The Impact of Density and Ratio on Object-Ensemble Representation in Human Anterior-Medial Ventral Visual Cortex

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

[1]Department of Psychology, University of Toronto Scarborough, Toronto, ON, Canada and [2]Visions Sciences Laboratory, Department of Psychology, Harvard University, Cambridge, MA, USA

Address correspondence to Jonathan S. Cant, 1265 Military Trail, Science Wing Room 411, Toronto, ON, Canada, M1C 1A4. Email: jonathan.cant@utoronto.ca

Behavioral research has demonstrated that observers can extract summary statistics from ensembles of multiple objects. We recently showed that a region of anterior-medial ventral visual cortex, overlapping largely with the scene-sensitive parahippocampal place area (PPA), participates in object-ensemble representation. Here we investigated the encoding of ensemble density in this brain region using fMRI-adaptation. In Experiment 1, we varied density by changing the spacing between objects and found no sensitivity in PPA to such density changes. Thus, density may not be encoded in PPA, possibly because object spacing is not perceived as an intrinsic ensemble property. In Experiment 2, we varied relative density by changing the ratio of 2 types of objects comprising an ensemble, and observed significant sensitivity in PPA to such ratio change. Although colorful ensembles were shown in Experiment 2, Experiment 3 demonstrated that sensitivity to object ratio change was not driven mainly by a change in the ratio of colors. Thus, while anterior-medial ventral visual cortex is insensitive to density (object spacing) changes, it does code relative density (object ratio) within an ensemble. Object-ensemble processing in this region may thus depend on high-level visual information, such as object ratio, rather than low-level information, such as spacing/spatial frequency.

Keywords: density, fMRI, object-ensemble perception, scene perception, ventral visual stream

Introduction

The perception and recognition of single objects has been a fruitful enterprise of cognitive neuroscientific research for decades. While we have learned much from this research, it remains that objects are rarely seen in isolation in everyday life, and are often part of a larger collection, or ensemble, of multiple similar objects. Indeed, object ensembles are ubiquitous in our visual world (e.g., leaves on a tree). Importantly, the representation of summary statistics from ensembles of multiple objects complements and guides object-specific processing since it allows the visual system to overcome the capacity limitation inherent in object-based attention (e.g., Luck and Vogel 1997; Pylyshyn and Storm, 1998; Xu 2002; Alvarez and Cavanagh 2004). Within the past decade, behavioral studies have demonstrated that observers can extract summary information from large collections of objects, such as their mean size, direction of motion, speed, orientation, and center location, without being able to provide fine details about any individual object in the ensemble (e.g., Williams and Sekuler 1984; Watamaniuk and Duchon 1992; Ariely 2001; Parkes, et al. 2001; Chong and Treisman 2003; Alvarez and Oliva 2008). For clarity, here "object ensemble" refers to a collection of objects

whose number exceeds the processing capacity of individual object-based attention (i.e., above 4 or 5 objects), that are perceptually grouped together using Gestalt principles such as proximity. Our definition is in agreement with a recent review on ensemble processing, which defines "ensemble representation" as a general process of computing information from multiple items, collapsing this information into useable and adaptive forms such as summary statistics (Alvarez 2011), and argues that ensemble processing applies to the extraction of statistical information from both low-level (e.g., mean size: Ariely 2001; mean brightness: Bauer 2009) and high-level visual information (e.g., mean emotion and gender of faces: Haberman and Whitney 2007).

Object ensembles resemble surface textures in that both contain repeating structures with slight variations in features such as size, orientation, and color (Portilla and Simoncelli 2000). Thus, the extraction of summary statistics is essential in the representations of both ensembles and textures. Indeed, in a recent series of fMRI experiments (Cant and Xu 2012), we found that ensembles and textures share similar neural processing substrates for summary statistics in anterior and medial regions of the ventral visual cortex, along the collateral sulcus and overlapping to a large extent with the parahippocampal place area (PPA). Specifically, we observed fMRI adaptation in this brain region whenever summary statistics repeated in object ensembles and surface textures, even when local shape features differed across images.

While the processing of surface texture in the general region around PPA has been noted previously (Peuskens et al. 2004; Cant and Goodale 2007, 2011; Cant et al. 2009), PPA is best known for playing a large role in scene perception, specifically by processing the 3D spatial structure, or geometry, of scenes (Epstein and Kanwisher 1998). Meanwhile, scene processing often requires the extraction of overall scene gist without representing the individual objects comprising the scene in great detail (e.g., Oliva and Schyns 2000; Oliva and Torralba 2001). Perhaps it is this aspect of scene processing that enables PPA to represent ensemble and texture stimuli even though they contain virtually no 3D spatial or scene information. Thus, PPA may play a greater role in extracting summary statistics from a variety of visual stimuli (including scenes, ensembles and textures), beyond its role in processing the 3D spatial structure of scenes. It is worth stating, however, that summary representation is not a general processing feature of all scene-sensitive regions in the brain, as it was not seen in the retrosplenial complex (RSC) and the transverse occipital sulcus (TOS) in our previous study (Cant and Xu 2012).

In the present study, we sought to further our understanding of the nature of the neural object-ensemble representation in

anterior-medial ventral visual cortex. There are a number of ensemble visual properties that need to be investigated to achieve this aim, and we have previously examined the representation of the size and color of ensembles as well as the shape and texture of individual elements in ensemble processing (Cant and Xu 2011, 2012). Other possible visual features include the overall brightness of object ensembles, and their location in the visual field, but such features are largely accidental and thus may not be diagnostic of perceived ensemble identity. In the present study, using the fMRI-adaptation approach (Grill-Spector, Henson, and Martin 2006) that we employed previously (Cant and Xu 2012), we conducted 3 experiments to investigate 2 specific visual features that may play a role in object-ensemble representation: absolute and relative density. In Experiment 1, we investigated whether or not PPA (see Fig. 1) would be sensitive to changes in the absolute density (or spacing) of the elements that comprise a homogeneous object ensemble (i.e., an ensemble containing only one type of object; see Fig. 2). On the one hand, naturally occurring real-world object ensembles vary in density (e.g., the amount of leaves on a tree over the course of a year), making density an informative feature that would matter in ensemble representation. Moreover, density covaries with number, spatial frequency and the level of clutter. As such, a change in density could evoke a strong neural response in this brain region. However, this also makes it difficult to attribute any neural response uniquely to absolute density changes and not to changes in number or spatial frequency. On the other hand, density changes may be encoded and processed by early visual areas, leaving later areas such as PPA to only encode higher-level and more abstract visual information. Indeed, we have previously shown that ensemble processing in PPA is not modulated by an overall size change of the ensemble images (Cant and Xu 2012). Density may also be considered as an accidental, rather than a diagnostic, feature of an ensemble as a change in density does not alter the mean features of the objects comprising the ensemble (such as mean size and mean texture). As such, a brain region computing and representing ensemble statistics may not be sensitive to changes in absolute density.

In Experiment 2, instead of focusing on the absolute density of homogeneous ensembles, we investigated whether or not PPA would show sensitivity to processing relative density by varying the ratio, or proportion, of 2 types of objects comprising a heterogeneous ensemble (see Fig. 3). Unlike the absolute density changes studied in Experiment 1, the relative density of 2 types of objects in an ensemble is a diagnostic and informative feature of an ensemble. For example, the ratio between leaves and fruit on a tree can inform us whether or not a particular tree is a good source of food, and the ratio between occupied and empty seats in a lecture hall can inform us whether or not a lecture is popular. A brain region computing and representing ensemble statistics should therefore be sensitive to changes in the relative density of 2 types of objects composing an ensemble.

In Experiment 3, we examined in detail whether potential sensitivity to ensemble ratio changes in PPA would be driven by changes in low-level visual features. To achieve this we held the shape and identity of objects within an ensemble constant and defined a ratio change solely by a change in the color of the ensemble elements (see Fig. 4). On the one hand, if PPA does represent ensemble ratio but shows no sensitivity to processing ratio changes defined by a change in color, then this would argue that the ensemble representation in anterior-medial ventral visual cortex may be based on relatively higher-level visual information. On the other hand, if PPA shows sensitivity to changes in the distribution of color within an ensemble, then this would argue that the ensemble representation in this region of cortex may be based on lower-level visual information.

In Experiments 1 and 2, besides PPA, we also examined patterns of adaptation in other scene-processing regions (RSC and TOS), and in all experiments we examined patterns of adaptation in the lateral occipital area (LO), a region known to play a key role in processing the shape of single objects (Malach et al. 1995; Grill-Spector, et al. 1998; Kourtzi and Kanwisher 2001; Cant and Goodale 2007, 2011). We previously showed that LO also extracts shape information from object ensembles (Cant and Xu 2012), such that this brain region showed a release from adaptation when local shape features varied across images, regardless of whether or not summary statistics repeated in the object ensembles. In addition to ROI-based analyses, we also conducted whole-brain analyses to identify other brain regions that may be involved in ensemble processing.
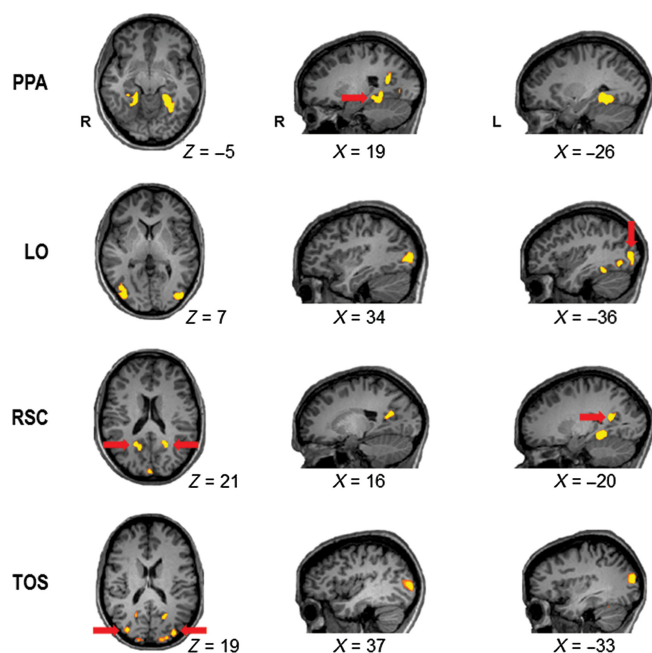

Figure 1. Examples of ROIs in individual observers. The scene-selective PPA (Talairach coordinates for the specific ROI example shown, *x, y, z* for right/left: +24/−26, −44/ −44, −1/−4) was defined by contrasting the activation for scenes against the activation for both faces and objects. The object-selective LO (+36/−40, −77/−77, +10/+4) was defined by contrasting the activation for objects against the activation for scrambled objects. The scene-selective RSC (+17/−20, −56/−53, +22/+18) and TOS (+35/−33, −78/−83, +10/+17) were defined by contrasting the activation for scenes against the activation for both faces and objects. PPA, parahippocampal place area; LO, lateral occipital area; RSC, restrosplenial complex; TOS, transverse occipital sulcus.

Materials and Methods

Observers
Experiment 1 included 8 paid observers (7 female, 1 male; mean age = 24.22, range = 19–31 years), Experiment 2 included 12 paid observers
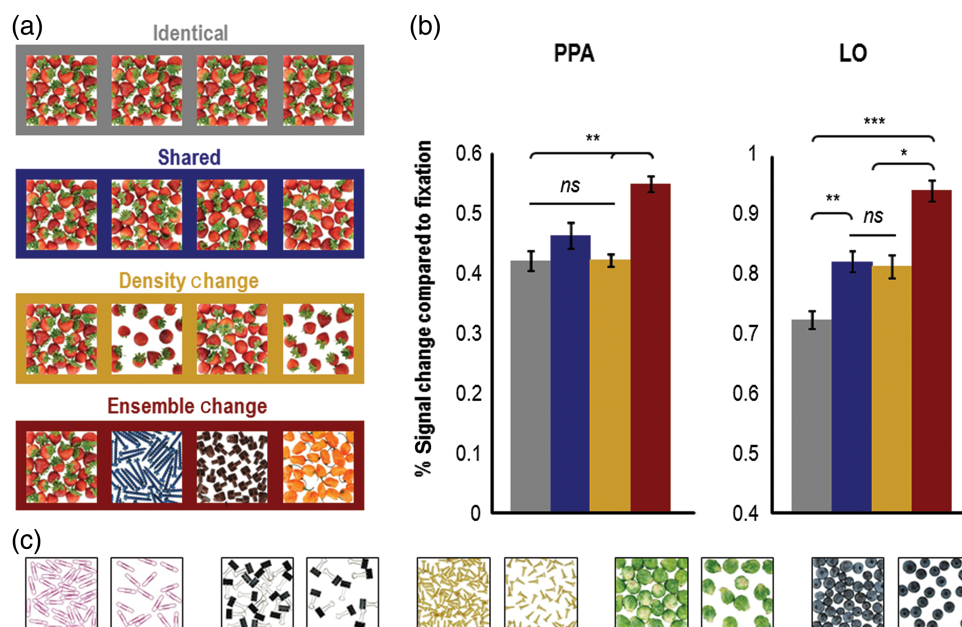
Figure 2. Example stimuli and results ($N = 8$) from Experiment 1. (*a*) Example stimuli used in the experiment. The stimuli used in the adaptation runs of Experiment 1 consisted of 20 different full-color photographs of homogeneous object ensembles. In each trial, observers saw a sequential presentation of 4 or 5 images that were either all identical (shown in the gray box), all different (shown in the red box), shared object-ensemble features (shown in the blue box), or contained density changes between successive images (shown in the orange box). To ensure attention to the images, observers were required to count the number of images presented in a trial and to press the appropriate button (i.e., either the "4" or the "5" button). (*b*) Results from Experiment 1. fMRI responses were extracted from independently localized object (LO) and scene-sensitive (PPA) areas of cortex. PPA showed equivalent levels of adaptation in the identical, shared, and density-change conditions when object-ensemble features were repeated, regardless of whether or not absolute density varied. In contrast, LO exhibited an equivalent release from adaptation in the shared and density-change conditions (compared with the identical condition), where changes to local shape information are evident, regardless of changes in absolute density, and showed an even higher release from adaptation when different ensembles were presented in the ensemble-change condition. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus and Mason 1994). (*c*) Additional examples of stimuli used in Experiment 1. PPA = parahippocampal place area; LO = lateral occipital area; *ns* = not significant. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

(7 female, 5 male; mean age = 24.85, range = 19–34 years; 4 of whom had also participated in Experiment 1), and Experiment 3 included 10 paid observers (5 female, 5 male; mean age = 25.6; range = 21–35 years; none participated in Experiment 1 but 4 participated in Experiment 2). All observers were recruited from the Harvard University community, and all were right-handed, reported normal color vision, 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.

One additional female observer was tested in Experiments 1 and 2 but was excluded due to extremely low (<0.1%) averaged PPA activations across all stimulus conditions. This overall low fMRI response made her data unreliable and difficult to interpret.

Stimuli and Procedures
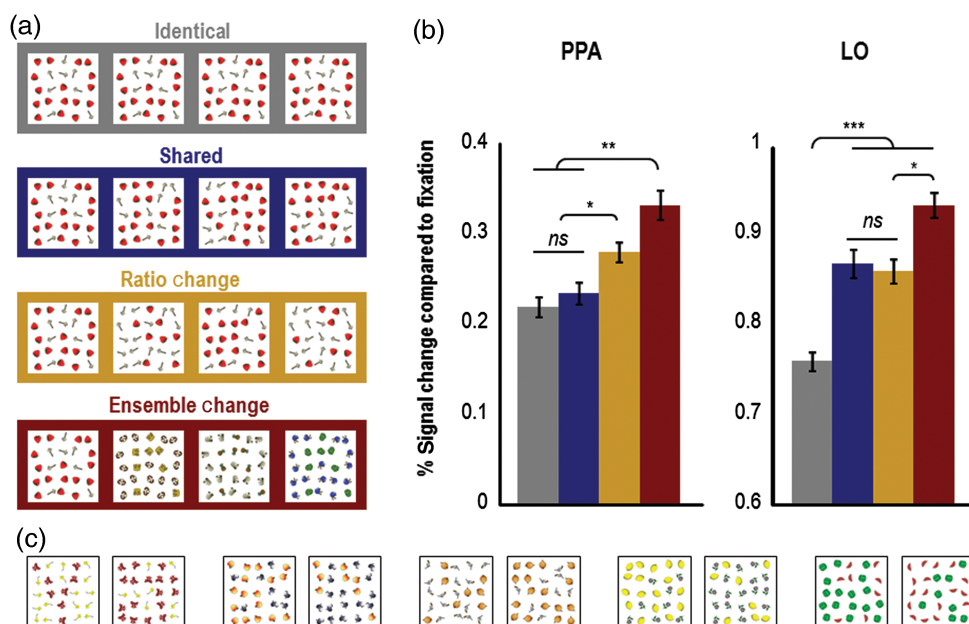
Adaptation Experiments

A fast event-related fMRI-adaptation paradigm was used in all experiments. Each trial contained a sequential presentation of either 4 or 5 images and observers were asked to report the exact number of images shown in a trial by pressing the appropriate response button. The large number of repetitions in this task allowed us to amplify the adaptation effect, and thus increased the power to detect any changes in activation in PPA and LO resulting from density or ratio changes. It should be noted that while this task involved enumeration, it should not affect the interpretation of our results because: (1) observers were not explicitly enumerating the number of objects presented in ensembles in all experiments, (2) the same enumeration task was used in all the stimulus conditions, so any difference in activation across conditions (and experiments) could not be attributed to the enumeration task used here, and (3) ensemble adaptation effects seen in Cant and Xu (2012)

in which different tasks were administered were replicated here with the enumeration task.
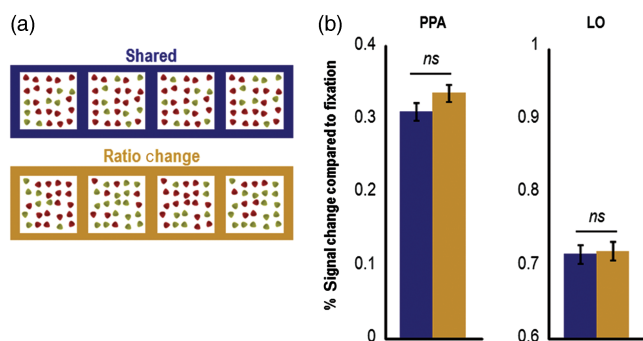
The stimuli used in Experiment 1 were colored photographs of 20 different homogeneous object ensembles, with each containing a repetition of the same type of object (see Fig. 2). All images subtended $12.5° \times 12.5°$ of visual angle (this also applies to all images used in Experiments 2 and 3 and in the object/scene localizer). The photographs were generated by a Nikon D3000 digital SLR camera (Nikon Corporation, Tokyo, Japan) using a desktop photo studio set up. Of the 20 different ensembles, 10 were composed of man-made objects, such as stone beads, screws, and paper clips, and 10 were composed of natural objects, such as nuts, spices, fruits, and vegetables. 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, USA). For each ensemble, we manipulated the spacing between objects and created both a high- and a low-density version of the ensemble, with the high-density version containing twice as many objects as the low-density one (see Fig. 2). Four different photographs of each version of each ensemble were then generated. There were a total of 4 stimulus conditions: (1) identical—repeated presentation of the same image (either dense or sparse); (2) shared—presentation of different images of the same ensemble with no density change (either dense or sparse); (3) density change—alternating presentation of dense and sparse ensemble images containing the same type of object; and (4) ensemble change—presentation of different sparse ensemble images or different dense ensemble images, each containing a different type of object. In a given run, a particular ensemble was shown on average 5.6 times in the ensemble change condition (as images from either 4 or 5 different ensembles were shown in each trial), and on average 1.3 times in the other conditions. Because each ensemble contained 8 different exemplar images (4 sparse exemplars and 4 dense exemplars), the chance of any particular ensemble image being shown once in each condition was <1.

In Experiment 2, instead of photographing real-world objects which were harder to manipulate, we created computer-generated object

Figure 3. Example stimuli and results ($N = 12$) from Experiment 2. (*a*) Example stimuli used in the experiment. The stimuli used in the adaptation runs of Experiment 2 consisted of full-color heterogeneous object-ensemble images, created by combining 2 different types of cartoon objects together in a ratio of roughly 2 to 1 (25 objects were shown in each ensemble, 17 of which were one type of object, and 8 of which were a different type of object). In each trial, observers saw a sequential presentation of 4 or 5 images that were either all identical (shown in the gray box), all different (shown in the red box), shared object-ensemble features (shown in the blue box), or contained ratio changes between successive images (shown in the orange box). The latter 2 conditions were matched in the amount of local object changes (see Materials and methods). To ensure attention to the images, observers were required to count the number of images presented in a trial and to press the appropriate button (i.e., either the "4" or the "5" button). (*b*) Results from Experiment 2. Replicating the results from Experiment 1, PPA again showed equivalent levels of adaptation when object-ensemble features were repeated in the identical and shared conditions, but showed equivalent releases from adaptation when either the ratio or the identity of the objects comprising the ensemble changed. In contrast, LO exhibited equivalent releases from adaptation when local shape information varied in the shared and ratio-change conditions (compared with the identical condition), regardless of whether or not relative density changed, and showed a second-level release from adaptation when the identity of the objects comprising the ensemble changed. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus and Mason 1994). (*c*) Additional examples of stimuli used in Experiment 2. PPA, parahippocampal place area; LO, lateral occipital area; *ns*, not significant. *$P < 0.05$; **$P < 0.01$;*** $P < 0.001$.

Figure 4. Example stimuli and results from Experiment 3 ($N = 10$). (*a*) Example stimuli used in the experiment. In Experiment 3, the change in the ratio of an ensemble was defined solely by a change in the color of ensemble elements, with shape and object identity being held constant. Only the shared and ratio-change conditions were included in the experiment. (*b*) In contrast to the results in Experiment 2, in Experiment 3 the levels of adaptation in the shared and ratio-change conditions in PPA were not significantly different, indicating that the release from adaptation in the ratio-change condition in PPA in Experiment 2 (compared with the shared condition) could not be driven entirely by a color ratio change of the ensembles. The same result was observed in LO. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus and Mason 1994). PPA, parahippocampal place area; LO, lateral occipital area; *ns*, not significant.

ensembles that were composed of 2 different types of objects (see Fig. 3). The 2 types of objects in an ensemble were roughly the same size but were otherwise highly distinguishable from each other, and

were drawn randomly from a pool of 40 different line-drawing objects (24 of which were man-made objects, and of 16 which were natural objects, i.e., fruits, vegetables, insects, and flowers). These 40 objects were a subset of the colored line-drawing objects developed by Rossion and Pourtois (2004). Since objects appeared in different orientations in the ensemble, only objects that could naturally appear in random orientations were included. Each ensemble contained 25 objects, with 8 of type A and 17 of type B objects. There were a total of 4 stimulus conditions: (1) identical—repeated presentation of the same image; (2) shared—presentation of different images of the same ensemble with no object ratio change; (3) ratio change—alternating presentation of 8A/17B and 17A/8B ensembles; and (4) ensemble change—presentation of different ensembles containing different objects with no ratio change. Conditions (1) and (4) provided anchor points for comparison of the amount of adaptation in a given condition, whereas conditions (2) and (3) allowed us to examine the impact of object ratio change on ensemble representation. To match the number of individual object changes between successive displays in conditions (2) and (3), 16 individual objects changed identity in condition (2) and 15 individual objects changed identity in condition (3) (see examples shown in Fig. 3). As such, if condition (3) elicited a higher brain response than condition (2), it could not be attributed to the difference in the amount of local object changes (i.e., shape and orientation changes of local contours).

In Experiment 3, we used the computer-generated ensembles that were employed in Experiment 2 but manipulated ratio change by changing the distribution of color in the ensembles, with shape and object identity held constant (see Fig. 4). All other aspects of this experiment were identical to those reported in Experiment 2, with the exception that only the shared and ratio-change conditions were included in this experiment as these were the 2 conditions critical to testing whether

ensemble representation in PPA is sensitive to processing changes in lower-level visual features (i.e., color).

Each trial lasted 6 s, beginning with a 500 ms fixation, followed by either 4 or 5 sequentially presented images (each consisting of a 200 ms image presentation and a 600 ms blank fixation), and ending with either a 2300 ms (for the 4-image trials) or 1500 ms (for the 5-image trials) blank screen. Observers were asked to report whether each trial contained 4 or 5 images by pressing the appropriate response button (the number of 4-image and 5-image trials were equal across all conditions). Besides the stimulus trials, there was also 6-s 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 Kourtzi and Kanwisher 2001; Xu and Chun 2006). 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 48 s and contained 15 trials for each stimulus condition. In Experiment 1, 4 observers took part in 3 adaptation runs, and the remaining 5 observes took part in 2 adaptation runs. In Experiment 2, all observers took part in 2 adaptation runs, and in Experiment 3, all observers took part in one adaptation run (note that the number of shared and ratio-change trials was equal in Experiments 2 and 3, since eliminating the identical and ensemble-change conditions in Experiment 3 enabled us to present twice as many shared and ratio-change trials within a single run).

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 4 blocks each of scenes, faces, intact objects, and phase-scrambled objects. Each stimulus block was 16-s long and contained 20 different images, each lasting 750 ms and followed by a 50 ms 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 out of every 10 images. Besides the stimulus blocks, there were also 8-s fixation blocks presented at the beginning, middle, and end of each run. Following Kanwisher et al. (1997) and Epstein and Kanwisher (1998), we used 2 unique and balanced run orders. Each run lasted 4 min and 40 s. All observers took part in 3 runs of this localizer. This localizer had already been acquired in a prior experiment in 5 of the observers in Experiment 1, 4 of the observers in Experiment 2, and 4 of the observers in Experiment 3. 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, CA, USA) 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 Corporation, PA, USA) onto a screen mounted behind the observer as he or she lay in the scanner bore. The observer viewed the images through a mirror mounted to the head coil directly above the eyes.

Imaging Parameters

This study was conducted on a 3.0 Tesla Siemens MAGNETOM Tim Trio (Erlangen, Germany) whole-body imaging MRI system at the Center for Brain Science, Harvard University (Cambridge, MA, USA). A Siemens radio-frequency (RF) 32-channel head coil was used to collect blood oxygen level-dependent (BOLD) weighted images (Ogawa et al. 1992). For high-resolution anatomical images, T_1-weighted 3-D

magnetization prepared rapid acquisition gradient echo (MPRAGE) sagittal slices covering the whole brain were collected (inversion time 1100 ms, echo time, or TE, 1.54 ms, repetition time, or TR, 2200 ms, flip angle 7°, 256 × 256 matrix size, 144 slices, 1.0 × 1.0 × 1.0 mm voxel size). For the functional runs, a T_2*-weighted echo-planar gradient echo pulse sequence (72 × 72 matrix size, field of view 21.6 cm) with TR of 1.5 s was used in all adaptation experiments (TE 29 ms, flip angle 90°, 312 volumes). Another pulse sequence with TR of 2.0 s was used for the localizer runs (TE 30 ms, flip angle 85°, 140 volumes). Twenty-four 5-mm-thick (3 × 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, 3D motion correction, linear trend removal, and Talairach space transformation (Talairach and Tournoux 1988).

Data from object/scene localizer was analyzed using a general linear model (GLM), 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 Fig. 1). Following Epstein and Higgins (2007), the RSC and TOS ROIs were defined as regions in retrosplenial 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 et al. (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. Finally, following known anatomical criteria, a retinotopic region in early visual cortex was defined as a region around the Calcarine sulcus whose activations were higher for phase-scrambled objects than for intact objects (e.g., Grill-Spector et al. 1998; James et al. 2003; MacEvoy and Epstein 2011) (see Fig. 6). 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 et al. 2006), we overlaid the ROIs from each observer onto their data from the main adaptation experiment and extracted time courses from that observer. The activation levels for all conditions were then converted to percentage BOLD signal change from baseline, by subtracting the corresponding activation from the fixation trials and then dividing by this value. Peak responses for each condition were obtained by collapsing the time courses for all of the conditions 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., Xu and Chun 2006; Xu 2010). This was done separately for each observer in each ROI, and these resulting peak responses were then averaged across all observers. Finally, the average levels of activation for each condition were subjected to a repeated-measures ANOVA, performed separately on each ROI (SPSS, Chicago, IL, USA). The amount of repetition suppression, or adaptation, for a given condition was evaluated by comparing the average level of activation for that condition against the average level of activation observed in the identical and ensemble-change conditions, using post hoc t-tests.

Behavioral Data Analysis

Behavioral performance measures of reaction time (only adaptation runs) and accuracy (both adaptation and localizer runs) were recorded by Matlab (running the Psychtoolbox) and were analyzed with SPSS (Chicago, IL, USA), by performing one-way repeated-measures ANOVAs to assess differences across the conditions in the adaptation and the localizer runs.

Results

### PPA is not Sensitive to Changes in the Absolute Density of Object Ensembles

Observers were presented with a sequence of real-world object-ensemble photographs that were either (1) all identical (i.e., the same sparse or dense object-ensemble photo was repeated successively), (2) shared object-ensemble and density features (i.e., different photos of the same ensemble, either all sparse or all dense), (3) differed in density but otherwise shared ensemble features (i.e., different images of the same ensemble, with density varying between sparse and dense on successive image presentations, or vice versa), or (4) were completely different, with density repeating (i.e., different ensembles, either all dense or all sparse; see Fig. 2A for examples). Each image was presented for 200 ms with a 600 ms inter-stimulus-interval. Observers were asked to report the number of images presented in each trial (either 4 or 5) by pressing the appropriate response key (due to issues of statistical power, we were not able to analyze neural activation in 4-image and 5-image trials separately, and thus restricted our analysis to activation across both of these responses combined). We used a counter-balanced trial history design and calculated percent-signal change compared with fixation directly from the raw MRI signal (see Kourtzi and Kanwisher 2001; Todd and Marois 2004; Xu and Chun 2006; Xu 2010; Dilks et al. 2011; Todd et al. 2011; Cant and Xu 2012).

We examined responses in independently localized LO and PPA ROIs. Left and right hemisphere ROIs were combined in both regions since no differences in activation were observed between the hemispheres. In PPA, the main effect of stimulus condition was significant ($F_{3,21} = 10.13$, $P < 0.001$). Planned pairwise comparisons revealed a similar level of adaptation when object-ensemble features were repeated (identical vs. shared: $t_7 = 1.11$, $P > 0.50$, one-tailed, and Bonferroni corrected for multiple comparisons; this applies to all subsequent planned pairwise comparisons except where noted), and a release from adaptation when the identity of the objects comprising the ensembles changed (identical vs. ensemble change: $t_7 = 5.38$, $P < 0.005$; shared vs. ensemble change (marginally significant): $t_7 = 3.07$, $P = 0.054$; see Fig. 2B). These results replicate our previous findings (Cant and Xu 2012) and show that PPA is insensitive to image changes so long as ensemble features remain the same. Interestingly, a change in object density did not evoke a release from adaptation and showed the same amount of adaptation as the identical and the shared conditions (identical vs. density change: $t_7 = 0.07$, $P > 0.50$; shared vs. density change: $t_7 < 1.41$, $P > 0.50$; and density change vs. ensemble change $t_7 = 5.77$, $P < 0.005$). These results indicate that absolute density is not a part of the ensemble representation formed in PPA. Since this claim rests on a null result from 8 observers, it is important to evaluate the possibility that we are simply underpowered in our ability to detect a release from adaptation in the density-change condition. To investigate the issue of power, we conducted a power analysis, and found that we would need 45 observers to find a difference between the shared and the density-change condition (with power = 0.90). However, since the effect is actually in the opposite direction (i.e., higher activation in the shared compared with the density-change condition), the inclusion of more participants will only further confirm the lack of ensemble density encoding in PPA. We are thus confident in concluding that we are not underpowered and that PPA is not sensitive to changes in the absolute density of object ensembles.

It is worth noting that the amount of activation for ensembles in PPA is lower than what is typically observed for natural scenes. We could not directly compare scene and ensemble activity in PPA in this study, due to the fact that scene images were shown in blocked design localizer runs whereas ensemble images were shown in event-related runs with different stimulus presentation durations and baselines. We did, however, directly compare PPA response to ensembles and scenes in our previous study (see Fig. 5 in Cant and Xu 2012) and found that scenes elicited roughly double the amount of activation in PPA compared with object ensembles. This is not surprising, given that natural scenes contain many more instances of object ensembles and are far more complex, both in terms of scene content and scene spatial boundary (Park et al. 2011), than the images used in our previous experiment and in all experiments of the present study. Moreover, the response to scenes in PPA is likely based upon both spatial and nonspatial information, and our ensemble images only contain the latter type of visual information. Thus, a difference in overall activation in PPA for object ensembles compared with natural scenes is to be expected, but this should not undermine the significance of object-ensemble representation in this brain region.

In LO, the main effect of condition was also significant ($F_{3,21} = 19.57$, $P < 0.001$), but planned pairwise comparisons revealed a different pattern of results compared with that of PPA. Specifically, there was a release from adaptation for object ensembles in LO when local shape or contours changed between successively presented images (identical vs. shared: $t_7 = 4.90$, $P < 0.01$; identical vs. density-change (marginally significant): $t_7 = 3.00$, $P = 0.057$; identical vs. ensemble change: $t_7 = 7.27$, $P < 0.001$; see Fig. 2B), again replicating our previous results (Cant and Xu 2012). With the present paradigm in which either 4 or 5 images were presented (as opposed to the presentation of only 2 or 3 images in our previous study), we also observed an additional release from adaptation when the objects comprising the ensemble changed identity, as greater response amplitude was seen in the ensemble change than in the shared and the density-change conditions (shared vs. ensemble-change: $t_7 = 4.29$, $P < 0.05$; density-change vs. ensemble-change: $t_7 = 4.92$, $P < 0.01$; the shared and density-change conditions did not differ: $t_7 = 0.23$, $P > 0.5$). This significant repetition attenuation in the shared condition (compared with the ensemble-change condition), which was not observed in our previous study (Cant and Xu 2012), may have arisen from the combination of a number of factors. First, presenting more images per trial in our present paradigm led to greater sensitivity in detecting adaptation effects (just as blocked fMRI designs, which present more images per block, have better detection power than event-related designs, which present fewer images per event; see Huettel et al. 2009). Second, in the shared condition, the arrangement/orientation of contours varied but the contours themselves repeated, whereas in the ensemble-change condition both the arrangement and the contours varied. This additional shape variation, in combination with the increased number of image repetitions, may account for the release from adaptation in the ensemble-change condition compared with the shared (and density-change) conditions. This may also explain why we failed to observe this effect in our previous study (Cant and Xu 2012), which had a fewer number of image repetitions per trial. Finally, it is also possible that differences in
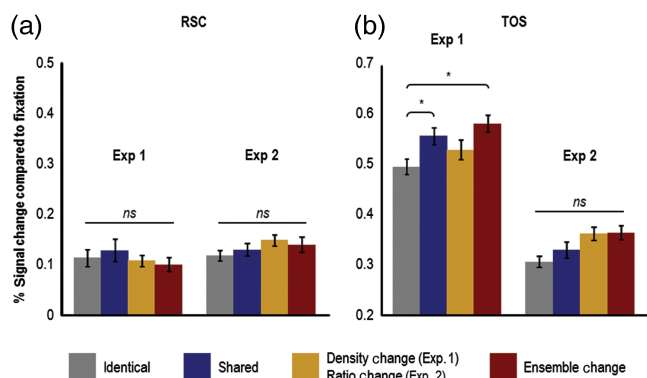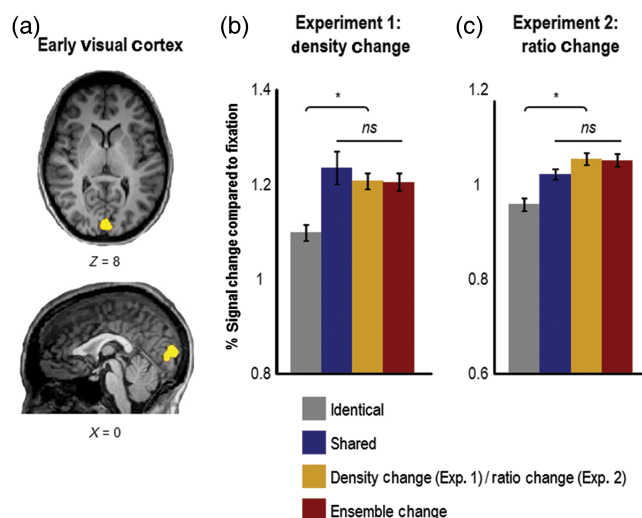
Figure 5. Adaptation results in RSC and TOS in Experiments 1 and 2. RSC and TOS were defined with the same contrast used to define PPA (i.e., contrasting the activation for scenes with the activation for faces and objects). (*a*) Results in RSC from Experiments 1 and 2. No differential adaptation effects were observed in RSC in either experiment. (*b*) Results in TOS from Experiments 1 and 2. In Experiment 1, only the shared and ensemble-change conditions showed a significant release from adaptation (compared with the identical condition), and in Experiment 2 no differential adaptation effects were observed in TOS. Taken together, these results are decidedly different than those observed in PPA, and suggest that sensitivity to processing object ensembles is not a general phenomenon seen throughout the entire human scene-processing network. Instead, these results suggest a functional dissociation between scene-processing regions, with PPA being sensitive to both spatial (e.g., spatial expanse; see Kravitz, Peng, and Baker 2011) and nonspatial (i.e., object ensembles and textures; see Cant and Xu 2012) aspects of visual scenes, whereas RSC and TOS may only participate in spatial aspects of visual scene processing. Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus and Mason 1994). RSC, retrosplenial complex; TOS, transverse occipital sulcus; *ns*, not significant. *$P < 0.05$.

Figure 6. Results from early visual cortex in Experiments 1 ($N = 8$) and 2 ($N = 12$). (*a*) Example of early visual cortex ROI (Talairach *x*, *y*, *z* coordinates: 0, −86, 8) from one observer (defined using the contrast of phase-scrambled objects > intact objects from the object/scene localizer and the calcarine sulcus as an anatomical landmark). (*b*) In both Experiments 1 and 2, there was an equivalent release from adaptation (compared with the identical condition) in early visual cortex whenever local shape information changed across successive images, which occurred in the shared, density-change (Experiment 1), ratio-change (Experiment 2), and ensemble-change conditions. Importantly, in both Experiments, the pattern of adaptation observed in early visual cortex was significantly different from the patterns observed in both LO and PPA, which suggests that early visual cortex extracts different kinds of visual information from object ensembles compared with LO and PPA. Specifically, low-level visual information (e.g., spatial frequency, oriented line segments) is likely extracted in early visual cortex, whereas higher-level visual information is likely extracted in LO (e.g., closed contours) and PPA (e.g., ensemble statistics). Error bars represent within-subject standard errors (i.e., with the between-subject variation removed; see Loftus and Mason 1994). *ns*, not significant. *$P < 0.05$.

the particular tasks used led to differences in the patterns of adaptation observed in LO here compared with our previous study, but we think this is unlikely because we found the same pattern of activation in LO in our previous study despite using 2 different tasks. In any regard, differences in the patterns of adaptation between PPA and LO in the present experiment reached significance (i.e., the interaction between brain region and stimulus condition, $F_{3, 21} = 4.78$, $P < 0.05$), demonstrating that these 2 brain areas extract different types of information from the same visual input.

To assess the reliability of these results, we also conducted a standard GLM analysis to derive beta weights for each condition in both PPA and LO. The patterns of adaptation obtained from these beta weight measures, and the results of the statistical tests applied to them, were identical to those obtained from the percent-signal change analysis in both PPA and LO. This confirms the reliability and validity of the counter-balanced trial history design and the use of the percent-signal change analysis. As such, in all subsequent analyses we report results using this analysis.

These data show that the density-change condition behaved just like the shared ensemble condition, with both showing adaptation in PPA but a release from adaptation in LO. In other words, PPA is insensitive to changes in the absolute density (i.e., spacing) of the objects comprising an ensemble, suggesting that absolute density is not part of the neural ensemble representation in anterior-medial ventral visual cortex. It is likely that density is treated as an accidental feature, rather than as an intrinsic and diagnostic feature of an ensemble in this brain region, as a change in density does not alter the mean features of the objects comprising an ensemble (such as mean size

and mean texture). This is consistent with our earlier findings showing that PPA is not modulated by an overall size change of an ensemble image (Cant and Xu 2012).

PPA is Sensitive to Changes in the Relative Density of Object Ensembles

Instead of varying absolute density by varying the spacing between the objects within a homogeneous ensemble as we did in Experiment 1, in Experiment 2, we varied relative density by varying the ratio, or proportion, of 2 types of objects comprising a heterogeneous ensemble to examine whether or not such a feature is part of the neural ensemble representation in anterior-medial ventral visual cortex (see Fig. 3*A*). Critically, we matched the number of local shape changes between the shared and the ratio-change conditions (see Materials and methods), ensuring that any activation differences obtained between these 2 conditions could not be attributed to a difference in the amount of local object shape changes. We used the same 4 conditions as in Experiment 1, but replaced the density-change condition with a ratio-change condition.

In PPA, the main effect of condition was significant ($F_{3,33} = 12.278$, $P < 0.001$). Planned pairwise comparisons revealed similar levels of fMRI-adaptation when object-ensemble features repeated (i.e., identical vs. shared: $t_{11} = 0.71$, $P = 0.50$; see Fig. 3*B*), and a significant release from adaptation when the identity of the objects comprising the ensembles changed (identical vs. ensemble-change: $t_{11} = 5.60$, $P < 0.001$; shared vs.

ensemble-change: $t_{11} = 4.04$, $P < 0.01$). These results replicate those from Experiment 1 and our previous study (Cant and Xu 2012) and show that PPA is insensitive to image changes so long as ensemble features within the images remain the same. Importantly for the purpose of this experiment, we observed a significant release from adaptation when the relative density of an object ensemble varied (identical vs. ratio-change: $t_{11} = 4.07$, $P < 0.01$; and shared vs. ratio-change: $t_{11} = 3.46$, $P < 0.05$). In fact, changing the relative density of an ensemble produced a release from adaptation that was comparable with that observed when changing the identity of the objects comprising an ensemble (ratio-change vs. ensemble-change: $t_{11} = 2.00$, $P > 0.20$). Since we matched the number of local shape and orientation changes across the shared and the ratio-change conditions, it is unlikely that the higher activation in the ratio-change condition is attributed to differences in local shape and orientation across these conditions.

One may note that because colorful object images were used, an object ratio change in Experiment 2 was always accompanied by a color ratio change. Thus, it is possible that any object ratio-related effect in PPA could be entirely attributed to a color ratio-related effect. We investigated this possibility in Experiment 3, where the change in the ratio of an ensemble was defined solely by a change in the color of ensemble elements (with shape and object identity being held constant; see Fig. 4A). In contrast to the results in Experiment 2, in Experiment 3 we did not observe a difference in activation between the shared and ratio-change conditions in PPA ($t_9 = 1.08$, $P > 0.30$ see Fig. 4B). Taken together, the results of Experiment 2 indicate that ratio, or the relative density of the 2 types of objects comprising an ensemble, is part of the neural ensemble representation formed in PPA, and the results of Experiment 3 rule out the possibility that this representation is solely based on differences in the color of the elements constituting an ensemble.

In Experiment 2, the main effect of condition in LO was significant ($F_{3,33} = 21.08$, $P < 0.001$). Planned pairwise comparisons showed a release from adaptation whenever local shape or contour varied in an ensemble, which occurred in the shared, ratio-change, and ensemble-change conditions (identical vs. shared: $t_{11} = 4.74$, $P < 0.005$; identical vs. ratio-change: $t_{11} = 5.94$, $P < 0.001$; and identical vs. ensemble-change: $t_{11} = 9.21$, $P < 0.001$; see Fig. 3B). There was no difference between the shared and the ratio-change conditions ($t_{11} = 0.32$, $P = 0.50$), confirming that our manipulation was successful in matching the total number of local shape changes between these 2 conditions. As in Experiment 1, we also observed an additional release from adaptation when the objects comprising the ensemble changed identity, as greater response amplitude was seen in the ensemble-change than in the shared (which approached significance; $t_{11} = 2.75$, $P = 0.063$) and the ratio-change conditions ($t_{11} = 3.08$, $P < 0.05$). Again, this is likely explained by the number of image repetitions used in this study (compared with Cant and Xu 2012, which used fewer repetitions and did not observe this effect) and the additional shape variation in the ensemble-change condition compared with the shared (and ratio-change) conditions. Differences in the patterns of adaptation between PPA and LO were significant (interaction between brain region and all stimulus conditions: $F_{3,33} = 4.94$, $P < 0.01$; and interaction between brain region and just the shared and the ratio-change conditions: $F_{1,11} = 5.14$, $P < 0.05$), again showing that these 2 regions extract different types of information from the same visual input. Finally, in Experiment 3, we observed similar levels of adaptation in the shared and ratio-change conditions in LO ($t_9 = 0.17$, $P > 0.86$; see Fig. 4B), a pattern that we also observed in Experiment 2. Differences in the patterns of adaptation between PPA and LO were not significant in Experiment 3 ($F_{1,9} = 0.88$, $P = 0.37$), suggesting that ratio changes defined solely on the basis of a change in the color of ensemble elements are not able to differentiate the processing carried out by these 2 regions.

Taken together, the results from Experiments 1 to 3 demonstrate that PPA is not sensitive to processing changes in the absolute density of homogeneous object ensembles, but is sensitive to processing changes in the relative density, or ratio, of heterogeneous ensembles. To provide a more direct measure of this functional difference, we used the data from the 8 observers who were unique to Experiments 1 and 2 and examined the patterns of adaptation in PPA across both experiments. Importantly, the patterns of adaptation in PPA across the 2 experiments were significantly different (interaction between Experiment and the shared and ratio-change conditions: $F_{1,14} = 7.98$, $P < 0.05$), providing strong evidence that PPA is sensitive to processing relative, rather than absolute, changes in density. Moreover, the patterns of adaptation in LO were not significantly different across the 2 experiments ($F_{1,14} = 0.27$, $P > 0.61$), suggesting that, compared with LO, PPA was more sensitive to the difference between the shared and ratio-change conditions in Experiment 2. Since this difference cannot be attributed solely to variations in visual features such as color ratio, this suggests that PPA is sensitive to processing higher-level changes in the ratio of the elements that constitute an object ensemble. Further investigation is needed to determine whether this higher-level ratio representation reflects a summary ratio representation of multiple low-level features or whether it instead reflects a high-level object identity ratio representation independent of low-level features.

The Encoding of Density and Ratio Outside PPA

There exists an extensive literature demonstrating that regions in posterior parietal cortex and prefrontal cortex participate in numerical perception (for review, see Nieder 2005; see also Jacob et al. 2012). Although results from Experiment 2 showed that PPA is sensitive to object ratio changes in an ensemble, this should not be taken as evidence that PPA is a number-processing region, as it did not show sensitivity to changes in absolute density in Experiment 1 in which the number of items either doubled or halved between successive displays. Moreover, a recent study reported that number representation may be distinct from texture processing (Stoianov and Zorzi 2012; see also Ross and Burr 2012). Since PPA responds similarly to object ensembles and textures (Cant and Xu 2012), it is likely that both texture and object ensembles are processed distinctly from number.

Nevertheless, given that both density and ratio manipulations are related to numerical processing, here we conducted whole-brain random effect analyses on the group adaptation data from Experiments 1 and 2 to investigate whether or not regions outside of occipito-temporal cortex, especially those in the parietal and prefrontal regions, participate in the processing of density and ratio. We should note that these analyses serve primarily exploratory purposes, as adaptation data

typically produce small effects and group-averaged data do not capture common activations if anatomical variations across observers are large.

In these analyses, we looked for regions that showed a release from adaptation in the density-change (Experiment 1) and the ratio-change (Experiment 2) conditions, compared with only the shared condition, to ensure that any release from adaptation that we observed in a region was driven by a change in density or ratio, and not simply by the basic adaptation effect (i.e., identical < different). In Experiment 1, with 8 observers, at $P < 0.01$, uncorrected, no regions showed sensitivity to density changes; at $P < 0.05$, uncorrected, the only significantly active region resided in the left anterior temporal cortex (x, y, z Talairach coordinates: −30, 2, −35, cluster size = 21 voxels, with each voxel being $1 \times 1 \times 1$ mm). In Experiment 2, with 12 observers, at $P < 0.01$, uncorrected, 4 regions showed sensitivity to ratio changes: one in the right posterior-medial parietal cortex (9, −61, 31, 263 voxels), 2 separate regions in right inferior frontal cortex (first region: 10, 35, −6, 64 voxels; second region: 18, 33, −11, 59 voxels), and one in left anterior temporal cortex (−39, 20, −29, 63 voxels).

The frontal and parietal regions found here are consistent with previous studies reporting that both of these brain regions are involved in absolute number and ratio representations (Dehaene et al. 1999; Ischebeck, Schocke, and Delazer, 2009; Piazza et al. 2007; Jacob and Nieder, 2009a,b; see Jacob et al. 2012 for a recent review), although our frontal activations are located a little more inferior than what is typically reported. Since numerical processing is not typically observed in anterior temporal cortex, it is unclear whether the anterior temporal activations observed in this study represent numerical processing of density and ratio, or instead reflect processing unrelated to number, such as the extraction of statistical information from ensembles, or the processing of semantic and/or visual features of the objects within the ensembles. Further studies are needed to fully understand this anterior temporal activation, and how anterior-medial ventral visual cortex and parietal and frontal regions coordinate to represent ratio in the human brain.

To understand whether ensemble ratio representation is common across the scene-processing regions, in addition to PPA, we also examined the patterns of adaptation in RSC and TOS, 2 regions that, along with PPA, comprise the human scene-processing network (see Epstein et al. 2005; Epstein 2008). The main results within RSC and TOS in Experiments 1 and 2 are shown in Figure 5. In Experiment 1, the main effect of condition in RSC was not significant ($F_{3, 24} = 0.08$, $P = 0.97$), and none of the 4 conditions differed from each other (all t's < 1.50). Similarly, in Experiment 2 the main effect of condition in RSC was not significant ($F_{3, 36} = 0.18$, $P = 0.907$), and none of the 4 conditions differed from each other (all t's < 1.50). Importantly, differences in the patterns of adaption between PPA and RSC were significant in both experiments (Experiment 1, region-by-condition interaction: $F_{3, 24} = 8.57$, $P = 0.001$; Experiment 2, region-by-condition interaction: $F_{3, 36} = 2.89$, $P = 0.049$), strongly suggesting that, unlike PPA, RSC is not sensitive to processing object ensembles.

In Experiment 1, the main effect of condition in TOS was significant ($F_{3, 24} = 5.09$, $P = 0.007$). Planned pairwise comparisons revealed that there was a significant release from adaptation, compared with the identical condition, in both the shared ($t_8 = 3.61$, $P = 0.021$) and ensemble-change ($t_8 = 3.22$, $P = 0.039$)

conditions. All other comparisons failed to reach significance (all other t's < 2.64). In contrast, in Experiment 2, the main effect of condition in TOS was not significant ($F_{3, 33} = 1.26$, $P = 0.303$), and none of the 4 conditions differed from each other (all t's < 2.18). Differences in the patterns of adaptation between PPA and TOS did not reach significance in Experiment 1 (region-by-condition interaction: $F_{3, 24} = 1.70$, $P = 0.194$) but approached significance in Experiment 2 (region-by-condition interaction: $F_{3, 33} = 2.41$, $P = 0.085$). These nonsignificant region-by-condition interactions, however, do not provide compelling evidence for the notion that PPA and TOS process object ensembles in similar manners. First, the activation in the shared condition (where the arrangement of ensemble elements varies but importantly, the high-level ensemble identity remains constant) in TOS in Experiment 1 is quite different from that observed in PPA in the same experiment: in TOS, the contrast of shared versus identical was significant but the contrast of shared versus ensemble-change was not; in PPA these significant and null results were switched. That is, shared versus identical was not significant but shared versus ensemble-change was. A region that is sensitive to processing high-order ensemble information should demonstrate equivalent levels of adaptation when ensemble information repeats (i.e., in the identical and shared conditions), and a release from adaptation when ensemble information varies (i.e., in the ensemble-change condition, compared with the shared condition). TOS did not show either of these effects in Experiment 1. Second, none of the 4 adaptation conditions differed from each other in TOS in Experiment 2, showing that, unlike PPA, TOS is not sensitive to processing object ensembles. [We did not examine patterns of adaptation in RSC and TOS in Experiment 3 because the stimuli used in Experiments 2 and 3 were similar (i.e., ensemble ratio stimuli), and results from Experiment 2 revealed that there was no differential sensitivity to processing ensemble ratio stimuli in RSC and TOS (i.e., no basic adaptation effect of identical < ensemble-change)].

Taken together, we have provided evidence that the contribution of object-ensemble information to scene representation is likely restricted to processing in PPA, and is not a general phenomenon seen throughout the entire scene-processing network. Together with our previous findings (Cant and Xu 2012), these results suggest the existence of a functional dissociation in the human scene-processing network. Specifically, PPA may be involved in both spatial (e.g., spatial expanse; Kravitz et al. 2011) and nonspatial aspects of visual processing (e.g., object-ensemble and texture processing; Cant and Xu 2012), whereas RSC and TOS may only participate in spatial aspects of visual processing. This suggestion is certainly consistent with the separate (but complementary) functional roles posited for these 3 regions in the representation of scenes (see Epstein 2008, for review).

Finally, it is likely that retinotopic regions of early visual cortex participate in the processing of ensemble density and ratio, since changes in absolute and relative density naturally entail changes in low-level visual information (e.g., spatial frequency, oriented line segments). Although we did not conduct retinotopic mapping in this study, we were able to localize early visual cortex by using known anatomical makers and our object/scene localizer. Specifically, we defined early visual cortex as a region encompassing the calcarine sulcus whose activation was higher for phase-scrambled than intact objects (e.g., Grill-Spector et al. 1998; James et al. 2003) (see Fig. 6A). We localized this region independently in all observers, and

Table 1

Percent correct accuracies and response latencies (in ms) of correct trials for the localizer and adaptation runs in all 3 experiments

Object/scene localizer accuracy

	Objects	Scenes	Faces	Scrambled objects
Experiment 1	94.79 ± 2.58	96.88 ± 2.05	96.35 ± 2.00	97.40 ± 1.35
Experiment 2	94.27 ± 1.80	92.88 ± 2.47	93.06 ± 1.80	93.40 ± 1.88
Experiment 3	93.75 ± 1.55	95.42 ± 1.70	95.00 ± 1.74	95.83 ± 1.96

Adaptation runs accuracy

	Same	Shared	Density change (Exp. 1) Ratio change (Exp. 2/3)	Ensemble change
Experiment 1	96.40 ± 1.01	95.65 ± 1.24	96.69 ± 1.33	95.92 ± 1.31
Experiment 2	94.08 ± 2.41	95.50 ± 2.32	96.33 ± 1.67	93.29 ± 2.18
Experiment 3		95.33 ± 1.87	96.33 ± 1.68	

Adaptation runs response latency

	Same	Shared	Density change (Exp. 1) Ratio change (Exp. 2/3)	Ensemble change
Experiment 1	1100 ± 46	1121 ± 46	1108 ± 48	1122 ± 49
Experiment 2	1159 ± 45	1149 ± 36	1146 ± 42	1161 ± 40
Experiment 3		1175 ± 27	1150 ± 25	

Note: Response latencies were not collected for the localizer runs. All values represent means with standard errors. There were no significant main effects of condition in any of the object/scene localizer (Experiment 1: $F_{3,21} = 0.42$, $P > 0.74$; Experiment 2: $F_{3,33} = 0.13$, $P > 0.94$; Experiment 3: $F_{3,27} = 0.45$, $P > 0.72$) or adaptation runs (accuracy in Experiment 1: $F_{3,21} = 0.26$, $P > 0.85$; accuracy in Experiment 2: $F_{3,33} = 1.44$, $P > 0.24$; response latency in Experiment 1: $F_{3,21} = 1.40$, $P > 0.26$; response latency in Experiment 2: $F_{3,33} = 0.85$, $P > 0.47$; there were no main effects of accuracy or response latency in the adaptation runs of Experiment 3 because only 2 conditions were used in each run). No pairwise comparison between conditions in any accuracy or response-latency analysis reached significance (all 2-tailed and Bonferroni corrected).

examined the patterns of activation across the 4 stimulus conditions in Experiments 1 and 2.

In Experiment 1, the main effect of stimulus condition was significant ($F_{3,21} = 5.35$, $P < 0.01$). Planned pairwise comparisons revealed a release from adaptation in early visual cortex, compared with the identical condition, whenever the local contours changed across images (identical vs. shared: $t_7 = 3.04$, marginally significant at $P = 0.058$; identical vs. density change: $t_7 = 5.45$, $P < 0.005$; identical vs. ensemble change: $t_7 = 3.93$, $P < 0.05$; see Fig. 6B). The activation observed in these latter 3 conditions did not differ ($ts < 0.63$, $Ps > 0.50$). Importantly, the adaptation pattern observed here is significantly different from that observed in PPA (interaction between brain region and all stimulus conditions: $F_{3,21} = 10.01$, $P < 0.001$) and LO ($F_{3,21} = 5.97$, $P < 0.005$). These differences indicate that the patterns of activation observed in PPA and LO are not direct reflections of low-level visual information encoded in early visual areas.

In Experiment 2, the main effect of stimulus condition was significant ($F_{3,33} = 6.11$, $P < 0.005$), and planned pairwise comparisons revealed a pattern of activation in early visual cortex similar to that seen in Experiment 1. Specifically, compared with the identical condition, there was a release from adaptation in all the other stimulus conditions (identical vs. shared: $t_{11} = 3.72$, $P < 0.05$; identical vs. ratio change: $t_{11} = 2.78$, marginally significant at $P = 0.054$; and identical vs. ensemble change: $t_{11} = 3.46$, $P < 0.05$). As in Experiment 1, the activation levels in these 3 latter conditions did not differ ($ts < 1.18$, $Ps > 0.50$). Finally, the patterns of adaptation in early visual cortex differed from those in LO (interaction between brain region and all stimulus conditions: $F_{3,33} = 3.33$, $P < 0.05$) and PPA (marginally significant, $F_{3,33} = 2.39$, $P = 0.087$). These results, together with those from Experiment 1, indicate that early visual cortex extracts different kinds of visual information from object ensembles compared with LO and PPA. This difference likely reflects the extraction of low-level visual information in early visual cortex (e.g., spatial frequency, oriented line segments) and more high-level information in LO (e.g., closed contours) and PPA (ensemble statistics).

Behavioral Results

In the adaptation runs of all Experiments, observers were asked to count the number of images presented in each trial (either 4 or 5 images), and in the object/scene localizer runs, observers were asked to detect an occasional spatial jitter of the images. All behavioral results are presented in Table 1. No behavioral comparison reached significance. This indicates that behavioral response patterns did not match the fMRI response patterns, making it unlikely that behavioral responses directly contributed to the observed fMRI results. This is consistent with the findings by Xu et al. (2007), who also showed that fMRI-adaptation responses in PPA are dissociable from behavioral responses.

Discussion

Object-ensemble perception is an important aspect of visual processing and involves the extraction of summary statistical information from large sets of objects, thus circumventing the capacity limitation inherent in object-based visual representation (e.g., Luck and Vogel 1997; Pylyshyn and Storm, 1998; see also Cowan 2001). We have previously demonstrated that anterior-medial ventral visual cortex, along the collateral sulcus and parahippocampal gyrus and overlapping with the scene-sensitive PPA, is involved in processing object ensembles (Cant and Xu 2012). The present study explored the nature of object-ensemble representation in this brain region by examining the encoding of the absolute and relative densities of object ensembles. Here, absolute density refers to the amount of spacing between the objects comprising an ensemble and relative density refers to the ratio, or proportion, of 2 different types of objects comprising an ensemble. Using the fMRI-adaptation method and an independent ROI-based analysis to define PPA, we found that while this brain region was not sensitive to changes in absolute density, it did respond to changes in relative density.

In contrast, the object-selective region LO responded whenever local shape contours changed, even when ensemble

features repeated. This replicates our previous findings (Cant and Xu 2012) and demonstrates that, in addition to processing the shape of single objects (e.g., Malach et al. 1995; Grill-Spector, et al. 1998; Kourtzi and Kanwisher 2001), LO is also involved in processing shape features from ensembles of multiple objects. Finally, patterns of adaptation in early visual cortex were different from patterns observed in both PPA and LO, suggesting that the former region extracts low-level visual information from object ensembles (e.g., spatial frequency, oriented line segments), whereas the latter 2 regions extract more high-level visual information (i.e., ensemble statistics in PPA and closed contours in LO).

The Nature of Object-Ensemble Representation in Anterior-medial Ventral Visual Cortex

A change in the absolute density of an object ensemble is often associated with a change in both the number of objects in the ensemble and the spatial frequency of the image. The finding that PPA is not sensitive to changes in absolute density (and is thus insensitive to changes in number and spatial frequency) shows that this brain region is tuned to process higher level, rather than lower level, aspects of object ensembles. This is consistent with our previous finding showing that PPA is not sensitive to a change in the size of an ensemble image (Cant and Xu 2012). Moreover, when we held absolute density, number and spatial frequency constant, but varied the ratio of the 2 types of objects comprising an ensemble, we observed a release from adaption in PPA, demonstrating a sensitivity to changes in relative density and lending further support to the notion that anterior-medial ventral visual cortex processes higher-level features of object ensembles.

What is the nature of the ratio representation in anterior-medial ventral visual cortex? A number of behavioral studies have shown that observers can quickly extract useful ensemble statistics, such as mean size, speed and orientation, from a display without encoding specific details of any single object within the display (e.g., Williams and Sekuler 1984; Watamaniuk and Duchon 1992; Ariely 2001; Parkes, et al. 2001; Chong and Treisman 2003; 2005a, b; Alvarez and Oliva 2008, 2009; also see Alvarez 2011). Note that although mean features are informative in describing a homogenous ensemble, they are not particularly useful in describing a heterogeneous ensemble (e.g., the average features from apples and oranges are not informative). Instead, the exact compositions and the amount of variations of the different visual features of objects can provide diagnostic characteristics of a heterogeneous ensemble. It is possible that anterior-medial ventral visual cortex represents precisely this type of ensemble statistical information (i.e., the statistical distribution of visual features in an ensemble), and that a ratio change simply alters such representations. Another possibility is that 'ratio' could be a high-level feature reflecting the composition of the individual object identities within an ensemble, and thus ratio representation is independent of the specific ensemble feature from which it was computed. In this sense, 'ratio' is associated more with numerical processing than it is with visual feature processing. Although we have evidence arguing against a purely numerical account of ratio representation in anterior-medial ventral visual cortex (see Results), further work is need to fully distinguish between these 2 possibilities.

Beyond mean ensemble features, summary statistics can also include features like the marginal distribution of luminance, luminance autocorrelation, correlations across location, orientation, and scale, and phase correlation. These latter summary features have been used successfully in texture synthesis algorithms (see Portilla and Simoncelli 2000) and in explaining how the visual system may represent the external environment, particularly outside of the fovea where visual resolution is degraded (Balas, Nakano, and Rosenholtz 2009; Rosenholtz 2011). Interestingly, PPA has been shown to be involved in texture processing (Steeves et al. 2004; Cant and Goodale 2007, 2011; Cant et al. 2009; Cant and Xu 2012) and to exhibit a peripheral-field bias (Levy et al. 2001). This suggests that anterior-medial ventral visual cortex may be involved in the processing of a variety of ensemble features beyond those investigated here and in our previous study (Cant and Xu 2012). We are agnostic regarding whether anterior-medial ventral visual cortex is the general "summary statistics" area of the human brain, or whether only a subset of summary statistics are processed and represented there. We believe that there are many exciting research opportunities for future studies to fully explore the various factors that may contribute to object-ensemble representation in the entire human brain.

The Link Between Object-Ensemble and Scene Processing

The processing of scenes in PPA has been shown to depend on 3D spatial layout (e.g., Epstein and Kanwisher 1998). Since our object-ensemble images do not depict 3D spatial layout, and in general do not invoke scene imagery, why does the processing of object ensembles and scenes involve PPA? Besides encoding 3D spatial layout, an important aspect of scene processing involves the extraction of gist (the overall meaning and perceptual structure of a scene), which can be obtained from zones of repeating information, without processing the individual objects within the scene in great detail (e.g., Oliva and Schyns 2000; Oliva and Torralba 2001). This type of processing bears a striking similarity to the processing of object ensembles and textures. Thus, anterior-medial ventral visual cortex may play a greater role in extracting higher-order statistical information from a variety of visual displays. This may explain why the processing of object ensembles, surface textures, and scenes all activate this common region. We should note, however, that the processing of summary statistics is restricted to PPA, as 2 additional regions in the human scene-processing network, RSC and TOS, are not sensitive to processing object ensembles and textures (see Fig. 5, and Cant and Xu 2012). This suggests that there may be a functional dissociation in the human scene-processing network, with PPA involved in processing both spatial (e.g., spatial expanse; Kravitz et al. 2011) and nonspatial aspects of visual processing (e.g., object-ensemble and texture processing; Cant and Xu 2012), whereas RSC and TOS are only involved in spatial aspects of visual processing. This suggestion is certainly consistent with the separate (but complementary) functional roles posited for these 3 regions in the representation of scenes (see Epstein 2008, for review).

Oliva and Torralba (2001) have put forward a model proposing that spatial-envelope representation is important in scene processing and that scenes can be represented as a collection of a number of diagnostic global features that describe both spatial boundary and scene content, such as openness, expansion, mean depth, roughness, and naturalness of content

(Oliva and Torralba 2001, 2002, 2006, 2007; Torralba and Oliva 2002, 2003; Greene and Oliva 2009, 2010; Park et al. 2011; for review, see Oliva et al. 2011). Since spatial frequency is important in the processing of a number of these global scene properties, how do we reconcile this with the lack of sensitivity to changes in absolute density/spatial frequency in PPA in Experiment 1? This latter finding, however, does not imply that spatial frequency is not an important aspect of scene representation. It merely shows that spatial frequency may not be encoded within the scene-sensitive PPA. Spatial frequency related to scene representation may instead be encoded in lower-level visual areas, such as V1/V2. Alternatively, PPA may indeed be sensitive to processing spatial frequency, but only in situations where this information can be used to form a representation of the spatial structure of the environment.

Park et al. (2015) have demonstrated that PPA encodes the level of clutter in a visual scene (i.e., different levels of clutter ranging from an empty room to a full room). At first blush, this seems inconsistent with the results of Experiment 1, which showed that PPA was not sensitive to changes in the absolute density of ensemble elements. In Park et al. clutter was manipulated by changing the total number of different types of objects in a room. Such a manipulation necessarily involves changing the composition (or ratio) of the different types of objects in a room, akin to the ratio manipulation we used in Experiment 2 where we found that PPA encodes the relative density of 2 types of objects comprising an ensemble. Thus, our results are in agreement with those of Park et al. and show that PPA is sensitive to the exact makeup of an object ensemble.

MacEvoy and Epstein (2011) have suggested that LO and PPA play complementary roles in scene recognition, with LO mediating object-based scene identification and PPA mediating scene identification via processing of more global scene properties (such as visual summary statistical information and spatial geometry). Specifically, MacEvoy and Epstein demonstrated that, in LO, single object representations are combined linearly to form the representation of scenes that contain multiple objects, whereas in PPA, single object representations would interact nonlinearly to give rise to a uniquely scene-specific representation, which could contain global scene properties such as summary statistics of visual features and spatial layout of the objects in a scene. The results of MacEvoy and Epstein are thus consistent with existing notions of object-specific representation in LO and scene-specific representation in PPA. Our present findings from LO and PPA with object ensembles are certainly consistent with this view as we found LO to be sensitive to individual object feature changes even when ensemble features remain the same, whereas PPA was sensitive to ensemble feature changes and not to individual object feature changes.

Two Independent and Complementary Visual Processing Mechanisms in the Brain

We previously posited that visual objects may be processed by 2 independent, but complementary, neural processing mechanisms (Cant and Xu 2012). One mechanism, involving higher visual object processing areas (such as LO) and regions in the parietal lobe (see Xu and Chun 2009), participates in the individuation and the encoding of the detailed features of both single objects and objects within an ensemble. The other mechanism, involving anterior-medial regions of the ventral processing stream (including the collateral sulcus and

parahippocampal gyrus and overlapping with the scene-sensitive PPA), extracts summary statistics from object ensembles without encoding the details of the individual objects comprising the ensemble. In this way, object-ensemble representation complements and guides object-specific processing as it allows the visual system to overcome the capacity limitation inherent in object-based attention (e.g., Luck and Vogel 1997; Pylyshyn and Storm, 1998; Xu 2002; Alvarez and Cavanagh 2004).

The present findings from PPA and LO are consistent with this model, and further demonstrate that the processing of statistical information in anterior-medial ventral visual regions is not based solely on low-level visual information, such as absolute density, number, spatial frequency, or color, but is instead based on high-level visual information, such as the relative density or ratio of different objects comprising an ensemble. Future experiments will need to investigate whether such higher-level ratio representation reflects a summary ratio representation of multiple low-level features or whether it instead reflects a high-level object identity ratio representation independent of low-level features.

Uncovering the neural underpinnings of object-ensemble representation is essential to understanding human vision since ensembles are ubiquitous in our everyday visual world, and their processing complements object-specific representations in the brain. Specifically, since visual processing has a limited capacity, collapsing multiple items into a single summary representation can enhance visual cognition (e.g., averaging multiple noisy elements can give a more precise representation than any individual element; Alvarez 2011), and can serve to guide object-based attention to salient regions of the environment. Moreover, given the connection between the processing of object ensembles, textures, and scenes via the processing of summary statistics, studying object-ensemble representation in the human brain provides us with the unique opportunity to deepen our understanding of the cognitive and neural mechanisms underlying multiple aspects of human visual perception.

Notes

Conflict of Interest: None declared.

References

Alvarez GA. 2011. Representing multiple objects as an ensemble enhances visual cognition. Trends Cogn sci. 15:122–131.

Alvarez GA, Cavanagh P. 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. Psychol Sci. 15:106–111.

Alvarez GA, Oliva A. 2008. The representation of simple ensemble visual features outside the focus of attention. Psychol Sci. 19:392–398.

Alvarez GA, Oliva A. 2009. Spatial ensemble statistics are efficient codes that can be represented with reduced attention. Proc Natl Acad Sci USA. 106:7345–7350.

Ariely D. 2001. Seeing sets: representation by statistical properties. Psychol Sci. 12:157–162.

Balas B, Nakano L, Rosenholtz R. 2009. A summary-statistic representation in peripheral vision explains visual crowding. J Vision. 9(12): 1–3, 1–18.

Bauer B. 2009. Does Steven's power law for brightness extend to perceptual brightness averaging? Psychol Rec. 59:171–186.

Brainard DH. 1997. The Psychophysics Toolbox. Spat Vis. 10:433–436.

Cant JS, Arnott SR, Goodale MA. 2009. fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. Exp Brain Res. 192:391–405.

Cant JS, Goodale MA. 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. Cereb Cortex. 17:713–731.

Cant JS, Goodale MA. 2011. Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. J Neurosci. 31:8248–8258.

Cant JS, Xu Y. 2011. Object ensemble coding is distinct from texture processing in the parahippocampal place area. J Vis (VSS Abstract). 11(11):877.

Cant JS, Xu Y. 2012. Object ensemble processing in human anterior-medial ventral visual cortex. J Neurosci. 32:7685–7700.

Chong SC, Treisman A. 2003. Representation of statistical properties. Vis Res. 43:393–404.

Chong SC, Treisman A. 2005a. Attentional spread in the statistical processing of visual displays. Percept Psychophys. 67:1–13.

Chong SC, Treisman A. 2005b. Statistical processing: computing the average size in perceptual groups. Vision Res. 45:891–900.

Cowan N. 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci. 24:87–185.

Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S. 1999. Sources of mathematical thinking: behavioral and brain-imaging evidence. Science. 284:970–974.

Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N. 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. J Neurosci. 31:11305–11312.

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

Epstein RA. 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn sci. 12:388–396.

Epstein RA, Higgins JS. 2007. Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. Cereb Cortex. 17:1680–1693.

Epstein RA, Higgins JS, Thompson-Schill SL. 2005. Learning places from views: variation in scene processing as a function of experience and navigational ability. J Cogn Neurosci. 17:73–83.

Friston KJ, Homes AP, Worsley KJ, Poline J-P, Frith CD, Frackwowiak RSJ. 1995. Statistical parametric maps in functional imaging: a general linear model approach. Hum Brain Mapp. 2:189–210.

Greene MR, Oliva A. 2010. High-level aftereffects to global scene properties. J Exp Psychol Hum Percept Perform. 36:1430–1442.

Greene MR, Oliva A. 2009. Recognition of natural scenes from global properties: Seeing the forest without representing the trees. Cogn Psychol. 58:137–176.

Grill-Spector K, Henson R, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci. 10:14–23.

Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzchak Y, Malach R. 1998. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. Human Brain Mapp. 6:316–328.

Grill-Spector K, Kushnir T, Hendler T, Malach R. 2000. The dynamics of object-selective activation correlate with recognition performance in humans. Nat Neurosci. 3:837–843.

Haberman J, Whitney D. 2007. Rapid extraction of mean emotion and gender from sets of faces. Curr Biol. 17:R751–R753.

Huettel SA, Song AW, McCarthy G. 2009. Functional magnetic resonance imaging. 2nd ed. Sunderland (MA): Sinauer Associates, Inc.

Ischebeck A, Schocke M, Delazer M. 2009. The processing and representation of fractions within the brain: an fMRI investigation. Neuroimage. 47:403–413.

Jacob SN, Nieder A. 2009a. Notation-independent representation of fractions in the human parietal cortex. J Neurosci. 29:4652–4657.

Jacob SN, Nieder A. 2009b. Tuning to non-symbolic proportions in the human frontoparietal cortex. Eur J Neurosci. 30:1432–1442.

Jacob SN, Valentin D, Nieder A. 2012. Relating magnitudes: the brain's code for proportions. Trends Cogn Sci. 16:157–166.

James TW, Culham J, Humphrey GK, Milner AD, Goodale MA. 2003. Ventral occipital lesions impair object recognition but not object-directed grasping: a fMRI study. Brain. 126:2463–2475.

Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci. 17:4302–4311.

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

Kravitz DJ, Peng CS, Baker CI. 2011. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. J Neurosci. 31:7322–7333.

Levy I, Hasson U, Avidan G, Hendler T, Malach R. 2001. Center-periphery organization of human object areas. Nat Neurosci. 4:533–539.

Loftus GR, Mason MEJ. 1994. Using confidence intervals in within-subject designs. Psychon B Rev. 1:476–490.

Luck SJ, Vogel EK. 1997. The capacity of visual working memory for features and conjunctions. Nature. 390:279–281.

MacEvoy SP, Epstein RA. 2011. Constructing scenes from objects in human occipitotemporal cortex. Nat Neurosci. 14:1323–1331.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A. 92:8135–8139.

Nieder A. 2005. Counting on neurons: the neurobiology of numerical competence. Nat Rev Neurosci. 6:177–190.

Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K. 1992. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci USA. 89:5951–5955.

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

Oliva A, Schyns PG. 2000. Diagnostic colors mediate scene recognition. Cogn Psychol. 41:176–210.

Oliva A, Torralba A. 2006. Building the gist of a scene: the role of global image features in recognition. Prog Brain Res: Visual Percept. 155:23–36.

Oliva A, Torralba A. 2001. Modeling the shape of the scene: a holistic representation of the spatial envelope. Int J Comput Vis. 42:145–175.

Oliva A, Torralba A. 2007. The role of context in object recognition. Trends Cogn Sci. 11:520–527.

Oliva A, Torralba A. 2002. Scene-centered description from spatial envelope properties. Lecture notes in computer science serie proceedings. In: Bulthoff H, Lee SW, Poggio T, Wallraven C, editors. Second international workshop of biologically motivated computer vision. Tuebingen (Germany): Springer. p. 263–272.

Park S, Brady TF, Greene MR, Oliva A. 2011. Disentangling scene content from its spatial boundary: complementary roles for the PPA and LOC in representing real-world scenes. J Neurosci. 31:1333–1340.

Park S, Konkle T, Oliva A. 2015. Parametric coding of the size and clutter of natural scenes in the human brain. Cereb Cortex. 25:1792–1805.

Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. 2001. Compulsory averaging of crowded orientation signals in human vision. Nat Neurosci. 4:739–744.

Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis. 10:437–442.

Peuskens H, Claeys KG, Todd JT, Norman JF, Hecke PV, Orban GA. 2004. Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. J Cogn Neurosci. 16:665–682.

Piazza M, Pinel P, Le Bihan D, Dehaene S. 2007. A magnitude code common to numerosities and number symbols in human intraparietal cortex. Neuron. 53:293–305.

Portilla J, Simoncelli EP. 2000. *A parametric texture model based on joint statistics of complex wavelet coefficients*. Int J Comput Vis. 40:49–71.

Pylyshyn ZW, Storm RW. 1988. Tracking multiple independent targets: evidence for a parallel tracking mechanism. Spat Vis. 3:179–197.

Rosenholtz R. 2011. What your visual system see where you are not looking. In: Rogowitz BE, Pappas TN, editors. Human vision and electronic imaging XVI. 7865:1–14.

Ross J, Burr D. 2012. Number, texture and crowding. Trends Cogn Sci. 16:196–197.

Rossion B, Pourtois G. 2004. Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic-level object recognition. Perception. 33:217–236.

Saxe R, Brett M, Kanwisher N. 2006. Divide and conquer: a defense of functional localizers. Neuroimage. 30:1088–1096; discussion 1097–1089.

Steeves JKE, Humphrey GK, Culham JC, Menon RS, Milner AD, Goodale MA. 2004. Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. J Cogn Neurosci. 16:955–965.

Stoianov I, Zorzi M. 2012. Emergence of a 'visual number sense' in hierarchical generative models. Nat Neurosci. 15:194–196.

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

Todd JJ, Han SK, Harrison S, Marois R. 2011. The neural correlates of visual working memory encoding: a time-resolved fMRI study. Neuropsychologia. 49:1527–1536.

Todd JJ, Marois R. 2004. Capacity limit of visua short-term memory in human posterior parietal cortex. Nature. 428:751–754.

Torralba A, Oliva A. 2002. Depth estimation from image structure. IEEE Pattern Anal. 24:1226–1238.

Torralba A, Oliva A. 2003. Statistics of natural image categories. Network Comp Neural. 14:391–412.

Watamaniuk SN, Duchon A. 1992. The human visual system averages speed information. Vis Res. 32:931–941.

Williams DW, Sekuler R. 1984. Coherent global motion percepts from stochastic local motions. Vis Res. 24:55–62.

Xu Y. 2002. Limitations in object-based feature encoding in visual short-term memory. J Exp Psychol Hum. 28:458–468.

Xu Y. 2010. The neural fate of task-irrelevant features in object-based processing. J Neurosci. 30:14020–14028.

Xu Y, Chun MM. 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. Nature. 440:91–95.

Xu Y, Chun MM. 2009. Selecting and perceiving multiple visual objects. Trends Cogn Sci. 13:167–174.

Xu Y, Turk-Browne NB, Chun MM. 2007. Dissociating task performance from fMRI repetition attenuation in ventral visual cortex. J Neurosci. 27:5981–5985.