trials per condition. Each run lasted approximately 8.5 min. BOLD responses for the four main conditions were examined in a functionally defined region of interest (ROI), the LOC composed of LO and pFs.

2.7. ROI procedure

Two block-designed localizer runs (approximately 4.8 min each) were also presented to define LOC. Each run was composed of twelve 16-s blocks, with the blocks alternating between those composed of intact objects and those composed of the scrambled images of the objects. Scrambling was accomplished by first overlaying a 20×20 square grid on the intact objects producing 400, 25.5×25.5 pixel boxes. Only the boxes that contained at least some portion of the intact object were shuffled in a random fashion, so that the resulting global shape of the scrambled image was comparable to the original object. Each block consisted of 45 images, each shown for approximately 360 ms.

Subjects performed an odd-man-out task during the localizer runs. At the end of each block, they pressed a key if they either saw an intact object during a block of scrambled image or a scrambled image during a block of intact objects, and a different key if they did not see the odd-man-out. The instance of the odd-man occurring (50% of the time) was counterbalanced across the two conditions. This task was included to maintain subjects' attention throughout the localizer runs.

LOC was defined by the greater BOLD activation of intact minus scrambled images for individual subject. The criterion was a thresholded t-map of $p < .05$, Bonferroni corrected (Fig. 2). This resulted in two bilateral regions, a posterior subregion, LO, (mean size = 12,128 mm^3, SD = 6358 mm^3) and a ventral and anterior subregion, pFs (mean size = 7488 mm^3, SD = 3161 mm^3). The mean peak Talairach coordinates were: −41.19 (SD = 4.46), −75.5 (SD = 4.15), −4.37 (SD = 4.99) for left LO; 38.62 (SD = 4.06), −74.06 (SD = 5.31), −5.06 (SD = 4.54) for right LO; −34.56 (SD = 4.03), −49.94 (SD = 7.27), −12.31 (SD = 2.94) for left pFs; 33.06 (SD = 3.92), −48.25 (SD = 8.77), −13.37 (SD = 2.68) for right pFs. BOLD activation for the four conditions was compared individually within LO and pFs.

2.8. Imaging parameters and data analysis

High resolution T1-weighted anatomical images and $T2^*$-weighted echo planar functional images were acquired. The MPRAGE sequence was used for the anatomical scan with TR = 2070 ms, 160 sagittal slices, and $1 \times 1 \times 1$ mm voxels. The functional runs for the main experimental runs consisted of 32 transversal slices covering the whole brain. The imaging parameters were as follows: TR = 2000 ms, TE = 30 ms, FOV = 224 mm, flip angle = 90°, voxel size = $3.5 \times 3.5 \times 3$ mm. The localizer runs consisted of 16 slices covering the occipital and temporal lobes with the following imaging parameters: TR = 1000 ms, TE = 30 ms, FOV = 192 mm, flip angle = 65°, voxel size = $3 \times 3 \times 3$ mm. The difference in imaging parameters for the localizer runs and the experimental runs was to allow for more precise spatial resolution in defining ROIs. This resulted in acquiring 16 slices across the occipito-temporal lobes per each TR. For the experimental runs, a TR of 2000 ms was used to increase the number of slices (32 slices) in order to cover more regions of the brain. This allowed for the whole brain contrast analysis, examining BOLD responses to the conditions outside of the functionally defined ROIs.

The functional imaging data were analyzed using the Brain Voyager QX software (Brain Innovation BV, Maastricht, The Netherlands). The default preprocessing package from Brain Voyager was used which included slice scan time correction with sinc interpolation, 3D motion correction with trilinear interpolation, spatial smoothing with 4 mm full-width at half-max Gaussian filter, and temporal filtering using a high pass filter of three cycles over the run's length for linear trend removal. The preprocessed functional images were coregistered to the anatomical scan. The anatomical scan for each subject was transformed into Talairach coordinates, where all the statistical analyses were performed.

The deconvolution analyses for each subjects' BOLD responses were performed in the pre-identified ROIs using the general linear model. The deconvolution was fitted with ten 2-s time points. The average beta values of the time points were computed into % BOLD signal change. The values for the peak points (5–8 s post stimuli onset) were then compared across conditions for statistical differences in LO and pFs by repeated-measures ANOVAs.

A whole brain contrast analysis was also performed to see if any areas outside of the pre-selected ROIs were differentially activated in the four conditions. Subjects' BOLD responses were concatenated, and the differences of averaged peak activations for the four conditions were calculated.

3. Results

3.1. Pilot study results

When performing a shape verification task, it is more difficult to respond "different" to exemplars that belong in the same superordinate category (e.g., a dog and a cat) than to images that do not (e.g., a dog and a desk) (Kelter et al., 1984; Smith & MaGee, 1980). This interference indicates that although semantics are not required to perform the task, the semantic relationship between the stimuli is, nonetheless, processed automatically.

A pilot behavioral study showed that this categorical similarity effect is present at the basic-level in the current task as responding "different" to two different breeds of dogs was more difficult than responding "different" to a dog and a cat. Subjects ($n = 9$) performed the shape judgment task (identical shape or not?) with image pairs that were either in the Ident, Within, or Across conditions. Both reaction time (RT), $t(8) = 3.34$, $p = .01$ and error rates, $t(8) = 2.92$, $p = .02$ were significantly higher for the Within than Across conditions, indicating greater difficulty in responding "different" to exemplars of the same basic-level class than to images of different, but closely related, basic-level classes. These results indicate that even though the physical identity judgment task did not require semantic classification of the images, nonetheless, the images were processed semantically.

The study also provided some independent confirmation of the relevance of the Gabor-jet scaling to the present task. As the physical similarity values increased for pairs of images, RTs for the Within and Across conditions also increased, as indicated by a significant positive correlation $r = 0.28$, $p < .002$ (df = 190). That is, the more similar the pairs as scaled by the Gabor-jet model, the more difficult it was to respond "different" to them. This is consistent with findings of Biederman and Subramaniam (1997) who showed a positive correlation with RT and error rates and Gabor-jet similarity values for pairs of line drawings of novel shapes (harmonics of a circle) in a same-different judgment task. As discussed above, the presence of nonaccidental differences between the stimuli probably accounts for the reduced magnitude of the correlation compared to what was reported for matching stimuli solely varying in metric properties.

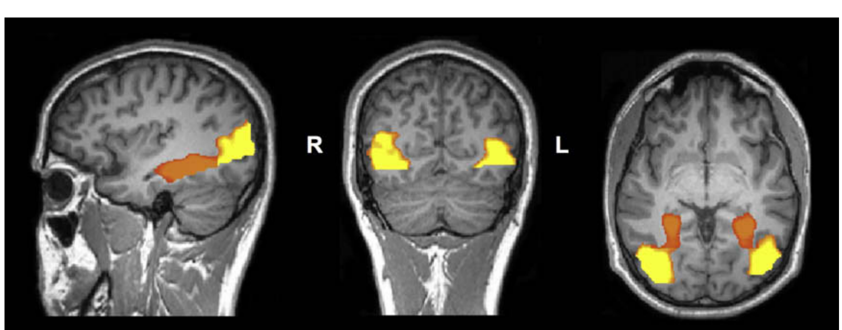

Fig. 2. Average ROIs (LO and pFs) of Talairach (Talairach & Tournoux, 1988) normalized brains ($n = 16$) mapped onto one representative subject's brain. LO (in yellow) and pFs (in orange) are defined by greater activation for intact than scrambled objects.

3.2. Behavioral results

The overall behavioral results were not significantly different for the subjects whose S2 was or was not translated with respect to S1. *ts* for both RTs and error rates were <1.00. Therefore, the results that we report are for the combined data from the two groups.

Behavioral responses for one subject were not recorded due to a malfunction of the button box. The subject was unaware of the situation, and therefore made a "same"/"different" response on each trial.

In general, subjects were highly accurate (94% correct) in performing the shape verification task. Separate repeated-measures ANOVAs were used to compare RTs, with correct trials only, and percent error for the four conditions with two contrasts (between Ident and Mirror where the correct response was "same" and between Within and Across, where the correct response was "different") (Fig. 3).

The overall ANOVA for RTs indicated a significant difference between the conditions, $F(3, 42) = 38.83$, $p < .001$. Subjects were slower at responding, "different" to exemplars within the same basic-level class (648 ms), than images of different classes (611 ms), $F(1, 14) = 38.96$, $p < .001$, an effect consistent with that of the pilot study. They were also slower at responding "same" to mirror reflected (603 ms) than identical images (569 ms), $F(1, 14) = 28.35$, $p < .001$.

ANOVA results for percent error also showed a significant difference across conditions, $F(3, 42) = 4.25$, $p = .01$. Subjects made more errors in the Within (8.4%) than Across (4.4%) conditions, $F(1, 14) = 9.56$, $p < .01$, which is consistent with the pilot study. Though subjects tended to have more errors in the Mirror (6.3%) than Ident (5.2%) conditions, this difference was not statistically significant ($p = .15$).

3.3. fMRI-a results

As with the behavioral data, the overall fMRI results were not significantly different for subjects whose S2 was or was not translated with respect to S1, LO: $t(14) = 1.65$, $p = .12$; pFs: $t(14) = 1.65$, $p = .12$. Therefore, the results reported below are for the combined data from the two groups. Further specification of the results for the two groups is reported in the Supplementary material section.

The average event-related hemodynamic response functions for bilateral pFs are shown in Fig. 4. Although, areas LO and pFs were examined separately, the orderings of the hemodynamic functions of the four conditions for all of the following analyses were identical, and therefore, results for only pFs is reported here. (The overall results for LO are reported in the Supplementary material section.).

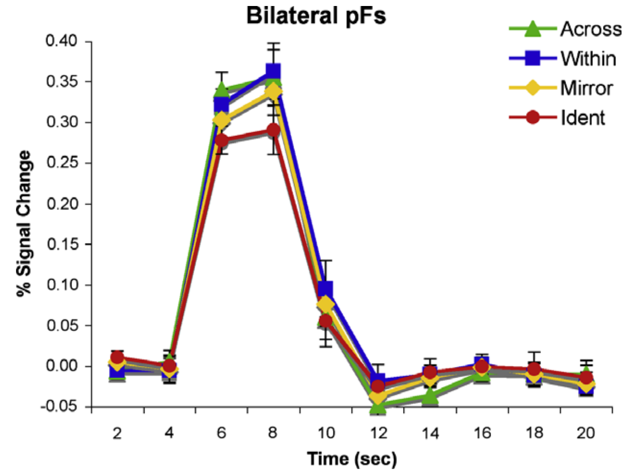

Fig. 4. Hemodynamic BOLD responses (% signal change over baseline) averaged across subjects in pFs. Peak activation responses (5–8 s) for the Within and Across conditions were both greater than the Ident condition, showing a significant release from adaptation. The magnitude of the peak BOLD responses for the Within and Across conditions were not different, indicating adaptation in pFs is not sensitive to basic-level semantic differences. The Mirror condition also showed a significantly greater BOLD response than the Ident condition.

A repeated-measures ANOVA revealed an overall significant difference among the conditions in pFs, $F(3, 45) = 19.83$, $p < .001$.

The peak responses of the Within and Across conditions both showed a significant release from adaptation compared to the Ident condition: Within, $F(1, 15) = 46.67$, $p < .001$; Across, $F(1, 15) = 48.05$, $p < .001$, indicating that the design was sufficiently sensitive to detect a release from adaptation. However, there was no difference in the magnitude of the BOLD responses for the Within and Across conditions, $F(1, 15) < 1$, $p = .44$, indicating that adaptation in pFs is not sensitive to basic-level semantic differences. The Mirror condition also showed a significant release from adaptation to the Ident condition, $F(1, 15) = 13.16$, $p = .002$.

The activation of the left and right pFs were examined in a 2×4 repeated-measures ANOVA, with hemispheres (left vs. right) and conditions as independent variables, to test if the Koutstaal et al. (2001) and Simons et al. (2003) finding of reduced activity in left pFs to within-category images could be replicated. There was no difference between the hemispheres for this comparison, $F(1, 15) = 1.26$, $p = .28$ or interaction of hemispheres and conditions, $F(3, 45) < 1$, $p = .59$. Specifically, there was no interaction between the Within and Across conditions and the left and right pFs, $F(1, 15) < 1$, $p = .98$. (The hemodynamic BOLD responses for left and right pFs are included in Supplementary Fig. 2.)

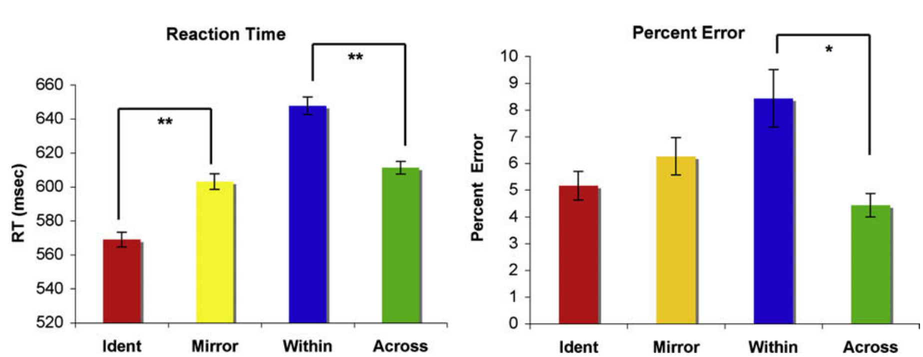

Fig. 3. Mean correct RTs and percent error. Subjects were slower and more prone to error in the Within than Across condition. RT for the Mirror condition was also higher for the Mirror than Ident condition. Asterisks denote significant difference across conditions (* = $p < .01$, ** = $p < .001$). The error bars are the standard errors computed from the deviation scores around each subject's overall mean.

3.4. fMRI-a results with matched reaction times for the Within and Across conditions

It is known that attention can modulate activation patterns in higher visual areas (e.g., Moran & Desimone, 1985; Murray & Wojciulik, 2004). RTs and error rates for the Within condition were significantly higher than those for the Across condition. Although there is evidence that repetition suppression observed in the ventral visual cortex is not affected by levels of task difficulty (Xu, Turk-Browne, & Chun, 2007), it is possible that the Within condition may have required additional attention, therefore masking the semantic sensitivities LOC would have otherwise shown, had the two conditions been equated for behavioral responses.

To assess this possibility, BOLD responses in LO and pFs were re-examined with only correct trials matched in RTs for the Within and Across conditions. For each subject, the Within trials from highest to lowest RTs were paired with Across trials from the lowest to the highest RTs. Each pair was eliminated until the mean RTs were either equal or slightly greater for the Across than the Within condition for each subject. This resulted in discarding an average of 18.3% and 14.5% of the trials for the Within and Across conditions, respectively. The mean RTs for the Within (620 ms) and Across (624 ms) conditions were now reversed, $t(14) = 8.74$, $p < .001$. If task difficulty was somehow masking the semantic effects that would otherwise be apparent in LO and pFs, now that the trials were chosen so that the Across condition was slightly more difficult than the Within condition, the Within condition should show significantly lower BOLD response than the Across condition. But this did not occur. Using only these trials, there was still no difference between the Within and Across conditions either in LO, $t(14) < 1$, $p = .63$, or pFs, $t(14) < 1$, $p = .73$. This selection of a subset of the data did not alter the equivalence in the mean Gabor-jet similarity values for the Within and Across conditions, $t(14) < 1$.

3.5. Voxel by voxel analysis

LO and pFs are relatively large regions comprised of many voxels. If LO and pFs consisted of two populations of voxels, a smaller population tuned to semantic information and a larger population tuned to physical information, then averaging the signals of the two populations would obscure the coding of semantics in this region.

To assess this possibility, a voxel-by-voxel analysis within each subject's ROIs was done to examine the BOLD responses in a smaller spatial scale. This was accomplished by reanalyzing each subject's fMRI data without the spatial smoothing step of the preprocessing procedure. For each voxel within left and right LO and pFs, a comparison of Across minus Within was done with a low threshold of $p < .01$, uncorrected. A liberal threshold was chosen to allow for the detection of any differences between the two conditions.

Only two of the 16 subjects showed voxels that gave greater BOLD responses to the Across than the Within condition. These voxels were small in size, both less than .01% of the total volumes of their ROIs (right pFs and left LO, respectively), and differed in their anatomical location ruling out the possibility that there is a consistent subpopulation of voxels tuned to semantic information in LOC.

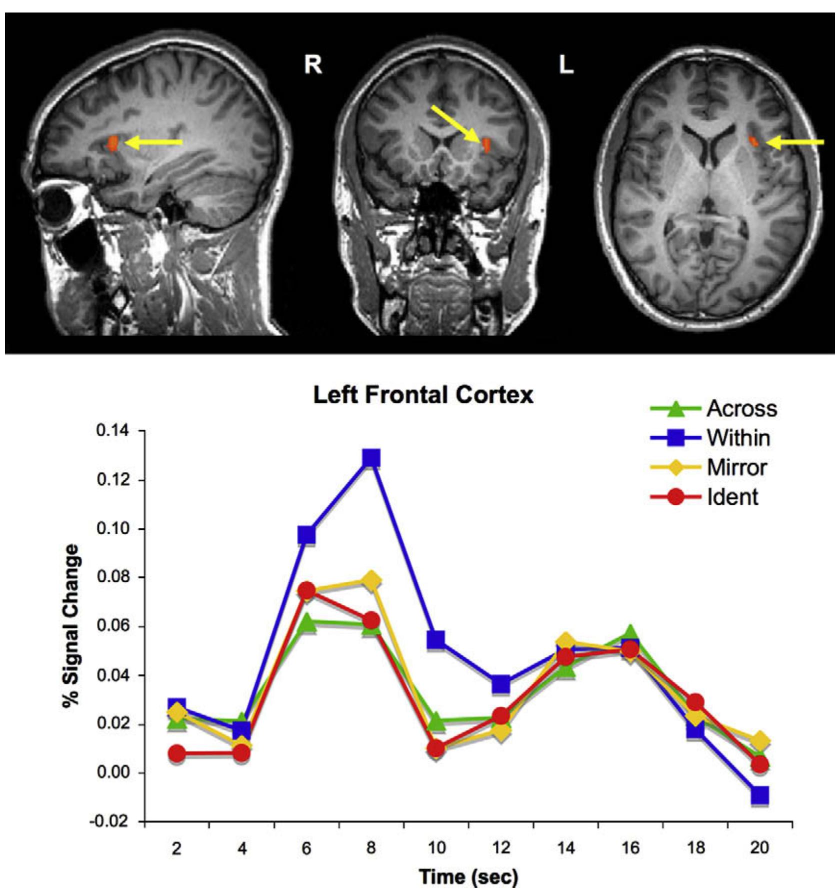

Fig. 5. Left frontal cortex (Talairach coordinates: −30, 16, 7: size = 226 mm^3). This area was defined by a Within minus Across contrast with a p value < .05, Bonferroni corrected. On the bottom are the hemodynamic BOLD responses (% signal change over baseline) for the four conditions. The most difficult condition, Within (as determined by highest RTs and percent error), showed the greatest % signal change in this region.

3.6. Whole brain contrast analysis

A whole brain contrast analysis of Within minus Across was performed to see if areas outside of LOC were differentially activated by the two conditions. Only one area was evident from this contrast, a region in the left frontal cortex (Fig. 5) as defined by a threshold of $p < .05$ Bonferroni corrected (Talairach coordinates: −30, 16, 7: size = 226 mm^3). This area is located near the left prefrontal cortex and middle frontal gyrus, which has been implicated in semantic working memory, selection and retrieval of semantic information (Gabrieli, Poldrack, & Desmond, 1998; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), perceptual decision making (Heekeren, Marrett, Ruff, Bandettini, & Ungerleider, 2006) and task complexity (Volle et al., 2008). It is likely that the greater BOLD response to the Within condition was reflecting its greater difficulty in responding "different" to images of the same basic-level category, rather than directly reflecting semantic similarity. If this area was sensitive to the basic-level semantics of the images, repetition of the semantic categories in the Within condition should have resulted in a lower BOLD response than in the Across condition, a result opposite to what we observed. In fact, the contrast of Across minus Within did not reveal any regions that exceeded threshold.

3.7. Early visual retinotopic areas

BOLD activity in retinotopic areas (V1–V4) was examined to see if a similar pattern of results as those of LO and pFs were evident in earlier areas. If the same pattern was to be found in retinotopic areas, the results in LOC could be interpreted as merely a feedforward effect rather than characteristic of the shape sensitive region.

Standard retinotopic maps were acquired for seven of the eight subjects who were run with S2 translated with respect to S1. Two functional runs with flashing checkerboard wedge stimuli were run to define V1, V2, V3, and V4 boundaries by utilizing a standard voxel-wise correlation method (Engel, Glover, & Wandell, 1997; Sereno et al., 1995). LO and pFs were redefined similarly as defined above, but with the exclusion of those voxels overlapping with retinotopic areas. Details of imaging parameters and visual stimuli used for the retinotopic mapping as well as the anatomical loci of these ROIs are described in the Supplementary materials section.

Since S2 was translated with respect to S1 in all conditions, early visual areas (V1–V4) composed of small receptive field sizes were expected to show no differential pattern of responses to the conditions. Statistical analyses were done with peak hemodynamic responses. Results are shown in Fig. 6. In all early visual areas, there were no differential BOLD responses to the four conditions.

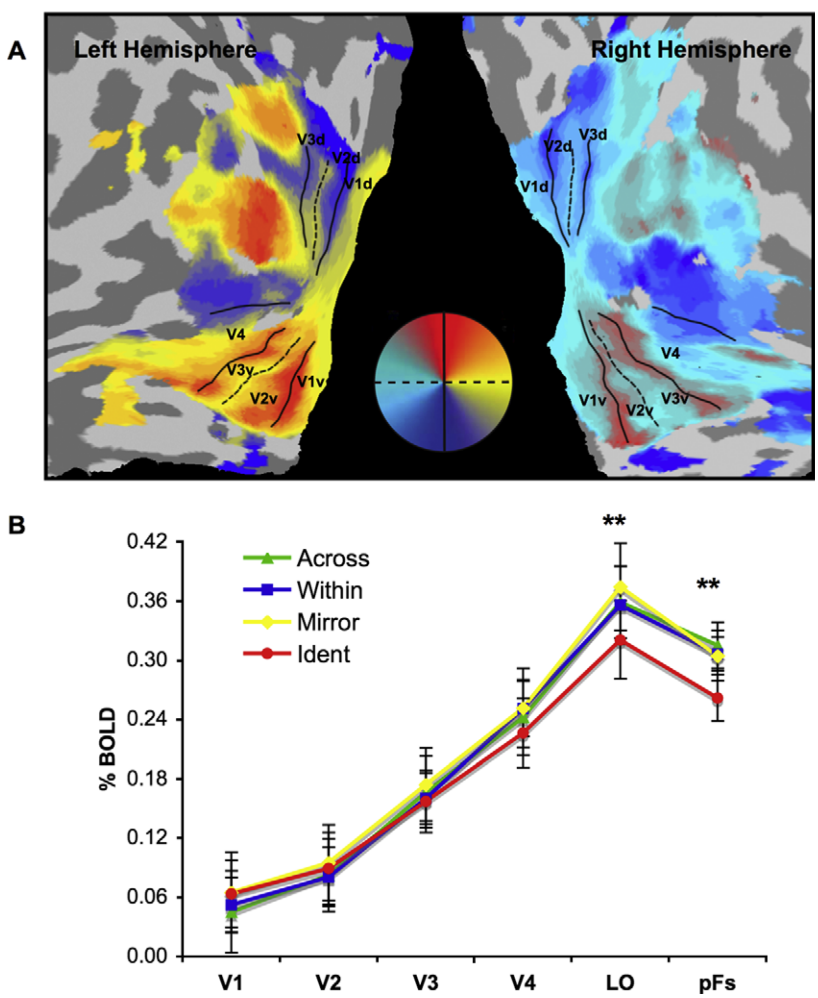

Fig. 6. Peak hemodynamic responses for early visual areas and redefined LO and pFs. (A) A representative subject's retinotopic correlation map ($r > 0.20$) displayed on flattened left and right hemispheres. (B) BOLD responses to the conditions did not differ in retinotopically defined areas V1–V4. The digression of the Ident from other conditions can be observed as visual areas progress in the ventral and anterior direction from V1 to pFs. Refer to the Supplementary materials section for the Talairach coordinates of these ROIs. Asterisks denote significant differences between the Ident and Across conditions ($** = p < .001$).

Repeated measures ANOVAs for V1, V2, V3, and V4 all showed *F*s $(3, 18) < 1.00$. However, significant differences across conditions were maintained in LO, $F(3, 18) = 6.38$, $p = .004$, and pFs, $F(3, 18) = 12.30$, $p < .001$. It is only in LO and pFs that a significant release from adaptation of the Across compared to the Ident condition was evident (comparison of responses between Ident and Across conditions in LO: $F(1, 6) = 22.30$, $p = .003$; and in pFs: $F(1, 6) = 89.38$, $p < .001$). Again, BOLD responses for the Within and Across conditions did not differ in LO and pFs: both *F*s $(1, 6) < 1.00$.

To summarize, in contrast to LOC, early visual areas (V1–V4) showed no overall difference in responses to the four conditions. That these retinotopic areas did not show differences among the conditions suggests that our results in LOC are characteristic of later ventral stream areas, and not merely feedforward from earlier stages.

4. Discussion

The present study indicates that (a) LOC is sensitive to changes in shape, and not basic-level semantics, and (b) left and right LOC are not differentially sensitive to semantic variation.

The current experiment utilized the Gabor-jet system (Lades et al., 1993) to equate the metric similarity of image pairs in the Within and Across conditions. Although this model is not sensitive to changes in nonaccidental properties, we had no basis to presume that image pairs in the Within and Across conditions differed in NAPs. Exemplars that were perceptually very distinct were chosen for the Within condition. Image pairs in the Across condition were of different, but closely related basic-level categories that belonged to the same superordinate class, in an effort to reduce the large physical differences that tend to occur between objects from distant semantic categories (Dixon et al., 1997; Humphreys & Forde, 2001).

As noted previously, the Gabor-jet system, in modeling V1 simple cell tuning, captures the orientation, scale, and 2D position information of shape. Previous studies have either omitted this step in examining semantic sensitivities of LOC (Koutstaal et al., 2001; Simons et al., 2003; Vuilleumier et al., 2002) or have used scaling methods not sensitive to the spatial layout of the objects (Chouinard et al., 2008; Grill-Spector et al., 1999; Humphreys et al., 1988; Kurbat, 1997; Laws & Gale, 2002). The scaling of metrically-varying shapes by Gabor-jet filtering has also been behaviorally validated.

One might argue that the absence of differences in the BOLD responses of the Within and Across conditions was due to a lack of power. However to achieve a significant difference, like the one observed for the comparison between the Mirror and Ident conditions, the difference in BOLD responses of the Within and Across conditions would have to be more than seven times greater than the observed difference. The current results are thus robust and consistent with studies in the past examining semantic vs. shape influences in LOC (Chouinard et al., 2008; Grill-Spector et al., 1999; Vuilleumier et al., 2002).

When performing the shape verification task, subjects were slower and more prone to error when responding "different" to two images that belonged in the same basic-level category than when they did not. Though the shape verification task did not require subjects to attend to the category information of S1 and S2, semantic processing was automatic and influenced performance (as it did in the pilot study). This semantic effect was not manifested in LOC, as the BOLD responses for the Within and Across conditions were essentially equivalent.

The Mirror condition showed a significant release from adaptation compared to the Ident condition, showing LOC's sensitivity to mirror reflection (but see the Supplementary material section for differentiation between responses for Mirror vs. Ident conditions when S2 was or was not translated). However this effect could merely be a consequence of the reduced physical similarity in the Mirror condition compared to the Ident condition, as assessed by the Gabor-jet system. Previous studies have shown invariance in long-term behavioral priming to mirror reversal (Biederman & Cooper, 1991a; Biederman & Cooper, 1991b; Fiser & Biederman, 2001). However in the present fast adaptation paradigm, both behavioral RTs and BOLD responses in LOC were increased by mirror reflection.

5. Conclusion

The critical methodological scaling by the Gabor-jet model allowed us to equate the physical similarity of the shapes of Within and Across object pairs. The results strongly indicate that LOC is sensitive only to changes in shape, not to changes in basic-level semantic categories associated with those shape changes.

Acknowledgments

This research was supported by NSF Grants 04-20794, 05-31177, and 06-17699 to IB. We thank Xiaomin Yue for the Gabor-jet scaling and Jianchang Zhuang for assistance in the MRI scanning.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2009.06.020.

References

Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research, 39*, 2885–2899.
Biederman, I., & Cooper, E. E. (1991a). Evidence for complete translational and reflectional invariance in visual object priming. *Perception, 20*, 585–593.
Biederman, I., & Cooper, E. E. (1991b). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. *Cognitive Psychology, 23*, 393–419.
Biederman, I., & Subramaniam, S. (1997). Predicting the shape similarity of objects without distinguishing viewpoint invariant properties (VIPs) or parts. *Investigative Ophthalmology & Visual Science, 38*, 998.
Chouinard, P. A., Morrissey, B. F., Kohler, S., & Goodale, M. A. (2008). Repetition suppression in occipital–temporal visual areas is modulated by physical rather than semantic features of objects. *NeuroImage, 41*, 130–144.
Dixon, M., Bub, D. N., & Arguin, M. (1997). The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cognitive Neuropsychology, 14*, 1085–1130.
Eger, E., Ashburner, J., Haynes, J.-D., Dolan, R. J., & Rees, G. (2008). FMRI activity in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, 356–370.
Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7*, 181–192.
Fiser, J., & Biederman, I. (2001). Invariance of long-term visual priming to scale, reflection, translation, and hemisphere. *Vision Research, 41*, 221–234.
Fiser, J., Biederman, I., & Cooper, E. E. (1996). To what extent can matching algorithms based on direct outputs of spatial filters account for human shape recognition? *Spatial Vision, 10*, 237–271.
Gabrieli, J. D. E., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences, 95*, 906–913.
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., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*, 187–203.
Grill-Spector, K., & Malach, R. (2001). FMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica, 107*, 293–321.
Hayworth, K. J., & Biederman, I. (2006). Neural evidence for intermediate representations in object recognition. *Vision Research, 46*, 4024–4031.

Heekeren, H. R., Marrett, S., Ruff, D. A., Bandettini, P. A., & Ungerleider, L. G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proceedings of the National Academy of Sciences, 103*, 10023–10028.

Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavior and Brain Sciences, 24*, 453–509.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5*, 67–103.

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.

Kelter, S., Grötzbach, H., Freiheit, R., Höhle, B., Wutzig, S., & Diesch, E. (1984). Object identification: The mental representation of physical and conceptual attributes. *Memory Cognition, 12*, 123–133.

Koutstaal, W., Wagner, A. D., Rotte, M., Maril, A., Buckner, R. L., & Schacter, D. L. (2001). Perceptual specificity in visual object priming: Functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *Neuropsychologia, 39*, 184–199.

Kurbat, M. A. (1997). Can the recognition of living things really be selectively impaired? *Neuropsychologia, 35*, 813–827.

Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., Von der Malsburg, C., Wurtz, R. P., et al. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers, 42*, 300–311.

Law, K. R., & Gale, T. M. (2002). Category-specific naming and the 'visual' characteristics of line drawn stimuli. *Cortex, 38*, 7–21.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, 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 Sciences, 92*, 8135–8139.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*, 782–784.

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience, 7*, 70–74.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*, 889–893.

Simons, J. S., Koustaal, W., Prince, S., Wagner, A. D., & Schacter, D. L. (2003). Neural mechanisms of visual object priming: Evidence for perceptual and semantic distinctions in fusiform cortex. *NeuroImage, 19*, 613–626.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology, 109*, 373–392.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Thieme: Stuttgart.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, 94*, 14792–14797.

Volle, E., Kinkingnehun, S., Pochon, J.-B., Mondon, K., Thiebaut de Schotten, M., Seassau, M., et al. (2008). The functional architecture of the left posterior and lateral prefrontal cortex in humans. *Cerebral Cortex, 18*, 2460–2469.

Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience, 5*, 491–499.

Xu, Y., Turk-Browne, N. B., & Chun, M. M. (2007). Dissociating task performance from fMRI repetition attenuation in ventral visual cortex. *Journal of Neuroscience, 27*, 5981–5985.

Yue, X., Subramaniam, S., Biederman, I. (2007). Predicting the psychophysical discriminability of faces and other complex stimuli based on measures of physical image similarity. Paper presented at the society for neuroscience 37th annual meeting, San Diego, California.