The Contribution of Object Shape and Surface Properties to Object Ensemble Representation in Anterior-medial Ventral Visual Cortex

Jonathan S. Cant[1] and Yaoda Xu[2]

Abstract

■ Our visual system can extract summary statistics from large collections of objects without forming detailed representations of the individual objects in the ensemble. In a region in ventral visual cortex encompassing the collateral sulcus and the parahippocampal gyrus and overlapping extensively with the scene-selective parahippocampal place area (PPA), we have previously reported fMRI adaptation to object ensembles when ensemble statistics repeated, even when local image features differed across images (e.g., two different images of the same strawberry pile). We additionally showed that this ensemble representation is similar to (but still distinct from) how visual texture patterns are processed in this region and is not explained by appealing to differences in the color of the elements that make up the ensemble. To further explore the nature of ensemble representation in this brain region, here we used PPA as our ROI and investigated in detail how the shape and surface properties (i.e., both texture and color) of the individual objects constituting an ensemble affect the ensemble representation in anterior-medial ventral visual cortex. We photographed object ensembles of stone beads that varied in shape and surface properties. A given ensemble always contained beads of the same shape and surface properties (e.g., an ensemble of star-shaped rose quartz beads). A change to the shape and/or surface properties of all the beads in an ensemble resulted in a significant release from adaptation in PPA compared with conditions in which no ensemble feature changed. In contrast, in the object-sensitive lateral occipital area (LO), we only observed a significant release from adaptation when the shape of the ensemble elements varied, and found no significant results in additional scene-sensitive regions, namely, the retrosplenial complex and occipital place area. Together, these results demonstrate that the shape and surface properties of the individual objects comprising an ensemble both contribute significantly to object ensemble representation in anterior-medial ventral visual cortex and further demonstrate a functional dissociation between object- (LO) and scene-selective (PPA) visual cortical regions and within the broader scene-processing network itself. ■

INTRODUCTION

A visual scene typically contains multiple objects, and often these objects are grouped together into a perceptual unit known as an object ensemble (e.g., leaves on a tree, grapes on a vine). Over the past decade, there has been considerable interest in the cognitive mechanisms that support object ensemble perception. This is likely because not only are object ensembles ubiquitous in everyday vision but also object ensemble representation has adaptive value to everyday behavior. For example, the representation of summary statistics from ensembles of multiple objects complements and guides object-specific processing because it allows the visual system to overcome the capacity limitation inherent in object-based attention (e.g., Alvarez & Cavanagh, 2004; Xu, 2002; Pylyshyn & Storm, 1988; Luck & Vogel, 1997). Indeed, numerous behavioral studies show that observers are able to extract summary information from ensembles of multiple objects, such as their mean size, direction of motion, speed, orientation, and central location. Interestingly, this ability comes at the expense of losing the ability to provide fine details about the individual objects in the ensemble (e.g., Alvarez & Oliva, 2008; Chong & Treisman, 2003; Ariely, 2001; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Watamaniuk & Duchon, 1992; Williams & Sekuler, 1984). This demonstrates that there is a tradeoff across these different visual perceptual processes, and consistent with this notion, we have recently demonstrated that there may be different underlying cognitive mechanisms supporting single-object versus object ensemble perception (Cant, Sun, & Xu, 2015).

The neural mechanisms of single-object perception have long been of interest to researchers and have been studied using various techniques such as human neuropsychological case studies, monkey electrophysiology, and, more recently, fMRI. In general, these studies have demonstrated that a number of different cortical regions are involved in processing various visual features of objects. For example, object shape activates a large lateral region of visual cortex, known as the lateral occipital complex (e.g., Kourtzi & Kanwisher, 2001; Grill-Spector

[1]University of Toronto Scarborough, [2]Harvard University

et al., 1999; Malach et al., 1995), and object texture activates medial and anterior regions of visual cortex, in parahippocampal cortex and overlapping with the scene-selective parahippocampal place area (PPA; e.g., Cant & Goodale, 2007, 2011; Steeves et al., 2004; James, Culham, Humphrey, Milner, & Goodale, 2003).

In contrast to the rich literature on the neural representation of single objects, much less is known about the neural mechanisms supporting object ensemble perception. In our first attempt to examine the neural processing of ensembles (Cant & Xu, 2012), we used the fMRI adaptation approach and found that a region of anterior-medial ventral visual cortex, along the collateral sulcus and overlapping the scene processing region PPA (Epstein & Kanwisher, 1998), was involved in object ensemble representation. We demonstrated that this region was also involved in texture representation, and interestingly, responses for object ensembles and textures were strikingly similar. This is likely explained by the fact that both types of stimuli contain multiple repeating structures that can vary slightly in features such as size, orientation, and color (Portilla & Simoncelli, 2000) and that the extraction of summary statistics is essential in the representations of both. Given that scene representation also involves the extraction of summary statistics (e.g., we do not encode each object in a scene in great detail when we comprehend a scene), anterior-medial ventral visual cortex may play a general role of extracting summary statistics from a host of different visual displays including scenes, textures, and ensembles (Cant & Xu, 2012).

We have since demonstrated that anterior-medial ventral visual cortex is not sensitive to changes in image size (Cant & Xu, 2012), absolute density (i.e., spacing between objects), or the color of object ensembles, but this region does represent information about the ratio or relative density of the different objects constituting a heterogeneous object ensemble (Cant & Xu, 2015).

To further understand the nature of ensemble representation in anterior-medial ventral visual cortex, here we examined how changes to the shape and surface properties (i.e., both color and texture) of objects within an ensemble affect the neural representation in this brain region. As this region overlaps a great deal with PPA (Cant & Xu, 2012), we targeted our investigation to PPA in this study. Previously, Cant and Goodale (2007) reported that anterior-medial ventral visual cortex showed sensitivity to changes in the surface properties of single objects, but not their shape. In our work, we have shown that PPA was not sensitive to local shape changes resulting from the rearrangement of the objects in an ensemble. Although PPA did show sensitivity when all the ensemble objects were changed from one type of object to another (e.g., from a pile of strawberries to a pile of apples), these changes almost always covaried with changes in both object shapes and surface properties, including color, material properties, and/or surface texture (Cant & Xu, 2012, 2015). Thus, whether or not object shape contrib-

utes to ensemble representation in PPA has not been properly evaluated.

In this study, we opted to study ensembles made from real-world materials with naturally occurring surface properties as opposed to using computer-generated stimuli, allowing us to more truthfully reflect the representation of real-world object ensembles in the human brain. In a fast event-related fMRI adaptation paradigm (Cant & Xu, 2012, 2015), we used photographs of semiprecious gem stone beads varying in both shape and surface properties (e.g., an ensemble of heart-shaped pink marble beads; see Figure 1B) to explore whether PPA would show a release from adaptation (i.e., a rise in fMRI activation) when we varied the shape, the surface properties, or both of these ensemble features compared with a baseline condition with no feature changes.

To provide a comparison and contrast to the activation observed in PPA, we also examined responses in the lateral occipital area (LO) and two other scene-selective regions, namely, the retrosplenial complex (RSC) and occipital place area (OPA; see Dilks, Julian, Paunov, & Kanwisher, 2013; this region was previously called transverse occipital sulcus). LO is sensitive to processing the shapes of both single objects (e.g., Kourtzi & Kanwisher, 2001; Grill-Spector et al., 1999; Malach et al., 1995) and object ensembles (Cant & Xu, 2012, 2015), but not surface properties (e.g., Cant & Xu, 2012; Cant & Goodale, 2007; Malach et al., 1995). We thus expected this region to show sensitivity to changes of the shape but not surface properties of the ensemble elements. We have previously reported that ensemble processing is unique to PPA and is not a general property of all scene-sensitive regions in the brain (Cant & Xu, 2012, 2015). We thus expected RSC and OPA to show no sensitivity to changes of either the shape or surface properties of the ensemble elements.

METHODS

Observers

Fifteen paid observers took part in this study (nine women, six men; mean age = 25.93, range = 19–34 years), all of whom were recruited from the Harvard University community, and all were right-handed, reported normal color vision and normal or corrected-to-normal visual acuity, had no history of neurological disorder, and gave their informed consent to participate in the study in accordance with the Declaration of Helsinki. The experiments were approved by the committee on the use of human subjects at Harvard University.

Four additional observers took part in the study but were excluded from further analysis for the following reasons: malfunctioning of the MRI scanner, which resulted in a loss of their fMRI data (1 observer); extremely low average PPA activation (i.e., less than 0.05%) across all conditions in the adaptation task, which made their data unreliable and difficult to interpret (1 observer);

Figure 1. Examples of ROIs
in individual observers
and stimuli used in the
adaptation experiment.
(A) The scene-selective PPA
(Talairach coordinates for
the specific ROI examples
shown, *x, y, z* for right/left:
+22/−23, −40/−40, −5/−5)
was defined by contrasting the
activation for scenes against
the activation for both faces and
objects. The object-selective LO
(+33/−35, −75/−75, +2/+2)
was defined by contrasting the
activation for objects against the
activation for phase-scrambled
objects. The scene-selective
RSC (+17/−20, −56/−53,
+22/+18) and OPA (+35/−33,
−78/−83, +10/+17) were
defined similarly as that of
PPA, namely, by contrasting
the activation for scenes against
the activation for both faces
and objects. (B) Examples of
all of the shape and surface
property combinations used in
the adaptation experiment.
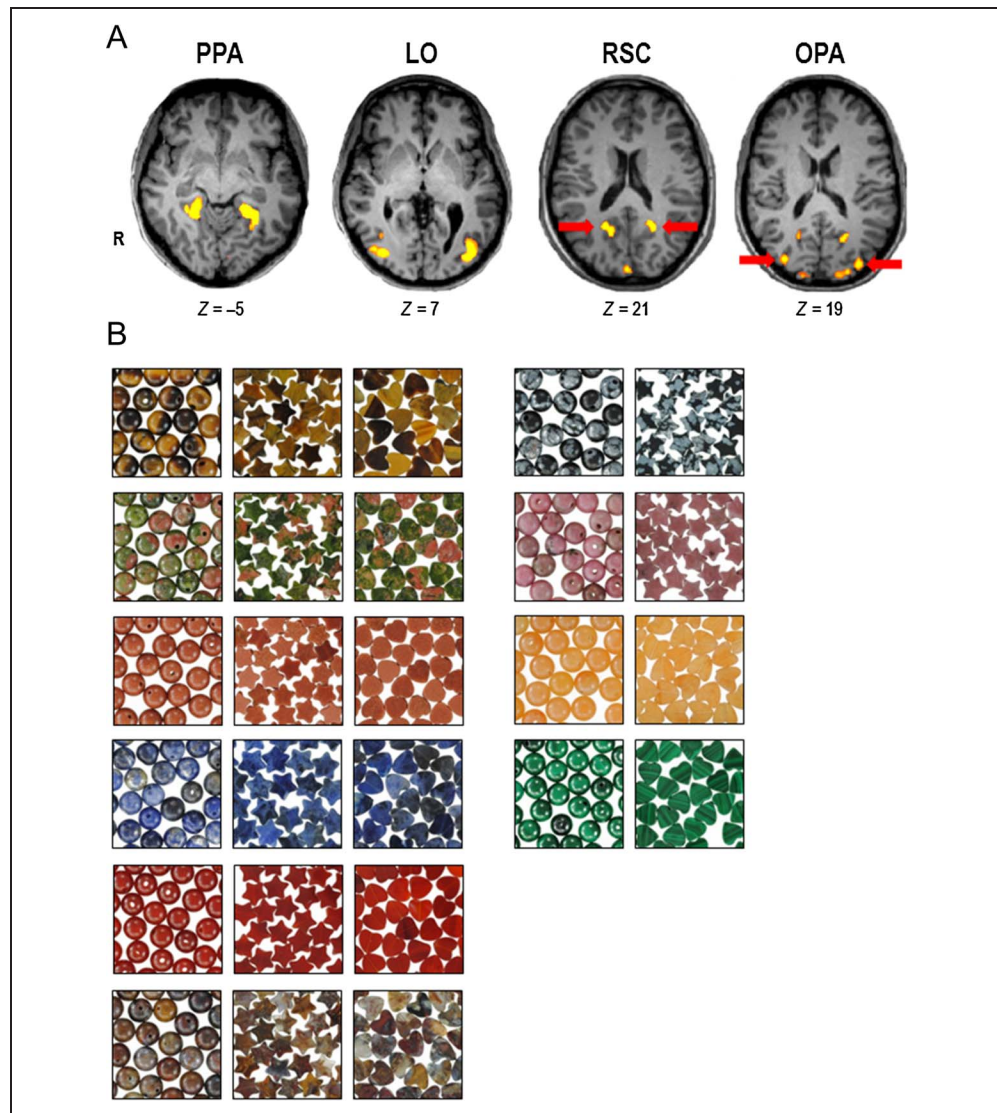
and poor behavioral performance in the adaptation task
(less than 60%, whereas accuracy ranged from 87% to
99% across conditions in the other observers, see Behav-
ioral Data Analysis section for more details), which ren-
dered their fMRI data unreliable and difficult to
interpret (two observers).

Stimuli and Procedures

Adaptation Experiment

A fast event-related fMRI adaptation paradigm, as was
used in Cant and Xu (2012), was used in this experiment.
Each trial contained a sequential presentation of two
images, and observers were asked to categorize the type
of trial encountered, from three possible alternatives:
identical (repeated presentation of the same image),
shared (presentation of different images of the same en-
semble), and different (presentation of different images
that either contained a change in ensemble shape, en-

semble surface properties, or both ensemble shape and
surface properties; see below and Figure 2A for more
details). Our previous study has shown that both a pas-
sive viewing and an active judgment task produce the
same fMRI adaptation results (Cant & Xu, 2012), indicat-
ing that the particular task used does not seem to affect
the adaptation results in PPA, LO, OPA, and RSC (see also
Xu, Turk-Browne, & Chun, 2007). But because an active
judgment task tended to produce more robust results
than a passive viewing task, we chose the same active
judgment task as used in Cant and Xu (2012) to increase
power.

The stimuli used in this experiment were colored pho-
tographs of 26 different object ensembles, with each
containing a repetition of different exemplars of the same
type of object (see Figure 1B). All images subtended
12.5° × 12.5° of visual angle (this also applied to all the
images used in the object/scene localizer). These images
were constructed using stone beads made from semi-
precious gems and photographed using a Nikon D3000

digital SLR camera (Nikon Corporation, Tokyo, Japan) and a desktop photo studio setup. The 26 different ensembles were created by using 10 different stone bead surface properties (e.g., pink marble, red agate, green jade, pink and green unakite, blue and white sodalite) that were available in at least two of three different stone bead shapes (balls, hearts, and stars; six of the surface properties were available in all three shapes, and four of the surface properties were available in two of the three shapes, yielding 26 total unique combinations of ensemble shape and surface properties). We ensured that the background of each image was the same uniform white by editing the images using Photoshop CS3 software (Adobe Systems, Inc., San Jose, CA). Four different photographs of each ensemble were then generated by rearranging the beads in each ensemble, yielding a total stimulus set of 104 unique ensemble images.

There were a total of five stimulus conditions (Figure 2A): (1) identical (the same image was presented two times successively; no change in the surface properties or shape of stone beads across image presentations), (2) shared (presentation of different images of the same ensemble, also with no change in the surface properties or shape of stone beads across images but with the ensemble elements rearranged), (3) different surface properties (the shape but not the surface properties of the stone beads repeated across successive ensemble images), (4) different shape (the surface properties but not the shape of the stone beads repeated across successive ensemble images), and (5) both different (both the surface properties and the shape of the stone beads varied across successive ensemble images). The stimulus images for each trial were randomly selected with the constraint that the same surface properties would not repeat within five consecutive trials, and thus, the same combination of surface properties and shape would not repeat within five consecutive trials. With regard to the behavioral categorization task, observers were told to categorize Condition 1

Figure 2. Example stimuli and results ($n = 15$) for all five fMRI adaptation conditions in PPA and LO. (A) The stimuli used consisted of full-color photographs of object ensembles, each with a unique shape and surface property combination. In each trial, observers saw a sequential presentation of two images that were either identical, shared ensemble features (i.e., contained the same shape and surface properties but in a different arrangement), shared the same shape but differed in surface properties, differed in shape but shared surface properties, or differed in both shape and surface properties. (B) Full average time course of activation from independently and individually localized scene- (PPA) and object-sensitive (LO) ROIs. The first 4 time points represent the length of a single event-related trial. (C) Averaged fMRI responses (representing the time point of greatest signal amplitude in the average response across all conditions) from PPA and LO, shown at the appropriate scales for better visualization of the differences (or lack thereof) across the adaptation conditions. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus & Masson, 1994).
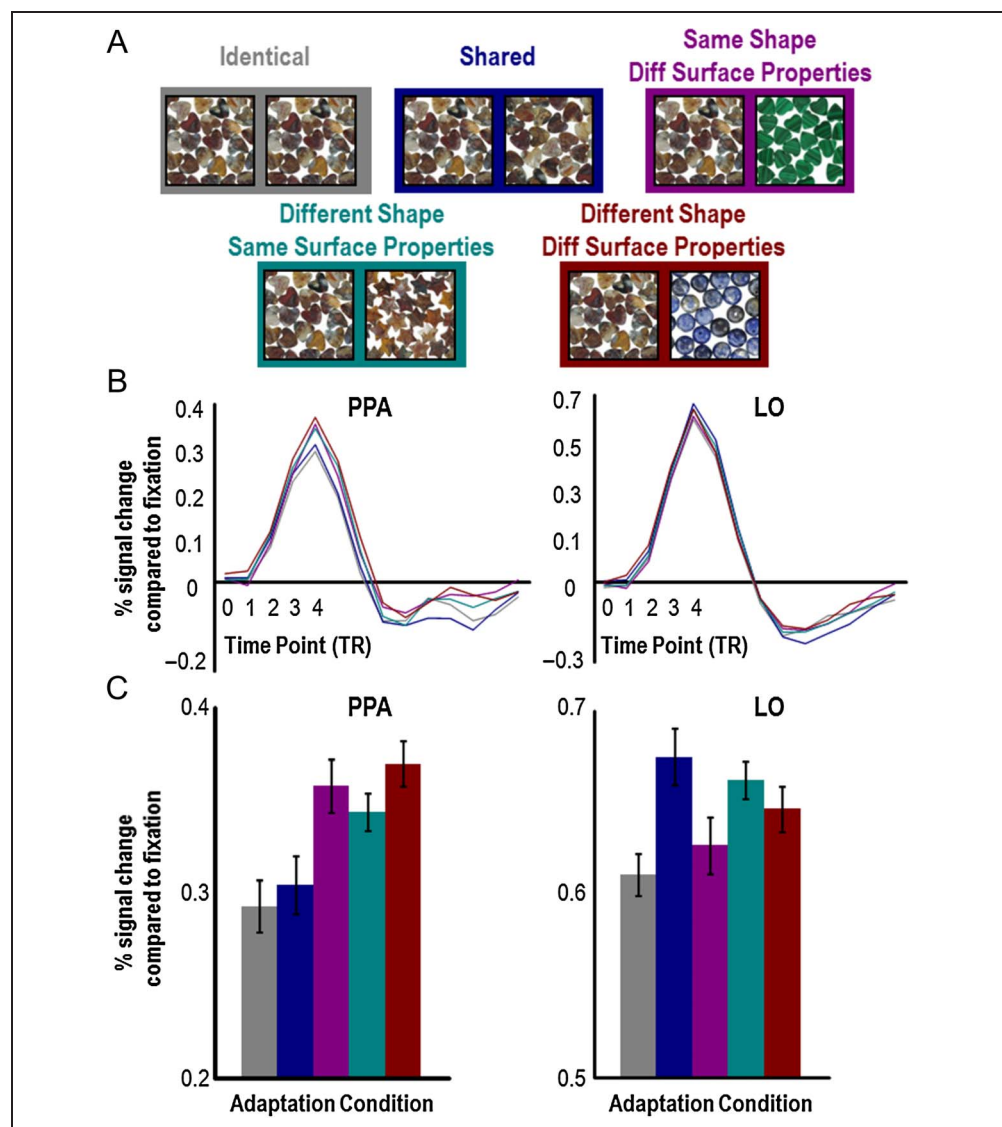
as identical, Condition 2 as shared, and Conditions 3, 4, and 5 as different.

Each trial lasted 6 sec, beginning with a 500-msec fixation, followed by two successively presented images (each presented for 200 msec, with an 800-msec blank fixation in between), and ending with a 4300-msec blank response screen. Observers were asked to categorize a trial as either "identical," "shared," or "different" by pressing the appropriate response button with their right index, middle, or ring finger, respectively. Although we did not counterbalance the finger button assignments, as will be discussed later, this did not lead to the behavioral results obtained. In addition to the stimulus trials, there was also 6-sec blank fixation trials in which no images were presented. Trial order was pseudorandom and balanced for trial history (e.g., trials from all conditions including fixation were preceded and followed equally often by trials from all the conditions, including itself, for one trial back and forward; see Cant & Xu, 2012, 2015; Xu & Chun, 2006; Kourtzi & Kanwisher, 2001). To further balance trial history, trial order was rotated among the conditions in different runs and among different observers. Each adaptation run lasted 7 min and 30 sec and contained 12 trials for each stimulus condition. All observers took part in three adaptation runs, with the exception of three observers who took part in two adaptation runs.

Object/Scene Localizer

The stimuli used to localize object and scene-sensitive areas of cortex consisted of photographs of various indoor and outdoor scenes (e.g., furnished rooms, buildings, city landscapes, and natural landscapes), both male and female faces, common objects (e.g., cars, chairs, food, and tools), and phase-scrambled versions of the common objects.

A single run consisted of presenting four blocks each of scenes, faces, intact objects and phase-scrambled objects. Each stimulus block was 16-sec long and contained 20 different images, each lasting 750 msec and followed by a 50-msec blank period. No images were repeated within or across blocks in a given run. To ensure attention to the displays, observers fixated at the center and detected a slight spatial jitter, occurring randomly in 1 of every 10 images. In addition to the stimulus blocks, there were also 8-sec fixation blocks presented at the beginning, middle, and end of each run. Following Kanwisher, McDermott, and Chun (1997) and Epstein and Kanwisher (1998), we used two unique and balanced run orders. Each run lasted 4 min 40 sec. All observers took part in three runs of this localizer. This localizer had already been acquired in a prior study in seven of the observers. For these observers, instead of repeating the localizer in this study, the localizer data from the prior scanning session were aligned with the adaptation data using our fMRI data analysis software.

Apparatus

Stimulus presentation and the collection of behavioral responses (via a response pad placed in the observer's right hand) were controlled by an Apple MacBook Pro (Apple Corporation, Cupertino, CA) running Matlab with Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). Each image was rear projected via an LCD projector (Sharp Notevision XG-C465X, resolution of 1024 × 768, Sharp Electronics Corporation, Mahwah, NJ) onto a screen mounted behind the observer as he or she lay in the scanner bore. The observer viewed the images through an angled mirror mounted to the head coil directly above the eyes.

Imaging Parameters

This study was conducted on a 3.0-T Siemens MAGNETOM Tim Trio (Erlangen, Germany) whole-body imaging MRI system at the Center for Brain Science, Harvard University (Cambridge, MA). A Siemens radio frequency 32-channel head coil was used to collect BOLD weighted images (Ogawa et al., 1992). For high-resolution anatomical images, T1-weighted 3-D magnetization-prepared rapid acquisition gradient-echo sagittal slices covering the whole brain were collected (inversion time = 1100 msec, echo time [TE] = 1.54 msec, repetition time [TR] = 2200 msec, flip angle = 7°, matrix size = 256 × 256, 144 slices, voxel size = 1.0 mm × 1.0 mm × 1.0 mm). For the functional runs, a T2*-weighted echo-planar gradient-echo pulse sequence (matrix size = 72 × 72, field of view = 21.6 cm) with TR of 1.5 sec was used in the adaptation experiment (TE = 29 msec, flip angle = 90°, 300 volumes). Another pulse sequence with TR of 2.0 sec was used for the localizer runs (TE = 30 msec, flip angle = 85°, 140 volumes). Twenty-four 5-mm-thick (3 mm × 3 mm in-plane, 0 mm skip) slices parallel to the anterior and posterior commissure line were collected in all the functional runs.

Data Analysis

fMRI Data Analysis

fMRI data were analyzed with Brain Voyager QX (Brain Innovation, Maastricht, The Netherlands). Data preprocessing included slice acquisition time correction, 3-D motion correction, linear trend removal, and Talairach space transformation (Talairach & Tournoux, 1988).

Data from the object/scene localizer was analyzed using a general linear model, accounting for hemodynamic lag (Friston et al., 1995). Following Epstein and Kanwisher (1998), the PPA ROI was defined as regions in the collateral sulcus and parahippocampal gyrus whose activations were higher for scenes than for faces and objects (false discovery rate $q < 0.05$; this threshold applies to all functional regions localized in individual observers; see Figure 1A). Following Epstein and Higgins (2007) and Dilks et al. (2013), the RSC and OPA ROIs were defined as regions in restrosplenial cortex–posterior cingulate–medial

parietal cortex and transverse occipital cortex, respectively, whose activations were higher for scenes than for faces and objects. Following Grill-Spector, Kushnir, Hendler, and Malach (2000), LO was defined as a region in lateral occipital cortex near the posterior inferotemporal sulcus whose activations were higher for intact objects than for phase-scrambled objects. All regions were successfully identified in both hemispheres separately for each individual that took part in the study.

Following the standard ROI-based analysis approach (see Saxe, Brett, & Kanwisher, 2006), we overlaid the ROIs from each observer onto their data from the main adaptation experiment and extracted time courses from that observer. The averaged activation levels for all conditions were then extracted and converted to percentage BOLD signal change from baseline by subtracting the corresponding activation from the fixation trials and then dividing by this value (see Cant & Xu, 2012, 2015; Dilks, Julian, Kubilius, Spelke, & Kanwisher, 2011; Todd, Han, Harrison, & Marois, 2011; Xu, 2010; Xu & Chun, 2006; Todd & Marois, 2004; Kourtzi & Kanwisher, 2001). Peak responses for each condition were obtained by collapsing the time courses for all of the conditions (over 14 TRs or 21 sec) and then identifying the time point of greatest signal amplitude in the average response, thereby ensuring that the time point selected was not biased to the level of activation for any one condition in particular (e.g., Cant & Xu, 2012, 2015; Xu, 2010; Xu & Chun, 2006). Moreover, note that this method does not bias the level of activation to a particular time point (i.e., TR) within the trial. The time point of greatest signal amplitude (i.e., the peak) was generally located four TRs into the trial (representing 4.5–6 sec from the beginning of the trial). See Figures 2B and 4B for average time courses of activation for all conditions in each of our ROIs, with the peak response for each condition aligned to the fourth TR (for a small minority of ROIs across participants, the peak was located at either the third or fifth TR; when this was the case, the peak for each condition was shifted forward or backward one TR, respectively). This was done separately for each observer in each ROI, and these resulting peak responses were then averaged across all observers. We first looked for differences in activation for all five adaptation conditions across hemispheres by conducting a 2 (Hemisphere: right vs. left) by 5 (Adaptation Condition: identical vs. shared vs. same shape/different surface properties vs. different shape/same surface properties vs. different shape/different surface properties) repeated-measures ANOVA, performed separately on each ROI. The data from the right and left hemispheres were collapsed if there was no evidence of hemispheric differences across the adaptation conditions. We next analyzed the average levels of activation for each condition (using the identical condition as baseline and excluding the shared condition; see Results for more details) using a two-way repeated-measures ANOVA, performed separately on each ROI (SPSS, Chicago, IL), with alpha = .05. Main effects of interest included

Shape (same vs. different) and Surface Properties (same vs. different). As the motivation for this study is to examine the representation of ensemble shape and surface properties in PPA, we then conducted two planned post hoc pairwise comparisons in this ROI: one t test examining differences in activation for changes in surface properties (i.e., same vs. different) while holding shape constant (i.e., same shape) and one t test examining differences in activation for changes in shape while holding surface properties constant.

We then assessed whether or not there were differences in initial baseline activation across the five adaptation conditions in two ways. Specifically, we looked for differences in activation between all conditions for the first TR in a trial and also when using the average of the first and second TRs. In both cases, we analyzed initial baseline levels of activation across all five conditions using a one-way repeated-measures ANOVA, again performed separately on each ROI, with a main effect of Adaptation Condition (identical vs. shared vs. same shape/different surface properties vs. different shape/same surface properties vs. different shape/different surface properties). Next, we assessed whether any potential differences in baseline activation contributed to differences in peak activation across the adaptation conditions in PPA (the ROI that is most relevant to the motivation for this study) by subtracting two different baseline measures from all time points within a trial (i.e., subtracting the value of first TR from all time points, including itself, and subtracting the average of the first and second TRs in a trial from all time points) and then performing the two-way repeated-measures ANOVA described above.

We also conducted two separate correlation analyses to examine if neural results in each cortical region could be explained solely by behavioral responses. In the first analysis, we examined the relationship between neural adaptation data and behavioral measures of accuracy and response latency across the five different conditions separately. In the second analysis, we examined relationships between neural and behavioral data for two specific adaptation effects: the release from adaptation for changes in surface properties while holding shape constant and the release from adaptation for changes in shape while holding surface properties constant.

Finally, we conducted two different repeated-measures ANOVAs to examine if patterns of adaptation differed across cortical regions: a 2 × 2 × 2 ANOVA with main effects of Region (e.g., PPA vs. LO), Shape (same vs. different), and Surface Properties (same vs. different), and a 2 × 5 ANOVA with main effects of Region (e.g., PPA vs. RSC) and Adaptation Condition (identical vs. shared vs. same shape/different surface properties vs. different shape/same surface properties vs. different shape/different surface properties). Both of these ANOVAs were performed separately for the comparison of each region with PPA (i.e., PPA vs. LO, PPA vs. RSC, PPA vs. OPA).

Behavioral Data Analysis

Behavioral performance measures of RT and accuracy for the adaptation runs were recorded by Matlab (running the Psychtoolbox) and were analyzed with SPSS (Chicago, IL), by performing two-way repeated-measures ANOVAs to assess differences across the conditions in the adaptation runs.

Before data were analyzed, two analyses were performed to remove outliers. First, observers who had an overall accuracy of less than 60% in any adaptation condition were excluded from further analyses. As accuracy ranged from 87–99% across conditions, an accuracy below 60% was a good indication that the observer was not properly engaged in the task, making their data unreliable and difficult to interpret This resulted in the removal of two participants. Second, for the remaining observers, following standard practice, response latencies that were 2.5 standard deviations above or below the mean RT for each stimulus condition were excluded from each observer separately.

Response latencies (for correct trials only) and the number of errors committed were analyzed using a 2 × 2 repeated-measures ANOVA, with alpha = .05. Main effects of interest included Shape (same vs. different) and Surface Properties (same vs. different).

RESULTS

Behavioral Results

In the adaptation runs, observers classified the type of trial that they encountered as either identical, shared, or different (see Methods for more details). All behavioral results are presented in Table 1. The identical and shared trials differed significantly in response latency ($t(14) = 3.00$, $p < .05$) but not in accuracy ($t(14) = 1.72$, $p = .11$). To be consistent with our fMRI analyses, we used data from the identical trials as our baseline condition in both the response latency and accuracy analyses. The analysis on response latencies revealed significant main effects of Shape ($F(1, 14) = 21.25$, $p < .001$) and Surface Properties ($F(1, 14) = 13.89$, $p < .005$), but the Shape-by-Surface Properties interaction was not significant ($F(1, 14) = 0.80$, $p = .39$). The significant main effects of Shape and Surface Properties for response latency reveal that

responses were longer when these features were repeated, which are findings that run counter to the neural adaptation effects that we report for PPA and LO below (see Table 1 and Figure 2). We found similar results for the analysis of accuracy, as the main effects of Shape ($F(1, 14) = 34.34$, $p < .001$) and Surface Properties ($F(1, 14) = 5.07$, $p < .05$) were both significant, but the Shape-by-Surface Properties interaction was not ($F(1, 14) = 1.24$, $p > .29$).

Although we did not counterbalance for the finger assignment when observers made their behavioral responses, we do not think this directly led to the RT results obtained. This is because the fastest (and most accurate) condition (the different trials) was assigned to the ring finger, which is the slowest finger to initiate button presses based on the biomechanics of finger responses. Moreover, the slower (and less accurate) condition (the same trials) was assigned to the index finger, which is the fastest finger to initiate button presses. Thus, the particular finger assignments we used actually worked against the finding of the RT results obtained.

### The Processing of Ensemble Shape and Surface Properties in PPA and LO

When examining differences in activation across each hemisphere for all five adaptation conditions in PPA, the main effects of Hemisphere and Adaptation Condition were significant ($F(1, 14) = 5.54$, $p < .05$, and $F(4, 56) = 4.31$, $p < .005$, respectively), but the interaction between these factors was not ($F(4, 56) = 1.34$, $p = .27$). Because the adaptation findings were similar in the left and right hemispheres, we averaged data between the two hemispheres. As in our previous studies (Cant & Xu, 2012, 2015), we found similar levels of repetition suppression for "identical" and "shared" adaptation trials in PPA ($t(14) = 0.72$, $p = .49$; Figure 2C), reflecting the processing of ensemble rather than the exact arrangement of local shape features in this brain region. To evaluate the contribution of shape and surface properties to ensemble representation, we used data from the identical trials as our baseline adaptation condition (i.e., the condition where neither shape nor surface properties changed; see below for results when the shared trials were used as baseline). In

Table 1. Percent Correct Accuracies and Response Latencies (in msec) of Correct Trials for the Adaptation Runs

			Adaptation Runs		
	Identical	Shared	Same Shape Different Surface Properties	Different Shape Same Surface Properties	Different Shape Different Surface Properties
Response latency	876 ± 40	942 ± 40	779 ± 37	797 ± 31	729 ± 33
Accuracy	91.69 ± 1.77	86.51 ± 2.58	95.54 ± 1.23	98.39 ± 0.49	99.55 ± 0.31

All values represent means with standard errors.

Figure 3. The contributions of shape and surface properties to ensemble adaptation in PPA and LO using identical trials as the baseline fMRI adaptation condition. (A) In PPA, changes in both shape and surface properties contributed significantly to ensemble adaptation, with no interaction between the two. (B) In contrast, in LO, only changes in shape but not surface properties contributed to ensemble adaptation, with no interaction between the two. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus & Mason, 1994).
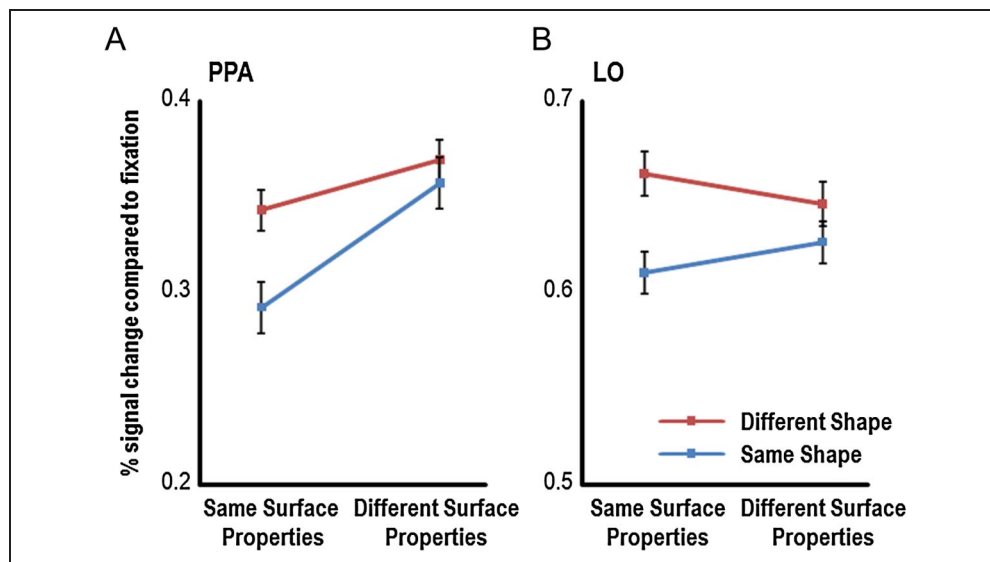
the resulting analysis, the main effects of Shape ($F(1, 14) = 7.19, p < .05$) and Surface Properties ($F(1, 14) = 14.21, p < .005$) were both significant, but the Shape-by-Surface Properties interaction was not ($F(1, 14) = 1.23, p = .29$; see Figure 3A; similar results were obtained in the analysis using the shared condition as baseline: main effect of Shape: $F(1, 14) = 4.50, p = .052$; main effect of Surface Properties: $F(1, 14) = 6.20, p < .05$; Shape-by-Surface Properties interaction: $F(1, 14) = 0.73, p = .41$). Planned pairwise comparisons (both two-tailed) revealed a significant release from adaptation when surface properties varied but shape did not (i.e., when changes in shape were held constant, there was higher activation in trials where surface properties varied, compared with trials where surface properties did not vary; $t(14) = 2.71, p < .05$; see Figures 2C and 3A) and when shape varied but surface properties did not ($t(14) = 2.43, p < .05$). This latter finding replicates our previous results (Cant & Xu, 2012) and, together with the release from adaptation observed with variations in surface properties, demonstrates that changing either feature alone is sufficient to cause a release from adaptation in PPA.

Although there may appear to be differences in initial baseline activation between conditions (i.e., activation levels for each condition at the beginning of a trial), we did not find any evidence of this, as a one-way repeated-measures ANOVA with factor Adaptation Condition revealed that differences in baseline activation were not significant (assessed using the first time point within a trial, T_1: $F(4, 56) = 0.24, p = .92$; assessed using the average of the first and second time points within a trial, T_{AVG}: $F(4, 56) = 0.41, p = .80$; see Figure 2B). Importantly, subtracting the baseline values for each condition from their resulting peak responses did not produce a significant Shape-by-Surface Properties interaction in the 2×2 ANOVA (subtracting T_1: $F(1, 14) = 2.07, p = .17$; subtracting T_{AVG}: $F(1, 14) = 3.29, p = .09$), but the planned

post hoc comparisons of a single feature change versus the identical condition remained significant (subtracting T_1, pairwise comparison when surface properties varied but shape did not: $t(14) = 2.25, p < .05$; subtracting T_1, pairwise comparison when shape varied but surface properties did not: $t(14) = 3.05, p < .05$; subtracting T_{AVG}, pairwise comparison when surface properties varied but shape did not: $t(14) = 2.28, p < .05$; subtracting T_{AVG}, pairwise comparison when shape varied but surface properties did not: $t(14) = 2.89, p < .05$). Thus, regardless of whether or not we correct for any potential differences in initial baseline activation, in PPA we observe significant contributions of both shape and surface properties to ensemble representation (i.e., changing either feature alone is sufficient to cause a release from adaptation) but no significant interaction between them. We return to these points later in the Discussion.

To explore whether the neural results in PPA could be explained solely by behavioral responses, we conducted two separate correlation analyses. In the first, we examined the relationship between neural adaptation data and behavioral measures of accuracy and response latency across the five different conditions separately. No significant correlations were found (all $rs < .33$ and $ps > .22$ for the relationship between neural data and accuracy, and all $rs > -.39$ and $ps > .15$ for the relationship between neural data and response latency). In the second analysis, we examined relationships between neural and behavioral data for the two adaptation effects that are reported above: namely, the release from adaptation when surface properties vary but shape does not (i.e., the difference between the identical and same shape/different surface properties conditions, or Adaptation$_1$) and the release from adaptation when shape varies but surface properties do not (i.e., the difference between the identical and different shape/same surface properties conditions, or Adaptation$_2$). Results for Adaptation$_1$

revealed no significant relationship between neural and behavioral data for accuracy ($r = .42, p > .11$), but a significant relationship for response latency ($r = −.55, p < .05$). No significant correlations were observed between neural and behavioral data for Adaptation$_2$ (accuracy: $r = .13, p > .31$; response latency: $r = −.40, p > .13$). Taken together, these results thus demonstrate that the shape and surface properties of the elements that make up an ensemble both contribute significantly to ensemble representation in PPA, that the two contributions may be additive (but see Discussion), and that behavioral responses alone cannot explain this pattern of neural adaptation results.

When examining differences in activation across each hemisphere for all five adaptation conditions in LO (see Figure 2C), the main effect of Hemisphere was not significant ($F(1, 14) = 0.37, p > .05$), but the main effects of Adaptation Condition ($F(4, 56) = 2.82, p < .05$) and the interaction between Hemisphere and Adaptation Condition ($F(4, 56) = 4.16, p < .01$) were both significant. Given the role of LO in shape processing, as in our previous studies (Cant & Xu, 2012, 2015), we again found a significant release from adaptation in LO for "shared" trials, compared with "identical" trials ($t(14) = 2.85, p < .05$; see Figure 2C). This likely results from the variations in the arrangement of local shape contours that are present in the shared, but not the identical, condition, and may indicate that LO plays a role in shape processing at multiple levels (i.e., from the contours of local individual objects to the arrangement of local shape contours and the geometry of the global background/surface they are situated upon, because rotating multiple objects in an ensemble changes local shape contour arrangement and the relationship between foreground and background elements in the image). To evaluate the contribution of shape and surface properties to adaptation in LO, we thus used data from the identical trials as our baseline. To compare directly with the results in PPA, we averaged data between the two hemispheres (but see below for the results in each hemisphere separately). With these measures, we found that the main effect of Shape was significant ($F(1, 14) = 6.92, p < .05$), but the main effect of Surface Properties ($F(1, 14) = 0.001, p = .99$) and the Shape-by-Surface Properties interaction ($F(1, 14) = 2.03, p = .18$) were not (see Figure 3B). Differences in baseline activation across the five adaptation conditions were not significant (using T$_1$: $F(4, 56) = 0.43, p = .79$; using T$_{AVG}$: $F(4, 56) = 0.23, p = .92$; see Figure 2B).

When examining relationships between neural and behavioral data for each condition separately, we found significant correlations between neural adaptation data and accuracy on same trials ($r = .63, p < .05$) and between neural data and response latency on same trials ($r = −.57, p < .05$), but all other correlations were nonsignificant. Moreover, no significant correlations were observed when examining relationships between neural and behavioral data for specific adaptation effects (accuracy for Adaptation$_1$: $r = .41, p > .12$; response latency

for Adaptation$_1$: $r = −.34, p > .21$; accuracy for Adaptation$_2$: $r = .33, p > .23$; response latency for Adaptation$_2$: $r = −.36, p > .18$). These results provide additional evidence for the notion that LO is sensitive to processing shape features from object ensembles but is not sensitive to processing surface properties (Cant & Xu, 2012; Cant & Goodale, 2007; Malach et al., 1995) and reinforces the idea that behavioral measures alone cannot explain our neural results.

When we analyzed the data in LO for each hemisphere separately, the main effect of Hemisphere interacted significantly with the main effect of Shape ($F(1, 14) = 5.47, p < .05$), such that the main effect of Shape was significant in the left but not in the right hemisphere ($F(1, 14) = 7.68, p < .05$; and $F(1, 14) = 0.07, p = .79$, respectively). The main effect of Hemisphere did not interact with the main effect of Surface Properties ($F(1, 14) = 0.023, p = .64$) and neither hemisphere showed a significant main effect of Surface Properties (both $ps > .76$). Finally, the three-way interaction of Hemisphere, Shape, and Surface Properties was significant ($F(1, 14) = 14.80, p < .005$), such that the two-way interaction of Shape and Surface Properties was marginally significant in the left but not the right hemisphere ($F(1, 14) = 4.42, p = .054$; and $F(1, 14) = 0.42, p = .53$, respectively). These results indicate that ensemble shape adaptation is much stronger in the left than the right LO, which differs from our previous findings (Cant & Xu, 2012, 2015). We will discuss this in detail in the Discussion.

Comparison between brain regions revealed that LO and PPA differed significantly with regard to surface property processing (Region-by-Surface Properties interaction: $F(1, 14) = 16.06, p < .001$), but not shape processing (Region-by-Shape interaction: $F(1, 14) = 0.15, p = .71$). Moreover, when examining differences in activation across LO and PPA for all five adaptation conditions, we observed a highly significant Region-by-Adaptation Condition interaction ($F(4, 56) = 9.70, p < .001$). Together, these results demonstrate that these regions process the same visual inputs using both distinct and potentially similar neural mechanisms, respectively, depending on the nature of the information being processed.

Comparing Ensemble Shape and Surface Property Processing in PPA, RSC, and OPA

We also examined the adaptation results in two other regions in the human scene-processing network: the RSC and OPA (see Dilks et al., 2013; Epstein & Higgins, 2007). In RSC, the analysis of all five adaptation conditions (see Figure 4C) revealed nonsignificant main effects of Hemisphere ($F(1, 14) = 0.99, p = .34$) and Adaptation Condition ($F(4, 56) = 0.45, p = .78$), but a significant interaction between these factors ($F(4, 56) = 2.64, p < .05$; however, further investigation revealed no significant main effects, interactions, or pairwise comparisons in either hemisphere

separately, revealing little evidence for hemispheric differences in RSC). To facilitate comparison with PPA, we collapsed the data across hemispheres and used data from the identical trials as our baseline adaptation condition. This analysis revealed no significant results (main effect of Shape: $F(1, 14) = 1.25, p = .28$; main effect of Surface Properties: $F(1, 14) = 0.30, p = .59$; Shape-by-Surface Properties interaction: $F(1, 14) = 0.67, p = .43$ see Figure 5A; baseline differences assessed at T_1: $F(4, 56) = 0.71, p = .59$; using T_{AVG}: $F(4, 56) = 1.03, p = .40$; see Figure 4B), only one significant correlation between neural and behavioral data when examining each condition separately (correlation between neural data and accuracy on shared trials: $r = .52, p < .05$; all other correlations were nonsignificant), and no significant correlations between neural and behavioral data when examining specific adaptation effects (i.e., Adaptation$_1$ and Adaptation$_2$; all $rs < .38$ and $ps > .16$ for the relationship between neural data and accuracy, and all $rs > -.16$ and $ps > .59$ for the relationship between neural data and response latency.

In OPA, the analysis of all five adaptation conditions (see Figure 4C) revealed no significant effect of Hemisphere ($F(1, 14) = 3.40, p = .087$) and Adaptation Condition ($F(4, 56) = 1.82, p = .14$) and a nonsignificant interaction between the two ($F(4, 56) = 1.33, p = .27$). After collapsing data from the right and left hemispheres and using data from the identical trials as the baseline adaptation condition, no significant results were observed (main effect of Shape: $F(1, 14) = 1.90, p = .19$; main effect of Surface Properties: $F(1, 14) = 0.57, p = .46$; Shape-by-Surface Properties interaction: $F(1, 14) = 0.02, p = .90$; see Figure 5B; baseline differences assessed at T_1: $F(4, 56) = 0.87, p = .49$; using T_{AVG}: $F(4, 56) = 0.71, p = .59$; see Figure 4B), and only one correlation between neural and behavioral data was significant when examining each condition separately (correlation between neural data and accuracy on same trials: $r = .54, p < .05$; all other correlations were nonsignificant). Similarly, a significant relationship between neural and behavioral data was observed for accuracy but not response latency for both Adaptation$_1$

Figure 4. Example stimuli and results ($N = 15$) for all five fMRI adaptation conditions in RSC and OPA. (A) The stimuli used, shown here again as a reminder of the adaptation conditions. (B) Full average time course of activation from independently and individually defined RSC and OPA ROIs. The first 4 time points represent the length of a single event-related trial. (C) Averaged fMRI responses (representing the time point of greatest signal amplitude in the average response across all conditions) from RSC and OPA, shown at the appropriate scales for better visualization of the differences (or lack thereof) across the adaptation conditions. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus & Masson, 1994).
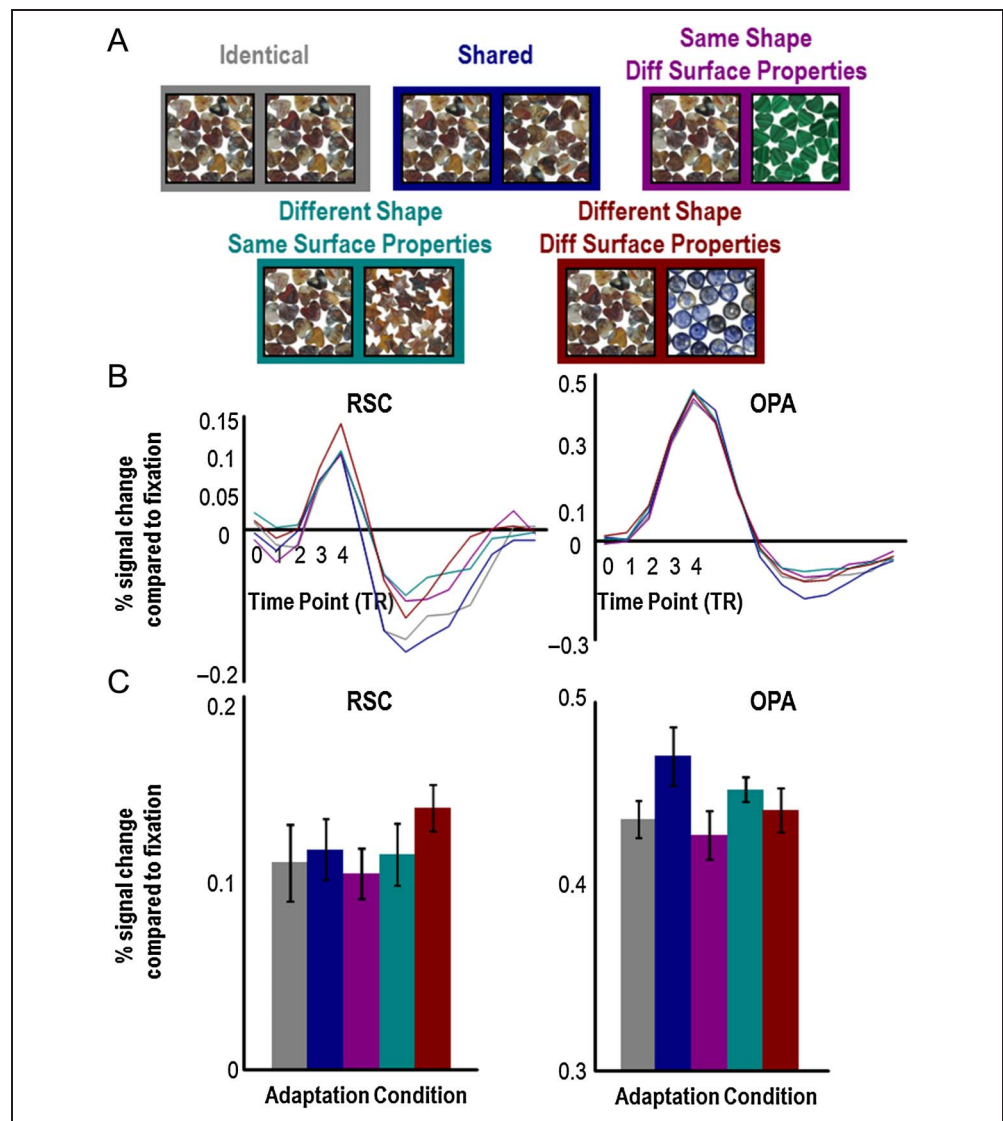
Figure 5. The contributions of shape and surface properties to ensemble adaptation in RSC and OPA using identical trials as the baseline fMRI adaptation condition. Neither shape nor surface properties contributed to ensemble adaptation in (A) RSC and (B) OPA. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus & Masson, 1994).
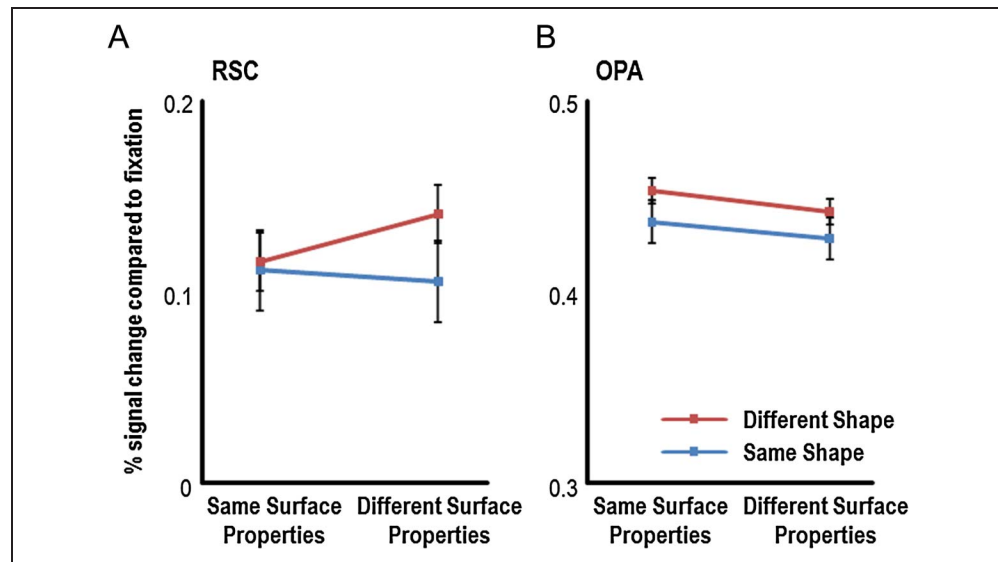
(accuracy: $r = .62, p < .05$; response latency: $r = -.38, p > .16$) and Adaptation$_2$ (accuracy: $r = .70, p < .005$; response latency: $r = -.50, p > .059$).

The data from RSC and OPA differ markedly from that observed in PPA, but to provide more direct evidence of this, we compared the patterns of activation observed in PPA with those in RSC and OPA. In both cases, we obtained evidence that the representation of ensembles in RSC and OPA are not mirror images of the representation in PPA (significant Region-by-Surface Properties interaction between PPA and RSC: $F(1, 14) = 7.56, p < .05$; significant Region-by-Shape-by-Surface Properties interaction between PPA and RSC: $F(1, 14) = 5.91, p < .05$; and significant Region-by-Surface Properties interaction between PPA and OPA: $F(1, 14) = 37.91, p < .001$; all remaining interactions involving the factor Region were not significant: all Fs < 2.16, all ps > .16). Moreover, when examining differences in activation across all five adaptation conditions, we observed significant Region-by-Adaptation Condition interactions for both the comparison of PPA versus RSC ($F(4, 56) = 4.65, p < .005$) and PPA versus OPA ($F(4, 56) = 11.36, p < .001$). Taken together, these findings demonstrate that, with regard to object ensemble processing, the activation in PPA may be functionally dissociated from the processing observed in other scene-selective regions of cortex, which is consistent with our previous findings (Cant & Xu, 2012, 2015).

DISCUSSION

Object ensemble processing involves the extraction of summary statistical information from large collections of objects at the expense of being able to provide fine details about any individual object within the ensemble (for a review, see Alvarez, 2011). In this way, object ensemble processing is adaptive as it allows the visual system to circumvent the capacity limitation involved with the pro-

cessing of single objects (e.g., Luck & Vogel, 1997; Pylyshyn & Storm, 1988; see also Cowan, 2001). We have previously shown that the processing of object ensembles recruits a region of anterior-medial ventral visual cortex that overlaps with the scene-selective PPA and that the ensemble representation in this region is not driven by changes in the size or absolute density of ensembles but is sensitive to processing changes in the ratio or relative density of elements in heterogeneous ensembles (Cant & Xu, 2012, 2015). Here we expanded upon our knowledge of the functional properties of this region and examined whether shape and surface properties, two important features of single object representation, may contribute to the neural representation of object ensembles in PPA. We found that PPA codes changes in both the shape and surface properties of ensemble elements, such that changing either feature alone was sufficient to cause a release from fMRI adaptation. This suggests that both shape and surface properties contribute significantly to the ensemble code in anterior-medial ventral visual cortex.

Object ensemble representation in PPA thus differs from single object representation in this brain region because only surface properties, but not shape, are coded in PPA during the processing of single objects (Cant & Goodale, 2007). This suggests that PPA may only be engaged in shape processing when the extraction of shape summary statistics is needed but is not involved in the detailed processing of object shape features per se. This is consistent with our earlier proposal that PPA may be generally involved in extracting summary statistics essential for ensemble representation (Cant & Xu, 2012, 2015).

In this study, to more truthfully reflect the representation of real-world object ensembles in the brain, we studied object ensembles made from real-world materials with naturally occurring surface properties. Because natural materials are usually defined by unique conjunctions

of colors and textures, we could not vary these two features independently. However, we have previously shown that holding ensemble shape constant but varying the color of ensemble elements does not produce a release from adaptation in PPA (Cant & Xu, 2015). Thus, the contribution of surface properties to ensemble representation in PPA likely comes from the effect of texture or a combined effect of color and texture. Given the presence of unique conjunctions of colors and textures in natural materials, there is likely a tight integration between these two features during surface property perception. Further studies are needed to fully understand the interactions between texture and color in ensemble representation.

In the object-selective LO, we observed a release from adaptation only when ensemble shape varied and adaptation when shape repeated but surface properties varied. This is consistent with previous reports examining the processing of shape and texture in LO in single-object perception (e.g., Cant, Arnott, & Goodale, 2009; Cant & Goodale, 2007; Malach et al., 1995) and in object ensemble perception (Cant & Xu, 2012, 2015). Unlike our previous findings (Cant & Xu, 2012, 2015), however, here we also found a hemisphere effect in LO such that ensemble shape adaptation was much stronger in the left than the right LO. One possibility for this finding relates to the putative roles of the left and right hemisphere in local and global processing, respectively (e.g., Fink et al., 1997; Van Kleeck, 1989). In this study, the processing of the shapes of the ensemble objects required processing at the local level. This likely recruited the left hemisphere more heavily than the right hemisphere and resulted in the observed response pattern in LO. However, as this hemisphere effect was not observed in our previous studies (Cant & Xu, 2012, 2015), it does not appear to be a robust and consistent effect. Further replication and validation of this effect is needed, possibly with the explicit manipulation of the scope of visual attention at the local and global levels.

Together with our previous findings, these results show that, with identical visual input, PPA and LO are engaged in different aspects of visual information processing, with PPA engaged in processing multiple ensemble features and LO in processing shape information both from single objects and object ensembles.

The Relationship between Shape and Surface Properties in Ensemble Representation

Our data do not unambiguously support either an additive or conjoint representation of ensemble shape and surface properties in anterior-medial ventral visual cortex. On the one hand, we found no significant interaction between the adaptation effect of shape and surface properties in PPA (both in our main analysis and in two analyses where initial baseline differences were subtracted from peak responses for each condition). This raises the possibility that the representation of ensemble shape and surface properties may be additive and independent in this brain region. Indeed, because shape and surface properties are computed somewhat independently during early stages of visual processing, it is possible that distinctive (but potentially intermixed) PPA neurons may be receiving these two types of inputs from early visual areas, making the representations of shape and surface properties relatively separate in this brain region.

On the other hand, however, it is also important to note that we only observed adaptation in PPA when both ensemble shape and surface properties were repeated compared with when one of these features varied or when both features varied. Compared to when both features varied, repeating one of these features did not result in significant adaptation (and in fact the levels of activation did not differ across these three conditions; compare the purple, green, and red bars for PPA in Figure 2C). This argues against a strong additive and independent account and indicates some conjoint representation of shape and surface properties in ensemble representation in PPA. Consistent with this, in a recent behavioral study, we found that observers could not ignore changes in shape while making discriminations of ensemble surface properties (i.e., texture) and vice versa, leading us to conclude that shape and surface properties may not be processed completely independently in object ensemble perception (Cant et al., 2015). If there is indeed a significant interaction between the neural representation of ensemble shape and texture, given that we failed to find evidence of this in three separate analyses with the 15 participants tested in this study, it is likely a small and weak effect. It is possible that we are underpowered to detect this significant interaction in this study, but we should note that our sample size is not small compared with many visual perception fMRI studies, and importantly, the results do not qualitatively change when using two versus three adaptation runs (a difference of 24 vs. 36 trials/condition, respectively). Because our data cannot unambiguously support the additive or conjoint representation hypotheses, more experiments are needed to fully understand the relationship between shape and surface properties in ensemble representation. Regardless, the main contribution of this study is to demonstrate that both of these features contribute significantly to ensemble representation in anterior-medial ventral visual cortex.

The Relationship among Object Ensemble, Texture, and Scene Representations in Anterior-medial Ventral Visual Cortex

Ensembles and texture patterns are clearly related, in that both contain repeating elements that can vary in visual features such as size and orientation (Portilla & Simoncelli, 2000) and both are represented by engaging in a global statistical extraction of numerous features in the visual

field, which often occurs in the visual periphery (see Cohen, Dennett, & Kanwisher, 2016; Rosenholtz, Huang, Raj, Balas, & Ilie, 2012; Rosenholtz, 2011). It then came as no surprise that texture processing also engages anterior-medial ventral visual cortex (Cant & Xu, 2012; Cant & Goodale, 2007). However, as we showed in this study, ensemble representation need not be viewed as simply representing the texture patterns or the surface properties of the ensemble elements, as keeping these properties constant but varying the shape of the ensemble elements still resulted in a release from adaptation in PPA (see also Cant & Xu, 2012, 2015). Thus, one possibility is that, although ensemble and texture representations may engage similar processing mechanisms, they are not identical but tap into representations at distinct levels of visual information processing. An alternative possibility, however, is that changing the shape of ensemble elements is perceived as a global texture change (e.g., the texture gestalt changes from multiple stars to multiple hearts), despite being produced by local changes to outline shape, and it is this global texture change that is responsible for the release from adaptation in PPA in this condition. Nevertheless, this account would still contend that texture processing occurs at multiple distinctive levels, at both the individual object level and the global level. It is then just a matter of semantics what should be considered a texture pattern, as according to Portilla and Simoncelli (2000), all images can be considered as texture patterns, including a face image. Future studies comparing models of texture and ensemble representations in PPA are required to help us understand how texture and ensemble, or local and global textures, are computed and represented in this brain region.

It is worth noting that another type of perceptual process that involves the statistical extraction of global features (as opposed to detailed processing of individual objects) is scene perception (Oliva & Torralba, 2001; for a review, see Oliva, Park, & Konkle, 2011). This commonality in processing likely explains why we observe sensitivity to both ensemble and texture features in PPA, a well-known scene-selective region (e.g., Epstein & Kanwisher, 1998). Thus, the extraction of summary statistics from visual inputs seems to occur at three distinct levels of visual processing: at the level of material and surface property perception, at the level of object ensemble perception, and at the level of scene perception. Given the relative ease of manipulating object ensembles compared with that of surface properties and scene contents, studying object ensemble processing is thus a bridge to furthering our understanding of the general mechanisms underlying both texture and scene perception and ultimately will increase our knowledge of the cognitive and neural mechanisms involved in these important aspects of human visual perception.

We should note, however, that sensitivity to processing object ensembles may be unique to PPA, as two additional regions in the human scene-processing network, RSC and

OPA, did not show the same pattern of adaptation when observers were processing the ensemble images. This is consistent with our previous investigations of object ensemble and texture processing (Cant & Xu, 2012, 2015) and suggests that there may be a functional dissociation in the human scene-processing network, with PPA involved in both spatial (e.g., spatial expanse; Kravitz, Peng, & Baker, 2011) and nonspatial aspects of visual processing (e.g., object ensemble and texture processing; Cant & Xu, 2012, 2015), whereas RSC and OPA may only be involved in spatial aspects of visual processing. This suggestion is certainly consistent with the separate (but complementary) functional roles posited for these three regions in the representation of scenes (see Epstein, 2008, for a review).

The Relationship between Neural and Behavioral Data

When using identical trials as the baseline, in both RT and accuracy data, we found significant effects of shape and surface properties but no interaction between the two. In general, responses were slower and less accurate on trials where ensemble features repeated, compared with trials where features varied. This is likely because a full inspection of the ensemble is needed to support a "same" or a "shared" response, whereas an inspection of part of the ensemble is sufficient to support a "different" response. Additionally, a two-stage processing procedure may be required to differentiate between the "same" and "shared" conditions, with the first stage involving the detection of ensemble feature repetition and the second stage requiring the inspection of the spatial arrangement of ensemble elements. This additional processing may have led to the slowest and least accurate responses for both the same and the shared trials.

Regardless, we do not believe that behavioral responses alone can explain our pattern of fMRI adaptation results, for a number of reasons. First, overall accuracy across the different conditions was quite high (averaging 96%), indicating that observers were alert, engaged in the task, and performing near ceiling. Second, there is no evidence of a speed–accuracy tradeoff in our results, as observers responded slowest and least accurately in the same conditions (i.e., the identical and shared trials). This reveals that there was likely no obvious cognitive strategy in operation during the behavioral task (e.g., slow responses down to respond more accurately), and thus, the fMRI results are not likely explained by the use of a particular behavioral response strategy. Third, despite sharing some similar patterns of significance, the direction of speed and accuracy responses do not match well with our fMRI adaptation results. For example, accuracy was lowest in conditions where a stimulus feature repeated and highest where a stimulus feature varied (but this relationship was not significant for surface property trials), yet we observed the lowest fMRI activation in trials where a stimulus feature repeated.

Likewise, response latency was longest in conditions where a stimulus feature repeated and shortest where a stimulus feature varied. It is difficult to conceive how trials that were more difficult and required more time and mental effort to classify would be associated with less fMRI activation. Fourth, and related to the previous point, there was not strong evidence of correlations between neural adaptation effects and behavioral measures of response latency or accuracy across two different correlation analyses, particularly in PPA. Although a small number of significant correlations were observed in lateral occipital complex, OPA, and RSC, together with the results in PPA, these results reinforce the idea that behavioral responses alone cannot explain our pattern of fMRI adaptation results. Finally, and perhaps most importantly, previous findings by Xu and colleagues (2007) have demonstrated that fMRI adaptation responses in PPA are dissociable from behavioral responses.

Conclusions

We have provided evidence that anterior-medial ventral visual cortex is sensitive to processing the shape and surface properties of the elements that constitute an ensemble. Moreover, this ensemble processing is distinct from that observed in lateral occipital cortex, which is sensitive to processing the shape, but not the surface properties, of object ensembles. This functional dissociation is consistent with our previous results (Cant & Xu, 2012, 2015) and reinforces an idea that we have put forward in recent years. Namely, there are (at least) two separate but complementary neural processing pathways in visual cortex, with each being involved in distinctive aspects of visual information processing. One pathway, which includes regions of lateral occipital (i.e., LO) and parietal cortex (Xu & Chun, 2009), is capacity limited and is involved in the individuation, perception, and identification of the detailed features of both single objects and objects within an ensemble. Another pathway, which includes regions of anterior-medial ventral visual cortex (i.e., the scene-selective PPA), is not strictly capacity limited in the classical sense and is involved in the statistical extraction of multiple visual features from our environment for use in tasks such as texture and material perception (Cant & Xu, 2012; Cant & Goodale, 2007, 2011), ensemble perception (Cant & Xu, 2012, 2015, and this study), scene perception, and navigation (Epstein, 2008). This proposal is consistent with Wolfe, Vo, Evans, and Greene's (2011) dual visual pathway model, with one selective and capacity-limited channel and one nonselective and capacity-unlimited channel. Together, both of these pathways, and the communication between them, contribute to the production of skilled and adaptive behavior.

Acknowledgments

REFERENCES

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15,* 122–131.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111.

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19,* 392–398.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,* 157–162.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Cant, J. S., Arnott, S. R., & Goodale, M. A. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research, 192,* 391–405.

Cant, J. S., & Goodale, M. A. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex, 17,* 713–731.

Cant, J. S., & Goodale, M. A. (2011). Scratching beneath the surface: New insights into the functional properties of the lateral occipital area and parahippocampal place area. *Journal of Neuroscience, 31,* 8248–8258.

Cant, J. S., Sun, S. Z., & Xu, Y. (2015). Distinct cognitive mechanisms involved in the processing of single objects and object ensembles. *Journal of Vision, 15,* 12.

Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *Journal of Neuroscience, 32,* 7685–7700.

Cant, J. S., & Xu, Y. (2015). The impact of density and ratio on object ensemble representation in human anterior-medial ventral visual vortex. *Cerebral Cortex, 25,* 4226–4239.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404.

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences, 20,* 324–335.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Dilks, D. D., Julian, J. B., Kubilius, J., Spelke, E. S., & Kanwisher, N. (2011). Mirror-image sensitivity and invariance in object and scene processing pathways. *Journal of Neuroscience, 31,* 11305–11312.

Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience, 33,* 1331–1336.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392,* 598–601.

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences, 12,* 388–396.

Epstein, R. A., & Higgins, J. S. (2007). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cerebral Cortex, 17,* 1680–1693.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1997). Neural mechanisms

involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain, 120,* 1779–1791.

Friston, K. J., Homes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackwowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear model approach. *Human Brain Mapping, 2,* 189–210.

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., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3,* 837–843.

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: A fMRI study. *Brain, 126,* 2463–2475.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293,* 1506–1509.

Kravitz, D. J., Peng, C. S., & Baker, C. I. (2011). Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *Journal of Neuroscience, 31,* 7322–7333.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review, 1,* 476–490.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281.

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, U.S.A., 92,* 8135–8139.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., et al. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences, U.S.A., 89,* 5951–5955.

Oliva, A., Park, S., & Konkle, T. (2011). Representing, perceiving, and remembering the shape of visual space. In L. R. Harris & M. Jenkin (Eds.), *Vision in 3D environments* (pp. 308–339). Cambridge, UK: Cambridge University Press.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42,* 145–175.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4,* 739–744.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision, 40,* 49–71.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197.

Rosenholtz, R. (2011). What your visual system see where you are not looking. In B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings of SPIE: Human Vision and Electronic Imaging XVI* (Vol. 7865, pp. 1–14). San Francisco, CA.

Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision, 12,* 14.

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *Neuroimage, 30,* 1088–1096; discussion 1097-1089.

Steeves, J. K. E., Humphrey, G. K., Culham, J. C., Menon, R. S., Milner, A. D., & Goodale, M. A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience, 16,* 955–965.

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

Todd, J. J., Han, S. K., Harrison, S., & Marois, R. (2011). The neural correlates of visual working memory encoding: A time-resolved fMRI study. *Neuropsychologia, 49,* 1527–1536.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751–754.

Van Kleeck, M. H. (1989). Hemispheric differences in global versus local processing of hierarchical visual stimuli by normal subjects: New data and a meta-analysis of previous studies. *Neuropsychologia, 27,* 1165–1178.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 32,* 931–941.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24,* 55–62.

Wolfe, J. M., Vo, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and non-selective pathways. *Trends in Cognitive Sciences, 15,* 77–84.

Xu, Y. (2002). Limitations in object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 458–468.

Xu, Y. (2010). The neural fate of task-irrelevant features in object-based processing. *Journal of Neuroscience, 30,* 14020–14028.

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature, 440,* 91–95.

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences, 13,* 167–174.

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.