Neural representation of geometry and surface properties in object and scene perception

Authors: Matthew X. Lowe[ab*], Jason Rajsic[b], Jason P. Gallivan[cde], Susanne Ferber[bf], and Jonathan S. Cant[a]

[a] Department of Psychology (Scarborough), University of Toronto, Toronto, ON, Canada, M1C 1A4
[b] Department of Psychology (St. George), University of Toronto, Toronto, ON, Canada, M5S 3G3
[c] Centre for Neuroscience Studies, Queen's University, Kingston, ON, Canada, K7L 3N6
[d] Department of Psychology, Queen's University, Kingston, ON, Canada, K7L 3N6
[e] Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON, Canada, K7L 3N6
[f] Rotman Research Institute, Baycrest, Toronto, ON, Canada, M6A 2E1

*Correspondence to: Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON, M1C 1A4, Email: matthew.lowe@utoronto.ca

Number of Figures: 7

Abstract: 200 words

Introduction: 846 words

Discussion: 2873 words

Accepted: June 19, 2017, NeuroImage

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements: This study was supported by an Ontario Graduate Scholarship (OGS) to M.X.L, a Natural Sciences and Engineering Research Council (NSERC) grant and a Connaught New Researcher Award to J.S.C, an NSERC grant and Canadian Institutes of Health Research (CIHR) grant to S.F, and J.P.G was supported by a CIHR postdoctoral fellowship.

Abstract

Multiple cortical regions are crucial for perceiving the visual world, yet the processes shaping representations in these regions are unclear. To address this issue, we must elucidate how perceptual features shape representations of the environment. Here, we explore how the weighting of different visual features affects neural representations of objects and scenes, focusing on the scene-selective parahippocampal place area (PPA), but additionally including the retrosplenial complex (RSC), occipital place area (OPA), lateral occipital (LO) area, fusiform face area (FFA) and occipital face area (OFA). Across three experiments, we examined functional magnetic resonance imaging (fMRI) activity while human observers viewed scenes and objects that varied in geometry (shape/layout) and surface properties (texture/material). Interestingly, we found equal sensitivity in the PPA for these properties within a scene, revealing that spatial-selectivity alone does not drive activation within this cortical region. We also observed sensitivity to object texture in PPA, but not to the same degree as scene texture, and representations in PPA varied when objects were placed within scenes. We conclude that PPA may process surface properties in a domain-specific manner, and that the processing of scene texture and geometry is equally-weighted in PPA and may be mediated by similar underlying neuronal mechanisms.

Keywords: structure, layout, texture, visual features, ventral visual, context, fMRI

Introduction

In only the briefest of moments, the human visual system is able to draw on a broad array of cues to efficiently identify and navigate complex environments. A fundamental question of visual perception has been how the brain represents scene information to perform this feat. Since its initial description, the parahippocampal place area (PPA) (Epstein and Kanwisher, 1998) has become a critical region for understanding the neural mechanisms underlying this ability, yet diverse claims to its function have produced ongoing debate. Emerging with the initial description of PPA, the influential spatial layout hypothesis posits this region represents the geometric structure of a scene as defined by its background elements. Evidence has since produced support for this hypothesis through the encoding of spatial features within a scene, such as structural geometry or layout (Epstein et al., 2003), spatial boundary (Park et al., 2011), and spatial depth (Kravitz et al., 2011). Conversely, a growing body of work suggests PPA plays a broader role in scene recognition, extending beyond the confines of the spatial layout hypothesis to include the processing of high-level conceptual scene categories (Walther et al., 2009; Walther et al., 2011; Dilks et al., 2011), non-spatial contextual associations of objects (Bar et al., 2008; Aminoff et al., 2007) and events (Diana, 2016), and the surface texture and material properties of isolated objects (Peuskens et al., 2004; Cant and Goodale, 2007; 2011). Evidence further suggests this region connects goal-states and context to construct a flexible neural representation of the environment by integrating multiple visual features diagnostic of scene identity (Lowe et al., 2016). Nevertheless, disentangling and directly comparing the unique contributions of individual visual elements to scene representation has been a central challenge,

and previous research has yet to elucidate the relative importance of individual visual features in shaping underlying neural responses, thus leaving these questions unanswered.

Akin to structural features, surface properties are ubiquitous within a scene, and inform our general perception and recognition of the world around us. For instance, Steeves and colleagues (2004) have shown that a patient with profound visual form agnosia (impairments in processing structure) was able to use visual texture and colour information for accurate scene recognition, suggesting these visual features play an important role in the formation of scene identity. In object perception, surface characteristics such as texture may facilitate visual search by defining edges (Biederman and Ju, 1988). Moreover, texture is instrumental in providing visual cues which aid in identification and action planning necessary for interacting with objects (Buckingham et al., 2009; Gallivan et al., 2014), and may form a contextual bridge linking an object to its surrounding environment (Lowe et al., 2015). Research has further highlighted the importance of surface properties in the perception of natural scenes, where this feature may be particularly important for the formation of scene identity (Lowe et al., 2016).

In light of the importance of both geometry and surface properties in object and scene perception, the present study aims to directly explore the relative contributions of these features across scene- and object-selective visual cortex in order to ascertain the importance of both geometry and surface properties in shaping representations of our visual world. To accomplish this, we use a novel set of images specifically designed to explore the relative weighting (i.e., levels of univariate activation) of geometry and surface properties in object and scene perception, and then compare neural representations of these features across objects and scenes. We first test the hypothesis that PPA will show equal weighting (i.e., equivalent levels of activation) to the processing of the geometry and surface properties of a scene, but greater sensitivity to the surface

properties of an object over its shape (Cant & Goodale, 2007; 2011), when scenes and objects are presented in isolation (Experiment 1). Building on previous behavioural research (Lowe et al., 2015), we next explore object-scene interactions and test the hypothesis that interactions between an object and its background context will modulate the neural relationship of shared visual features (Experiment 2). In this experiment, we combine object and scene images from the previous experiment to form a new set of scenes. Across the first two experiments, we use multivoxel pattern analyses (MVPA) to examine if the processing of scene geometry and surface properties in PPA are mediated by shared or distinct neuronal mechanisms, and also predict that the processing of these visual features in PPA is domain specific to scenes, and thus PPA would show greater activation when processing the surface properties of scenes compared with objects.

Finally, we use the fMR-adaptation approach to obtain a sensitive measure of the relative weighting of geometry and surface properties solely within scene perception in PPA (Experiment 3). Here, we predict equivalent releases from adaptation for variations in scene geometry or surface properties and an interaction (i.e., non-additivity) between the processing of these features, which would imply that their representations are not independent. In addition to examining the PPA, in all experiments we explore how geometry and surface properties contribute to neural representations in regions sensitive to processing scenes (RSC, OPA), objects (LO), and faces (FFA, OFA).

Materials and Methods

Observers

Thirty-six paid observers with normal or corrected-to-normal visual acuity were recruited from the University of Toronto community, consisting of ten paid observers (6 male; mean age

26.2 ± 4.92) in Experiment 1, twelve paid observers (6 male; mean age 25.83 years ± 3.61) in Experiment 2, and fourteen paid observers (6 male; mean age 24.21 ± 3.26) in Experiment 3. All Observers gave informed consent in accordance with the University of Toronto Ethics Review Board. One observer in Experiment 3 was removed prior to analyses due to excessive head motion (i.e., rotation and or translation in excess of 3 mm or 3°, respectively) which could not be motion-corrected within acceptable limits.

Stimuli and Procedure

E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA; Experiment 1; Experiment 2) and Matlab (MathWorks, Natick, MA; Experiment 3) were used to control stimulus presentation and collect behavioural responses. Images for all three experiments were rear-projected onto a screen in the MRI scanner (subtending 17.1° x 12.8 of visual angle), and observers viewed stimuli through a mirror mounted to the head coil directly above the eyes. In Experiment 1, stimuli were 512 unique full-colour 3-dimensional indoor scenes and objects rendered using Blender 2.0 software (Stichting Blender Foundation, Amsterdam; **Fig. 1A**) and created by varying a counterbalanced combination of scene-shape (circular; square), scene-texture (wood; brick), object-shape (circular; square), and object-texture (wood; brick). Textures were heterogeneous within a category (i.e., wood and brick), such that each category contained multiple exemplars of the same type of texture, as would be experienced in real-world environments (see **Fig. 1A**). One exemplar was used for each type of shape (i.e., circular and square; but see Experiment 3 for results when the number of texture and shape exemplars were matched). A blocked fMRI experimental paradigm was used wherein 24 images were presented in blocks of 20-s each. Each block was preceded by a 12-s fixation period and a 4-s written

instruction to attend to changes in either the texture or shape of the scene or object in the ensuing block.

In each trial (12 per block, lasting 1666 ms each), two images were presented for 300 ms (separated by a 200-ms blank interval), and the task of the observers was to decide if the attended feature (shape or texture) of the stimulus (scene or object) was the same or different across the two images, responding during a 1166-ms period following the onset of the second image (via a response pad placed in the observer's right hand). Each block contained an equal number of "same" and "different" trials. Observers were instructed to maintain central fixation and respond as quickly and as accurately as possible. Images in each condition were presented randomly within each block. Each observer took part in 5 runs (7 min 24 s each), and each run contained a unique and counterbalanced order of 12 stimulus blocks (i.e., three of each condition: scene-shape; scene-texture; object-shape; object-texture). Run order was randomized across observers and each condition (i.e., what was attended, shape or texture) was held constant per block. For Experiment 2, the procedure was identical, but the stimuli were combined to create new images containing both scenes and objects (**Fig. 1B**), with an object presented in the center of each scene.

In Experiment 3, 100 unique grayscale 3-dimensional indoor scenes were generated using Blender 2.0 software (Stichting Blender Foundation, Amsterdam; **Fig 1C; Supplementary Materials Fig. S1**). Stimuli were fully counterbalanced with 10 unique scene shapes, and 10 unique scene textures (i.e., variation in scene shape and scene texture were matched). Differences in low-level image features (luminance, contrast, and hue) were controlled using the SHINE toolbox (Willenbocket et al., 2010). A fast, event-related fMR-adaptation design was used, wherein 75 trials (6 s each) were presented in each run (7 min 42 s each), with a total of 5

runs per observer. Observers were asked to respond whether two consecutively displayed images were, as a whole, the "same" or different", without attending to any particular stimulus dimension. Conditions examined included a "no-change" condition where neither feature changed (baseline), a "both-change" condition where both features changed, and two conditions where either feature could change independently of the other ("same-shape-different-texture"; "different-shape-same-texture"). In each trial, following an initial fixation of 1-s, two images were displayed (300 ms each) and were separated by an interstimulus interval of 800 ms. Following these images, observers responded during a 3600 ms response window (via a response pad placed in the observer's right hand). Trial order was counterbalanced across five unique runs, and the order of these runs was counterbalanced across participants.

Localizer Scan

For Experiments 1 and 2, stimuli used to localize object-, scene-, and face-sensitive areas of cortex were photographs of various scenes, faces, common objects, and tile-scrambled images. Stimuli were presented in 16-s blocks of 32 images at a resolution of 375 x 375 pixels (7.8° x 7.8°), and were displayed for 400 ms each, with an interstimulus interval of 50 ms. Observers fixated on a centrally-presented black fixation cross, and were instructed to respond with a button press when the fixation cross changed from black to red (randomly occurring once or twice per stimulus block). There were 4 blocks for each stimulus category within a run, and there were two unique run orders. Sixteen-second long fixation periods were presented after each stimulus block. Each observer took part in three localizer runs (6-min 40-sec each). For Experiment 3, stimuli used to localize object and scene-sensitive areas of cortex consisted of photographs of various scenes, faces, common objects, 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 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.

MRI Acquisition

Scanning was performed at the Center for Addiction and Mental Health using a 3T GE Discovery MR750 whole-body MRI scanner equipped with an 8-channel head coil. T1-weighted anatomical images were acquired using a 3D SAG T1 BRAVO spiral pulse sequence [repetition time (TR), 6736 ms; echo time (TE), 3 ms; inversion time, 650 ms; flip angle 8°, 256 x 256 matrix size, 200 slices, 1 mm isovoxel]. For the functional runs, T2*-weighted images sensitive to blood oxygenation level-dependent (BOLD) contrasts were acquired using a spiral pulse sequence (64 x 64 matrix size; field of view 22 cm; TR 2000 ms; TE 30 ms; flip angle 60°; 200 volumes for the localizer runs in Experiments 1 and 2 and 140 volumes for the localizer runs in Experiment 3, 134 volumes for the main experimental runs in Experiments 1 and 2 and 231 volumes for the adaptation runs in Experiment 3). Thirty-one slices (3.4 mm x 3.4 mm x 5 mm, no gap) parallel to the anterior and posterior commissure line were collected in all functional runs.

Univariate Data Analysis

fMRI data were processed and analyzed using BrainVoyager QX 2.8 (Brain Innovation, Maastricht, the Netherlands). Data preprocessing included slice acquisition time correction, 3D motion correction, temporal filtering (linear trend removal and high-pass filtering set at 3 cycles/run), and Talairach space transformation (Talairach and Tournoux, 1988). Data from the functional localizer was analyzed using a general linear model (GLM), accounting for hemodynamic response lag (Friston et al., 1994). Regions of interest (ROIs) can be seen in Figure 2. The PPA ROI was defined as a region in the collateral sulcus and parahippocampal gyrus (see Epstein and Kanwisher, 1998) whose activation was 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; identified in all observers in all Experiments). In addition, the RSC (see Epstein and Higgins, 2005) and OPA (also known as transverse occipital sulcus; see Dilks et al., 2013) ROIs were functionally defined as regions in restrosplenial cortex–posterior cingulate–medial parietal cortex and transverse occipital cortex, respectively, whose activations were higher for scenes than for faces and objects (identified in eleven, and nine, observers, respectively, in Experiment 1, 7 and 9 observers in Experiment 2, and 12 and 13 observers in Experiment 3). In accordance with Grill-Spector et al. (2000), LO, a sub-division of the lateral occipital complex (LOC), was defined as a region in the lateral occipital cortex near the posterior inferotemporal sulcus, with activation higher for objects than for scrambled objects (identified in all observers in all experiments). The fusiform face area (FFA) was selected as a control region (as our stimuli did not contain any faces) and following Kanwisher et al. (1997), this area was defined as a region in the extrastriate cortex whose activations were higher for faces than scenes or objects (identified in 11 observers in Experiment 1, all observers in Experiment 2, and 12 observers in Experiment 3). As an additional control region, the occipital face area (OFA) was

defined as a region in the inferior occipital gyrus (Gauthier et al., 2000) whose activations were higher for faces than scenes or objects (identified in 9 observers in Experiment 1, 11 observers in Experiment 2, and 11 observers in Experiment 3).

Following the standard ROI-based analysis approach (Saxe et al., 2006), we overlaid the ROIs onto the data from our main experiment and extracted time courses from each observer. 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 (Xu and Chun, 1997; Xu, 2010; Cant and Xu, 2012). This was done separately for each observer in each ROI, and the resultant peak responses were then averaged across all observers. The average levels of peak activation (measured in percent BOLD signal change from baseline fixation) for each condition across observers were subjected to a 2 (Experiment 1 and 2: Stimulus: object vs. scene; Experiment 3: Texture: same vs. different) x 2 (Experiment 1 and 2: Feature: texture vs. shape; Experiment 3: Shape: same vs different) repeated-measures ANOVA (SPSS, Chicago, IL, USA) for each ROI (PPA, RSC, OPA, LO, FFA, OFA). Planned pairwise comparisons (Bonferroni-corrected for multiple comparisons) were then conducted to compare the processing of texture and shape for both objects and scenes in each ROI in Experiments 1 and 2, and to compare same versus different conditions for scene texture and scene shape in Experiment 3. Left and right hemispheres were combined for each ROI in all analyses (see Supplementary Materials).

Multivoxel Pattern Analysis (MVPA)

Support Vector Machine Classifiers. Pattern classification was performed in Experiments 1 and 2 with a combination of in-house software (using Matlab) and the Princeton MVPA Toolbox for Matlab (http://code.google.com/p/princeton-mvpa-toolbox/) using a Support Vector Machines

(SVM) classifier (libSVM, http://www.csie.ntu.edu.tw/~cjlin/libsvm/). The SVM model used a linear kernel function and a constant cost parameter, C=1, to compute a hyperplane that best separated the block/condition responses. To test the accuracy of the SVM classifiers we used a "leave-one-run-out" N-fold cross-validation, in which a single fMRI run was reserved for classifier testing. We performed this N-1 cross-validation procedure until all runs were separately tested, and then averaged across N-iterations in order to produce a representative classification accuracy measure for each participant, ROI, and pattern discrimination (Duda et al., 1995).

Multiclass and Pairwise Discriminations. SVMs are designed for classifying differences between two stimuli and LibSVM (the SVM package implemented here) uses the so-called 'one-against-one method' for classification (Hsu and Lin, 2002). With the SVMs we performed two complementary types of classification analyses; one in which the multiple pairwise results were combined in order to produce multiclass discriminations (distinguishing among all 4 of our condition types) and another in which the individual pairwise discriminations were examined and tested separately.

The multiclass discrimination approach allowed for an examination of the distribution of the classifier guesses through visualization of the resulting 'confusion matrix'. In a confusion matrix, each row (*i*) represents the instances of the actual condition and each column (*j*) represents the predicted condition. Their intersection (*i*, *j*) represents the (normalized) number of times a given condition *i* is predicted by the classifier to be condition *j*. Thus the confusion matrix provides a direct visualization of the extent to which a decoding algorithm confuses (or correctly identifies) the different classes. All correct classifications are located in the diagonal of the matrix (with classification errors represented by non-zero values outside of the diagonal) and

average decoding performance is defined as the mean across the diagonal. The values in each row sum to 1 (i.e., 100% classification). If decoding is at chance levels, then classification performance will be at 1/4 = 25%. For all multiclass discriminations, we statistically assessed decoding significance across participants (for each ROI and condition epoch) using one-tailed t-tests versus 25% chance decoding. For pairwise discriminations, we statistically assessed decoding significance across participants using one-tailed t-tests versus 50% chance decoding. Importantly, an FDR correction of $q \leq 0.05$ was applied to the pairwise comparisons based on the number of comparisons examined per ROIs, and for the multiclass discriminations based on the number of ROIs examined (Benjamini and Hochberg, 1995).

Inputs to the SVM Classifier. BOLD percent signal change values for each ROI provided inputs to the SVM classifier. The percent signal change response was computed from the time-course activity for the task-evoked responses with respect to the time-course of a run-based averaged baseline fixation value, for all voxels in the ROI. The baseline fixation window was defined as a time point prior to the 4-sec instruction period before each stimulus block (6 sec prior to block onset, averaged across all blocks within an experimental run). For the block-evoked activity we extracted, for each condition, the average of imaging volumes 3-10 (i.e., 6-20 sec), which are time points encompassing the first peak of the hemodynamic response until the end of the experimental block. This windowed-average percent signal change classification approach corresponds with that used in recent work using the same technique (Gallivan et al., 2013; Gallivan et al., 2014; Lowe et al., 2016). Following the extraction of each block's activity, these values were rescaled between -1 and +1 for each voxel pattern within an ROI (Misaki et al., 2010).

Behavioural data analysis

Behavioural performance measures of accuracy and reaction time were recorded using E-Prime 2.0 software (Experiments 1 and 2) and Matlab (Experiment 3), and analyzed with SPSS, by performing a 2 (stimulus: object vs. scene) x 2 (feature: texture vs. layout) repeated-measures ANOVA for Experiments 1 and 2, and a 2 (texture: same vs. different) by 2 (shape: same vs. different) repeated-measures ANOVA for Experiment 3, with subsequent pairwise comparisons (all two-tailed and Bonferroni corrected) conducted based on a priori theoretical motivation (i.e., examining differences between shape and texture processing for objects and scenes separately in Experiments 1 and 2, and examining potential releases from adaptation when scene texture varied but scene shape was held constant, and when scene shape varied but scene texture was held constant, in Experiment 3). See supplementary material for results of the behavioural analyses for all three experiments.

Experiment 1: Feature processing in objects and scenes separately

In Experiment 1, observers viewed either indoor scenes or objects, separately (**Fig. 1A**). Each scene and object could change along two dimensions: shape (circular; square) and texture (wood; brick). Given previous findings highlighting the influence of both surface properties (Lowe et al., 2016) and spatial layout (Epstein and Kanwisher, 1998) on activity in scene-selective cortex, we hypothesize similar univariate neural activity between scene-texture and scene-shape conditions in PPA (i.e., equal weighting). We further predict, based on previous results, increased univariate activity when attending to object-texture compared with object-shape in PPA (Cant and Goodale, 2007; 2011), and vice versa in the object-selective LO, which

has been shown to represent higher level object-shape information (Kourtzi and Kanwisher, 2001). To test the hypothesis that PPA processes surface properties in a domain specific manner, we examine neural representations of this visual feature across object and scene perception, predicting that PPA will show increased activity to the surface properties of a scene over an object. Moreover, to test the prediction that geometry and surface properties are processed similarly within scene perception (i.e., are potentially mediated by shared neuronal mechanisms), and distinctly within object perception, we use multivariate techniques to explore whether these features can be discriminated from one another in both objects and scenes. As the stimuli used here were not tailored to the known functional properties of FFA (Kanwisher et al., 1997) and OFA (Gauthier et al., 2000), these regions were used as controls (see **Fig. 2** for all ROIs).

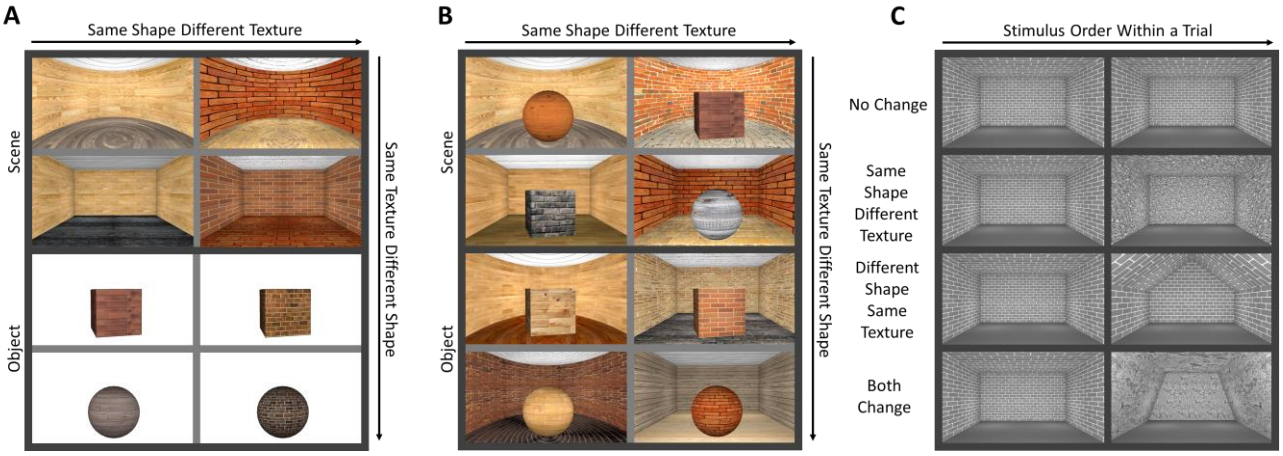

Figure 1. Experimental Stimuli. (A) Examples of stimuli used in Experiment 1. Scenes and objects are defined by their shape (circular vs. square) and texture (wood vs. brick). Observers attended to the shape or texture of the object or scene, either of which could change while the other was held constant. (B) Examples of stimuli used in Experiment 2. The stimuli and procedure were identical to Experiment 1, with the exception that objects were placed within scenes. (C) Examples of the stimuli used in Experiment 3. Scenes could vary across 10 different shapes, and 10 different textures. In Experiment 3, observers attended only to overall changes across images, and did not attend directly to any one particular feature. For additional examples of the stimuli used in Experiment 3, see Supplementary Materials **Fig. S1**.

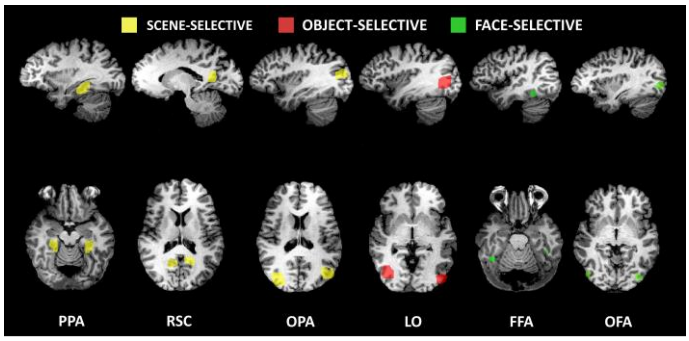

Figure 2. Regions of interest. Functionally defined ROIs are shown on the brains of two representative observers. Talairach coordinates for peak voxels of each ROI in observer one (PPA – FFA) and two (OFA) are as follows: LPPA, -29, -34, -10; RPPA, 27, -34, -5; LRSC, -15, -56, 14; RRSC, 14, -59, 17; LOPA, -41, -72, 14, ROPA, 32, -80, 19; LLO, -45, -80, 0; RLO, 39, -71, 5; LFFA, -43, -36, -17; RFFA, 41, -47, -13; LOFA, -36, -79, 4; ROFA, 43, -75, -3.

We first examined the univariate response amplitudes (percent signal change compared to fixation) of ROIs for each condition (scene-shape; scene-texture; object-shape; object-texture) using 2 x 2 repeated-measures ANOVAs with stimulus (object; scene) and feature (texture; shape) as factors **(Fig. 3;** collapsing left and right hemispheres for each region; this applies to all Experiments), and conducted simple main effects analyses to examine interactions between stimulus category and feature in each region. We found a significant main effect of stimulus for all regions except FFA (PPA: $F_{1,9} = 143.98$, $p < 0.001$; RSC: $F_{1,6} = 381.14$, $p < 0.001$; OPA: $F_{1,8} = 101.42$, $p < 0.001$; LO: $F_{1,9} = 22.08$ $p = 0.001$; FFA: $F_{1,9} = 0.36$ $p = 0.566$; OFA: $F_{1,9} = 16.17$ $p = 0.003$), and a main effect of feature for PPA ($F_{1,9} = 24.76$, $p = 0.001$), LO ($F_{1,9} = 11.00$, $p = 0.009$), and OFA ($F_{1,9} = 6.99$ $p = 0.027$), but not for the remaining regions (all Fs < 4.59; all ps ≥ 0.061). We observed significant stimulus-by-feature interactions in all scene-selective regions (PPA: $F_{1,9} = 35.03$, $p < 0.001$; RSC: $F_{1,6} = 9.12$, $p = 0.023$; OPA: $F_{1,8} = 20.63$, $p = 0.002$), but not in object-selective cortex (LO: $F_{1,9} = 0.06$, $p = 0.811$) or face-selective cortex (FFA: $F_{1,9} = 0.89$ $p = 0.369$; OFA: $F_{1,9} = 0.28$ $p = 0.609$).

These findings confirm a dissociation between object and scene processing across scene-selective and object-selective areas of cortex. In line with our predictions, subsequent planned pairwise comparisons (two-tailed and Bonferroni-corrected; this applies to all Experiments) revealed no significant differences between scene-shape and scene-texture processing in PPA (t_9 = 0.17, p = 0.866). Similarly, this was also observed in RSC (t_6 = 1.75, p = 0.131), yet higher BOLD response for scene-shape over scene-texture was observed in OPA (t_8 = 3.09, p = 0.015). Moreover, consistent with our predictions, analysis of object-processing revealed significantly higher activation for object-texture over object-shape in PPA (t_9 = 12.61, p < 0.001) but not RSC (t_6 = 1.60, p = 0.170), replicating previous findings (Cant and Goodale, 2007; 2011). We further observed higher activation for object-shape over object-texture in LO (t_9 = 2.30, p = 0.047), as expected. Finally, higher activation for object-texture over object-shape was also found in OPA (t_8 = 3.15, p = 0.014). This latter result was unexpected given the sensitivity of OPA to scenes, yet we speculate this finding may speak to the involvement of OPA in the processing of local elements, which may contain cues for scene recognition and navigation (Kamps et al., 2016).

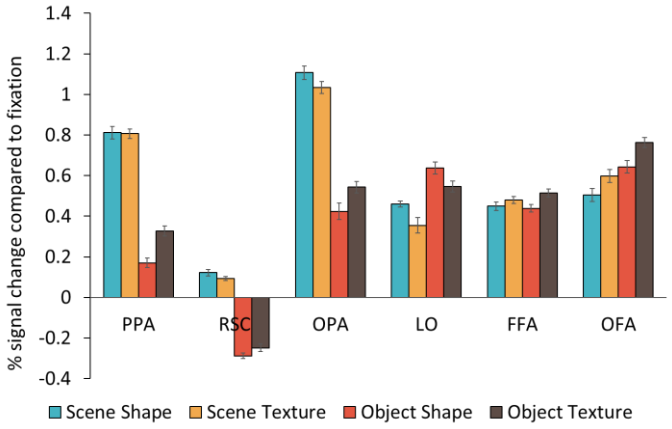

Figure 3. Univariate response amplitudes for Experiment 1. BOLD signal activation for all conditions (attend scene shape; attend scene texture; attend object shape; attend object texture) in each ROI. Data are represented as mean ± SEM.

Interactions across ROIs were investigated using a 2 (region) by 2 (stimulus) by 2 (feature) repeated-measures ANOVA. This analysis revealed significant region-by-stimulus interactions for PPA with RSC ($F_{1,6} = 45.43$, $p = 0.001$), LO ($F_{1,9} = 142.00$, $p < 0.001$), FFA ($F_{1,9} = 117.08$, $p < 0.001$), and OFA ($F_{1,9} = 134.46$ $p < 0.001$); for RSC with OPA ($F_{1,5} = 14.68$, $p = 0.012$), LO ($F_{1,9} = 98.15$, $p < 0.001$), FFA ($F_{1,6} = 160.90$, $p < 0.001$), and OFA ($F_{1,6} = 208.34$ $p < 0.001$); for OPA with LO ($F_{1,8} = 118.29$, $p < 0.001$), FFA ($F_{1,8} = 76.99$, $p < 0.001$) and OFA ($F_{1,8} = 122.78$ $p < 0.001$); for LO with FFA ($F_{1,8} = 19.93$, $p = 0.002$), and for FFA with OFA ($F_{1,9} = 27.81$ $p = 0.001$). Significant region-by-feature interactions were found for PPA with RSC ($F_{1,6} = 6.13$, $p = 0.048$), OPA ($F_{1,8} = 5.51$, $p = 0.047$), and LO ($F_{1,9} = 40.08$, $p < 0.001$); for RSC with OFA ($F_{1,6} = 6.11$ $p = 0.048$), for OPA with LO ($F_{1,8} = 17.77$, $p = 0.003$); and for LO with FFA ($F_{1,9} = 50.70$, $p < 0.001$) and OFA ($F_{1,9} = 21.67$ $p = 0.001$). A significant three-way interaction was found for PPA with LO ($F_{1,9} = 7.51$, $p = 0.023$), FFA ($F_{1,9} = 6.09$, $p = 0.036$), and OFA ($F_{1,9} = 12.34$ $p = 0.007$), and for OPA with OFA ($F_{1,8} = 5.52$ $p = 0.047$) These results are consistent with previous work demonstrating a functional dissociation between processing within scene-selective, object-selective, and face-selective cortex, but also within the scene-processing network itself (Cant & Goodale, 2007, 2011; Cant & Xu, 2012, 2015, 2016; Lowe et al., 2016).

Since a null result in a univariate analysis does not necessarily imply that a given region cannot distinguish between two properties, we conducted multivariate analyses to examine whether each of our ROIs could discriminate between stimulus category and feature. The aims of these analyses were twofold. Firstly, we investigated the extent to which each ROI could successfully discriminate generally across all conditions (object shape, object texture, scene shape, scene texture), and secondly, we conducted a number of pairwise comparisons in order to directly explore whether these regions could successfully discriminate between the geometry and

surface properties of objects and scenes. We first extracted multivoxel fMRI activity and used linear SVM classifiers in each region to create confusion matrices representing the distribution of classifications (and misclassifications) across conditions (**Fig. 4A**). These matrices demonstrate misclassifications in scene- and object-cortex are largely contained within the same stimulus category (e.g., scene-texture is more likely to be misclassified as scene-shape than object-texture). We then examined the extent to which each of our conditions could be decoded above chance (25%). These multiclass discriminations revealed classification accuracies that were significantly above chance for all ROIs (all ts ≥ 3.21; all ps ≤ 0.005; all qs ≤ 0.005), indicating that each region classifies conditions with above-chance accuracy (**Fig. 4B)**. To investigate these findings in greater detail, we performed subsequent analyses using planned pairwise-comparisons corrected for multiple comparisons using the FDR procedure (*q;* based on the number of comparisons per ROI) to examine whether shape and texture could be decoded at greater-than-chance accuracy (50%) when attending to either objects or scenes, separately (**Fig. 4C**).

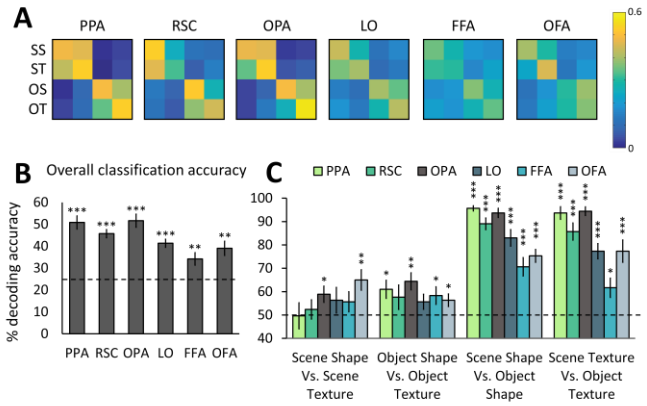

Figure 4. (Above) Confusion matrices for Experiment 1. (A) Confusion matrices (chance = 0.25) for Experiment 1 generated from multiclass discriminations showing the distribution of classification errors across all four conditions (SS = Scene Shape; ST = Scene Texture; OS = Object Shape; OT = Object Texture) for each ROI. The average classifier response proportions across participants are shown. When decoding is perfect, the confusion matrix will have a diagonal containing values of 1 and the rest of the matrix will be 0. Note that the average decoding performance (shown in Fig. 4B) is defined as the mean across the diagonal. To highlight differences in decoder performance, the matrices have been rescaled (rather than being scaled from 0-1) **(Below) Multivariate results for Experiment 1.** (B) Overall classification accuracy (chance = 25%; dashed line) of all four conditions for each ROI. (C) Decoding accuracy (chance = 50%; dashed line) for scene feature discriminations (shape vs. texture), object feature discriminations (shape vs. texture), shape category discriminations (scene vs. object) and texture category discriminations (scene vs. object) for each ROI. Data are represented as mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001, all p-values shown have been FDR-corrected

Findings revealed no significant decoding of scene-shape versus scene-texture in both PPA and RSC (both ts ≤ 0.58; both ps ≥ 0.29), replicating previous findings using real-world scene stimuli (Lowe et al., 2016), and thus provides additional evidence that scene-texture and scene-shape may be processed similarly in these regions. In contrast, we found significant decoding of scene-shape versus scene-texture in OPA ($t_8 = 2.56$, $p = 0.017$, $q = 0.017$). Significant discrimination of these features was not observed in LO ($t_9 = 1.16$, $p = 0.139$), nor in FFA ($t_9 = 1.32$, $p = 0.110$), but it was found in OFA ($t_9 = 3.44$, $p = 0.004$, q = 0.005). Continuing our analysis, we found significantly above-chance classification accuracy for object-shape versus object-texture in PPA ($t_9 = 2.83$, $p = 0.010$, $q = 0.013$), providing mounting evidence consistent with previous univariate findings that the processing of these object features is dissociated in PPA (Cant and Goodale, 2007; 2011). Significant decoding of these object features was found in OPA ($t_8 = 3.97$, $p = 0.002$, $q = 0.010$), but only marginally significant decoding of these features was found in LO ($t_9 = 1.76$, $p = 0.056$), and no significant decoding was observed in RSC ($t_6 = 1.46$, $p = 0.098$). We found significant decoding in both FFA ($t_9 = 2.21$, $p = 0.027$, $q = 0.045$) and OFA ($t_9 = 2.35$, $p = 0.022$, q = 0.022) across these conditions. Finally, to ensure a null finding in one of our main ROIs (e.g., PPA) could not be attributed to a problem with the classification procedure itself, we conducted control pairwise comparisons in which we expected to find significantly above-chance classification. Specifically, we examined the classification of scenes versus objects when holding stimulus feature constant, and found significant decoding of scene-shape versus object-shape for all regions (all ts ≥ 5.25; all ps ≤ 0.001; all qs ≤ 0.001), as

well as significant decoding of scene-texture versus object-texture for all regions (all ts ≥ 2.82; all ps ≤ 0.01; all qs ≤ 0.01).

Experiment 2: Feature processing in scenes containing an object

Experiment 1 provides strong evidence that feature processing of a scene extends beyond spatial features to include surface properties such as texture, and that these features elicit similar neural activity in PPA. In other words, the processing of scene texture and scene shape may be weighted equally in PPA. Scenes rarely exist independently of objects, however, and previous research has indicated object-processing may have an interactive relationship with scene-processing (Joubert et al., 2007; Mullin et al., 2013) and that geometry and surface properties may influence this relationship (Lowe et al., 2015). To examine this interaction (i.e., the presence of a surrounding scene influencing object representation, and vice versa), we expanded upon the findings of Experiment 1 by examining the processing of shape and texture while observers viewed new images depicting an object placed within a scene (**Fig. 1B**). This allowed us to attempt to replicate the findings from Experiment 1, but importantly, to test the prediction that shared visual object and scene features are not processed independently, but interact across stimulus categories. Specifically, since the surface properties of a scene can exhibit global precedence over the surface properties (but not the geometry) of an object within that scene (Lowe et al., 2015), we might observe decreased sensitivity to object surface properties relative to the processing of object geometry in PPA (e.g., equivalent activation for object texture and shape, or less activation for object texture). This would differ from the results in Experiment 1, in which objects were not presented within scenes.

Results and Discussion

Consistent with Experiment 1, we examined univariate response amplitudes in each region by conducting a 2 x 2 repeated-measures ANOVA with stimulus (object; scene) and feature (texture; shape) as factors (**Fig. 5**), and found significant main effects of stimulus for all regions except one of our control regions, FFA (PPA: $F_{1,11} = 88.40$, $p < 0.001$; RSC: $F_{1,10} = 16.74$, $p = 0.002$; OPA: $F_{1,10} = 110.83$, $p < 0.001$; LO: $F_{1,11} = 11.62$ $p = 0.006$; FFA: $F_{1,10} = 2.20$ $p = 0.169$; OFA: $F_{1,10} = 70.96$ $p < 0.001$), but only a main effect of feature for LO (PPA: $F_{1,11} = 2.25$, $p = 0.162$; RSC: $F_{1,10} = 0.004$, $p = 0.952$; OPA: $F_{1,10} = 0.334$, $p = 0.576$; LO: $F_{1,11} = 12.53$, $p = 0.005$; FFA: $F_{1,10} = 3.361$ $p = 0.097$; OFA: $F_{1,10} = 0.15$ $p = 0.711$). Interestingly, the stimulus-by-feature interaction was non-significant in all regions (all Fs < 3.43; all ps ≥ 0.094), which differs from the results of Experiment 1, where scenes and objects were presented separately. Subsequent pairwise comparisons found that activation levels for scene-shape and scene-texture did not differ in any scene selective region (PPA: $t_{11} = 0.35$, $p = 0.734$; RSC: $t_{10} = 0.19$, $p = 0.855$; OPA: $t_{10} = 1.00$, $p = 0.341$), and neither did activation for object-shape and object-texture in PPA ($t_{11} = 1.69$, $p = 0.118$). In contrast, levels of activity for these features were dissimilar in LO, with significantly higher activity observed when attending to object-shape compared with object-texture ($t_{11} = 2.53$, $p = 0.028$), as predicted. In line with our predictions and consistent with the results of Experiment 1, these findings reveal that PPA exhibits similar BOLD responses across scene features. In contrast to the results of Experiment 1, however, we observed equal sensitivity to the processing of object-texture and object-shape in PPA, which is consistent with our prediction regarding object-scene interactions.

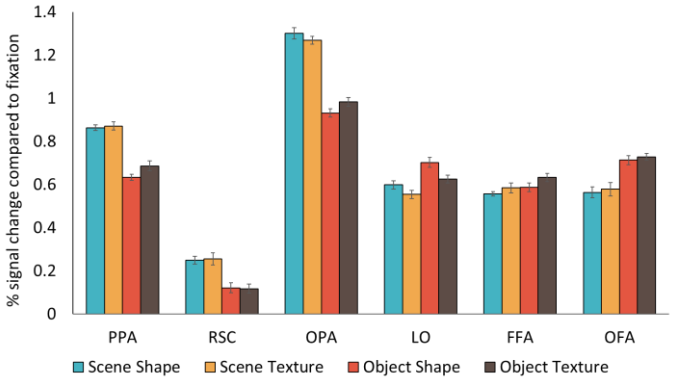

Figure 5. Univariate response amplitudes for Experiment 2. BOLD signal activation for all conditions (attend scene shape; attend scene texture; attend object shape; attend object texture) in each ROI. Data are represented as mean ± SEM.

To explore these interactions more directly, we conducted a 2 (stimulus: object vs. scene) by 2 (feature: shape vs. texture) by 2 (Experiment: Experiment 1 vs. Experiment 2) mixed-design ANOVA for PPA across Experiments 1 and 2. Significant main effects of Experiment ($F_{1,20}$ = 5.19, $p < 0.034$) and stimulus were observed ($F_{1,20}$ = 246.85, $p < 0.001$), along with a significant stimulus-by-Experiment interaction ($F_{1,20}$ = 52.05, $p < 0.001$). Moreover, a significant main effect of feature was found ($F_{1,20}$ = 16.41, $p = 0.001$), but the feature-by-Experiment interaction was not significant ($F_{1,20}$ = 3.02, $p = 0.098$). A significant stimulus-by-feature interaction was found ($F_{1,9}$ = 18.12, $p < 0.001$), and importantly, a significant three-way interaction was also observed ($F_{1,20}$ = 5.89, $p = 0.025$), warranting further analyses. Subsequent pairwise comparisons revealed no significant differences between the representation of either the geometry (t_{20} = 0.43, $p = 0.671$) or surface properties (t_{20} = 0.51, $p = 0.611$) of a scene across experiments. In contrast, we found a significant difference between the representation of the geometry (t_{20} = 4.98, $p <$ 0.001) and surface properties (t_{20} = 4.13, $p = 0.001$) of an object across experiments. Not surprisingly, higher response amplitudes in PPA for both of these object features was observed when the object was placed within a scene. But importantly, these results demonstrate that the

representation in PPA changes when objects are placed within the context of a scene (specifically, the representation of object, but not scene, features changes).

Significant region-by-stimulus interactions were found for PPA with OPA ($F_{1,10}$ = 17.293, p = 0.002), LO ($F_{1,11}$ = 67.07, p < 0.001), FFA ($F_{1,10}$ = 126.45, p < 0.001), and OFA ($F_{1,10}$ = 242.61 p < 0.001); for RSC with OPA ($F_{1,9}$ = 21.25, p = 0.001), LO ($F_{1,10}$ = 35.46, p < 0.001), FFA ($F_{1,10}$ = 20.52, p = 0.001), and OFA ($F_{1,9}$ = 58.17 p < 0.001); for OPA with LO ($F_{1,10}$ = 61.81, p < 0.001), FFA ($F_{1,9}$ = 74.95, p < 0.001), and OFA ($F_{1,9}$ = 170.46 p < 0.001), and for FFA with OFA ($F_{1,9}$ = 95.96 p < 0.001). Significant region-by-feature interactions were found for PPA with LO ($F_{1,11}$ = 35.60, p < 0.001), for RSC with LO ($F_{1,10}$ = 10.60, p = 0.009), for OPA with LO ($F_{1,10}$ = 15.42, p = 0.003), and for LO with FFA ($F_{1,10}$ = 25.89, p < 0.001). A significant three-way interaction was found for RSC with OPA ($F_{1,9}$ = 12.24, p = 0.007) and OPA with LO ($F_{1,10}$ = 7.69, p = 0.020) and OFA ($F_{1,9}$ = 19.00 p = 0.002). Taken together, the results from Experiments 1 and 2 demonstrate that scene-, object-, and face-sensitive regions of cortex process the same visual input in appreciably different ways, and speak to varying levels of functional specificity within ventral visual cortex (Cant & Goodale, 2007, 2011; Cant & Xu, 2012, 2015; Lowe et al., 2016).

To further explore the processing of geometry and surface properties across ventral-visual cortex, we conducted a number of multivariate analyses consistent with Experiment 1. After creating confusion matrices (**Fig. 6A)**, multiclass discriminations revealed classification accuracies that were significantly above chance (25%) for all ROIs (all ts ≥ 4.70; all ps ≤ 0.001, all qs ≤ 0.001), indicating that each region classifies conditions with above-chance accuracy (**Fig. 6B)**. To investigate these findings in greater detail, we performed subsequent analyses using planned pairwise-comparisons to examine whether shape and texture could be decoded at

greater-than-chance accuracy (50%) when attending both objects and scenes, separately (**Fig. 6C**). Critically, and consistent with the results of Experiment 1, we found no evidence for significant discrimination of scene-shape and scene-texture conditions in both PPA and RSC (both $ts \leq 1.05$; both $ps \geq 0.159$), suggesting these features are processed similarly within these regions. In contrast, we found significant decoding of scene-shape versus scene-texture in OPA ($t_{10} = 2.09$, $p = 0.031$, $q = 0.031$), and observed similar results in LO ($t_{11} = 2.15$, $p = 0.027$, $q = 0.027$). We further found no difference across these conditions in one control region, FFA ($t_{10} = 0.88$, $p = 0.201$), but did find a difference in OFA ($t_{10} = 3.08$, $p = 0.006$, q = 0.008).

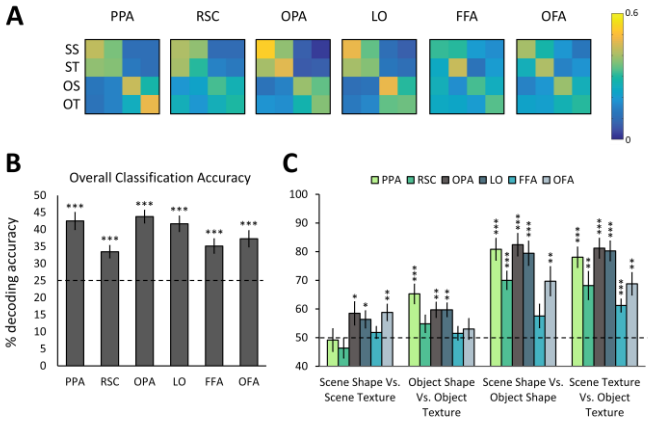

Figure 6. (Above) Confusion matrices for Experiment 2. (A) Confusion matrices (chance = 0.25) for Experiment 2 generated from multiclass discriminations showing the distribution of classification errors across all four conditions (SS = Scene Shape; ST = Scene Texture; OS = Object Shape; OT = Object Texture) for each ROI. The average classifier response proportions across participants are shown. When decoding is perfect, the confusion matrix will have a diagonal containing values of 1 and the rest of the matrix will be 0. Note that the average decoding performance (shown in Fig. 6c) is defined as the mean across the diagonal. To highlight differences in decoder performance, the matrices have been rescaled (rather than being scaled from 0-1). (Below) **Multivariate results for Experiment 2.** (B) Classification accuracy (chance = 25%; dashed line) of all four conditions for each ROI. (C) Decoding accuracy (chance = 50%; dashed line) for scene feature discriminations (shape vs. texture), object feature discriminations (shape vs. texture), shape category discriminations (scene vs. object) and texture category discriminations (scene vs. object) for each ROI. Data are represented as mean ± SEM. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, all p-values shown have been FDR-corrected

We next examined the decoding of object-shape versus object-texture, and found significantly above-chance classification accuracy in PPA (t_{11} = 4.50, p < 0.001, q = 0.001), despite similar univariate response amplitudes. These results confirm a dissociation of these object features in PPA consistent with previous findings (Cant and Goodale, 2007; 2011), and further highlight the advantage of an approach using both univariate and multivariate analyses: A null result in univariate response does not necessarily imply a null multivariate result. Thus, examining both enabled us to better characterize the relationship between the weighting of visual features and the degree to which they are represented by shared or distinct neuronal populations. We further observed significant decoding of these object features in OPA (t_{10} = 3.68, p = 0.002, q = 0.003) and LO (t_{11} = 4.10, p < 0.001, q = 0.001), but not in RSC (t_{10} = 1.60, p = 0.070), FFA (t_{10} = 0.64, p = 0.267), and OFA (t_{10} = 0.84, p = 0.210). Finally, we found significant decoding of scene-shape versus object-shape in all regions (all ts ≥ 3.90; all ps ≤ 0.002, all qs ≤ 0.003) except FFA (t_{11} = 1.86, p = 0.046, q = 0.093). For the decoding of scene-texture versus object-texture, we found significant classification in all regions (all ts ≥ 3.72; all ps ≤ 0.002, all qs ≤ 0.004). These latter sets of results demonstrate that any null result in decoding accuracy (e.g., scene shape vs. scene texture in PPA) cannot be explained by imperfections in the classification algorithm itself.

Experiment 3: fMR-adaptation of scene features

The results of the first two experiments suggested that the processing of geometry and surface properties are weighted equally in scene representation, and are possibility mediated by shared neuronal mechanisms. To provide an additional test of this hypothesis, we utilized a fast event-related fMR-adaptation paradigm, which allowed us to examine the neural representation of one feature (e.g., scene texture), independent of changes in a second feature (e.g., scene

shape). If scene shape and scene texture are not weighted equally in PPA, then changing each feature in isolation should result in significantly different releases from adaptation (compared to a no feature change baseline). If these features are weighted equally, however, then we should observe equivalent releases from adaptation. Based on the results from the previous experiments, we predict the latter scenario. Moreover, by examining potential interactions between scene shape and scene texture processing, we are able to assess whether these features are represented additively or non-additively. Given our previous results suggesting equal weighting of these features in PPA (mediated by potentially shared underlying neuronal mechanisms), we predict an interaction (i.e., non-additivity) between the processing of these features, which would imply that their representations are not independent in PPA.

Furthermore, previous work has shown that goal-states and attentional task demands directly influence activity in ventral visual cortex (Harel et al., 2014; Lowe et al., 2016). In addition to this potential influence, there is strong evidence to suggest that scene perception may be impacted by lower-level stimulus properties such as colour (Oliva and Schyns, 2000; Steeves et al., 2004; Goffaux et al., 2005; Castelhano and Henderson, 2008). Together, these findings raise the possibility that the results of our previous experiments, which utilized full-colour stimuli and manipulated attention to particular stimulus features, may be partially explained by these factors. Thus, we controlled for these potential caveats by examining the response properties of ventral-stream regions when observers were not explicitly attending to a particular stimulus dimension, but instead were performing a more general same-different judgement (**Fig. 1C**). We focus this investigation exclusively on scenes to avoid potential modulation by interactions between objects and scenes (see Experiment 2). Additionally, we controlled for low-level image properties by using grayscale scenes and then processing these images by using the

SHINE toolbox (Willenbockel et al., 2010), which equates low-level image attributes across images by normalizing luminance, contrast, and hue.

Finally, it is possible that the equal univariate activation for scene shape and scene texture in PPA in Experiments 1 and 2 resulted from unmatched variation across these features. That is, because we used only two instances of scene shape (i.e., round vs. square), but many instances of scene texture (i.e., many different types of brick and wood textures), blocks where participants attended to scene shape would result in more adaptation (i.e., less activation) compared with blocks where they attended to scene texture. If the representation of these features is not equally weighted in PPA (i.e., attending to scene shape normally elicits greater activation then attending to scene texture), then over adapting scene shape compared with scene texture would give the appearance of equivalent univariate activation and thus equal weighting. To investigate this possibility, we matched the variation across scene geometry and texture by using ten different scene-shapes that could each be rendered in ten different scene-textures. If equivalent releases from adaptation are observed when changing scene shape and scene texture independently, then the results of Experiments 1 and 2 cannot be explained by unmatched variation across these features.

Results and Discussion

fMR-adaptation to geometry and surface properties was analyzed using a 2 x 2 repeated-measures ANOVA with shape (same; different) and texture (same; different) as factors (**Fig. 7**). Importantly, this design allowed us to examine the interaction between scene shape and texture processing, but similar results (with regard to releases from adaptation for variations in scene shape or scene texture) were obtained when we used a one-way repeated measures ANOVA with the factor condition (no change vs. same shape, different texture vs. different shape, same texture

vs. both change) to analyze the data. Results from the 2 x 2 analysis revealed a significant main effect of shape for PPA ($F_{1,12} = 8.22$, $p = 0.014$), RSC ($F_{1,11} = 8.13$, $p = 0.016$), and LO ($F_{1,12} = 16.59$, $p = 0.002$), but not for OPA ($F_{1,12} = 1.04$, $p = 0.328$) or face-selective regions (FFA: $F_{1,11} = 1.28$, $p = 0.283$; OFA: $F_{1,10} = 0.01$ $p = 0.928$). In addition to a main effect of shape, a main effect of texture was found in PPA ($F_{1,12} = 8.21$, $p = 0.014$), but not the remaining regions (all Fs < 3.90; all ps ≥ 0.074). Significant shape-by-texture interactions were found in PPA ($F_{1,12} = 14.55$, $p = 0.002$), OPA ($F_{1,12} = 16.90$, $p = 0.001$), and FFA ($F_{1,11} = 8.00$, $p = 0.016$), but not the remaining regions (all Fs < 4.56; all ps ≥ 0.058). Next, we conducted planned pairwise comparisons to examine fMR-adaptation effects across conditions. Compared to the no-change condition, the same-shape-different-texture condition resulted in a release from adaptation for all regions (all ts ≥ 2.21; all ps ≤ 0.047), except OFA ($t_{10} = 1.79$, $p = 0.104$). Similarly, the different-shape-same-texture condition showed a release when compared to the no-change condition in all regions except FFA and OFA (PPA: $t_{12} = 5.61$, $p < 0.001$; RSC: $t_{11} = 2.69$, $p = 0.021$; OPA: $t_{12} = 3.40$, $p = 0.005$; LO: $t_{12} = 4.20$, $p = 0.001$; FFA: $t_{11} = 1.46$, $p = 0.173$, OFA: $t_{10} = 1.63$, $p = 0.135$).

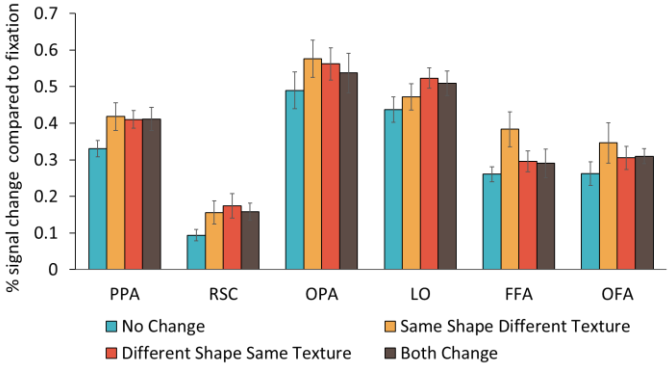

Figure 7. fMR-Adaptation response amplitudes for Experiment 3. Levels of adaptation across conditions (no change; same-shape-different-texture; different-shape-same-texture; both change) for each ROI. Data are represented as mean ± SEM.

Interactions between regions were analysed using a 2 (region) by 2 (shape) by 2 (texture) repeated-measures ANOVA. Significant region-by-shape interactions were found for PPA with FFA ($F_{1,11} = 7.16$, $p = 0.022$), for RSC with FFA ($F_{1,11} = 9.15$, $p = 0.012$), for OPA with FFA ($F_{1,11} = 7.01$, $p = 0.022$) and for LO with FFA ($F_{1,11} = 7.36$, $p = 0.020$). Significant region-by-texture interactions were found for PPA with RSC ($F_{1,11} = 4.91$, $p = 0.049$) and LO ($F_{1,12} = 6.80$, $p = 0.023$), and for RSC with FFA ($F_{1,11} = 4.90$, $p = 0.049$). No significant three-way interactions were found.

Finally, we should note that despite controlling for multiple low-level visual properties in our stimulus set (e.g., color, luminance, contrast), it is likely that some low-level differences persist. For example, changing scene shape in our stimuli, while keeping scene texture constant, produced low-level changes in monocular depth cues from surface properties. Thus, neural responses in the different-shape-same-texture condition likely result from a combination of differences in scene shape and differences in low-level depth cues from texture. We observed equivalent releases from adaptation across all 'change' conditions, however, and the adaptation results are consistent with the results from the previous two experiments. Together, this makes it unlikely that differences in low-level properties alone would account for a significant proportion of the neural response in the different-shape-same-texture condition, especially since numerous studies have found robust sensitivity of the PPA to variations in scene shape (e.g., Epstein and Kanwisher, 1998). Nevertheless, it is possible that changes to global scene shape and layout in previous studies were correlated with changes to low-level monocular depth cues from surface properties, as they were in the different-shape-same-texture condition here. Future studies are thus required to better understand the perceptual and neural impact of covariations in different scene features. We would contend, however, that the contribution of depth-from-texture

information to these results is not inconsistent with our general conclusion that both geometry and surface properties contribute to scene representation in PPA.

General Discussion

Our findings indicate that the scene-selective PPA responds equally in strength to the surface properties and geometry of a scene, and these features could not be discriminated from one another. In object and scene perception, evidence revealed preferential sensitivity to an object's shape within the object-selective LO, but the opposite was true for these object properties within the PPA when objects were presented in isolation (Experiment 1), suggesting that the PPA may be particularly sensitive to processing surface properties. An interaction between objects and scenes, however, modulated this relationship within the PPA such that the surface properties and geometry of objects were processed with similar sensitivity (i.e., similar univariate response amplitudes) when objects were presented within scenes (Experiment 2), despite demonstrating distinct underlying neural representations (i.e., multivariate spatial patterns of activation). Furthermore, varying scene geometry and surface properties independently of each other led to equivalent releases from adaptation in PPA (Experiment 3). Together with the findings that PPA can discriminate geometry and surface features of objects but not scenes, this evidence suggests that the processing of scene geometry and surface properties are weighted equally and may share similar underlying neuronal mechanisms in PPA, which jointly contribute to the construction of scene identity. The present findings also provide evidence that the processing of surface features in PPA may be mediated by domain-specific, rather than domain-general mechanisms, emphasizing the importance of this property for scene perception. Evidence across three experiments further highlights how these global scene properties are processed in the broader scene-processing network, albeit in ways which are

dissociated from each other, suggesting a complementary relationship between regions in this network for the purposes of perceiving and navigating the world around us.

Neural representation of geometry and surface properties

While extensive evidence has supported a primary role for the PPA in the encoding of spatial information (Epstein and Kanwisher, 1998; Epstein et al., 2003; Kravitz et al., 2011; Park et al., 2011), there is considerably less research investigating the encoding of surface information within this region. In anticipation of the present neural findings, however, behavioural results have implicated surface characteristics such as colour and texture as being instrumental in mediating early-stage processes responsible for successful scene recognition (Schyns and Oliva, 1994; Oliva and Schyns, 2000; Goffaux et al., 2005; Castelhano and Henderson, 2008). Similarly, PPA activation in an individual with profound visual form agnosia was modulated by the presence of appropriate scene colour (Steeves et al., 2004), and a case of topographical disorientation (wayfinding difficulties) related to landmark agnosia has revealed that geometry and surface properties may interact and jointly contribute to scene perception (Robin et al., 2017). Converging neuroimaging results have further shown that the surface and material properties of objects are processed in regions overlapping with scene-selective areas of parahippocampal cortex in humans (Peuskens et al., 2004; Cant and Goodale, 2007; 2011), and evidence has demonstrated sensitivity to both object and scene material properties in nonhuman primate visual cortex, including areas selective to scene processing (Kornblith et al., 2013; Goda et al., 2014). Surface properties may play a meaningful role in humans' ability to distinguish one scene from another. In the present study, by using strictly-controlled environments which allow for the direct comparison of visual features, we demonstrate that scene-selective cortex responds just as strongly to distinctive surface properties as it does to geometry. Together with previous

findings, these results raise a question of how multiple cues are used to construct the visual world around us across inherently complex and vastly different environments.

One account to explain how multiple cues are utilized in scene perception may include the role of diagnostic features and goal states (Lowe et al., 2016). In fact, an existing behavioural scene recognition framework centers on the notion of feature diagnosticity: the idea that specific visual cues are used for specific types of categorizations and an interaction between task demands and available visual information can explain how different cues are used to recognize scenes (Oliva and Schyns, 1997). Thus, for a complete view of scene understanding, it is necessary to account for both the contributions of diverse scene properties and differing observer goals (Lowe et al., 2016, Malcolm et al., 2016). Altering diagnostic aspects of a scene may in turn alter place information responsible for perceived novelty and the subsequent encoding of a scene in memory, consistent with findings suggesting the PPA is involved in encoding novel place information in memory (Epstein et al., 1999). Surface properties may represent the unique identity of a scene in much the same way that geometry does, yet this representation is dependent on the context in which the environment is perceived. The strongest empirical support for this account lies in the asymmetrical response of PPA to visual features across real-world scene environments. Building on the observation that natural, compared with manmade, environments typically contain large distinctive surface areas useful for recognition (Rao and Lohse, 1993; Loschky et al., 2010), evidence has shown that neural activity in PPA associated with attending to diagnostic surface properties in real-world scenes was higher in natural than manmade environments, while the opposite was found for geometry (Lowe et al., 2016). Here, we would argue that PPA may not necessarily be responsible for the selection of diagnostic information, but rather responds in accordance with top-down mechanisms to facilitate scene recognition.

An alternative explanation for our findings may lie in the sensitivity of PPA to low-level visual statistical changes, including those elicited by changes in both spatial structure and surface properties (e.g., monocular cues to depth). This view would deem PPA an area purely devoted to processing the low-level visual statistics of a scene. Real-world scenes contain a high degree of statistical regularity (Torralba and Oliva, 2003; Oliva and Torralba, 2001) and evidence has demonstrated that PPA, and scene-selective cortex more generally, is sensitive to processing low-level image features such as high spatial frequency content (Rajimehr et al., 2011; Watson et al., 2016). This may help to explain how regions distributed across scene-selective cortex show similar response amplitudes to changing features (both spatial and non-spatial). Yet given findings that (1) PPA plays a more direct role in humans' ability to categorize real-world scenes compared with areas of early visual cortex (Walther et al., 2011), (2) PPA exhibits higher activation for object-texture over object-shape and vice-versa in LO, and (3) the functional dissociations between PPA and control regions (FFA and OFA) observed here, we believe that appealing to the processing of low-level features alone is unlikely to explain the present series of results.

Interactions across object and scene perception

An intriguing finding from the present study concerns the modulation of activity in PPA by the presence of an object within a scene. Specifically, attention to object texture resulted in greater activation in PPA compared with attention to object shape when scenes and objects were presented separately (Experiment 1), but this difference was weakened and was no longer significant when scenes and objects were presented together (Experiment 2). Previous behavioural research has demonstrated an interaction between the processing of shared visual features (i.e., geometry and surface properties) across object and scene perception (Lowe et al.,

2015), and neuroimaging investigations have further highlighted a role for object-selective cortex in modulating activity within PPA (Mullin and Steeves, 2011; Rafique et al., 2015). These findings emphasize that feature perception in PPA is not static, but dynamic and dependent on the contextual relationship between the surrounding environment and the objects within it. The exact nature of the object-scene interaction observed in this study is unclear. Moreover, the general pattern of univariate results across Experiments 1 and 2 was qualitatively similar. As such, we hesitate to make strong conclusions regarding the nature of object-scene interactions from the present results. Factors such as ease of figure-ground segmentation, object size, location of the object within the scene, and congruent versus incongruent relationships between scene and object geometry or surface properties may all play a role in explaining object-scene interactions specifically and our results more generally. At the very least, the results of the present study represent a first step in establishing the finding that placing an object within a scene modulates object-related, but not scene-related activity in PPA. The degree to which this finding is explained by the various factors described above, and whether scene-related activity can be modulated by the presence of an object, are interesting empirical questions that deserve further study.

Interactions between geometry and surface properties

Visual features within our local visual environment can interact to support a common goal. For instance, surface properties (on both objects and scenes) may contribute to the perception of spatial information through cues provided from depth and contours (Torralba and Oliva, 2003), as well as edge information (Renninger and Malik, 2004), while geometric features may provide boundary information utilized in texture segmentation (Mohan and Nevatia, 1992). Support for the idea that geometry and surface properties are often used to accomplish a common

goal comes from the results of Experiment 3, where the processing of scene shape and scene texture was found to be non-additive. This suggests that these features are not represented independently in PPA. For active navigation, surface properties provide affordance-related information pertaining to how an observer should move through an environment, and boundaries provided from geometric cues are similarly critical for navigation (e.g., obstacle avoidance). Indeed, how could an individual efficiently navigate across a landscape without knowledge of the material differences between snow and ice, or sand and grass, and the most efficient path in space to move from one place (i.e., the origin) to another (i.e., the goal)? The results presented here suggest that neural representations of these features may not be spatially independent within the PPA, but interestingly, there is some evidence to suggest that these features may be temporally distinct. For instance, electrophysiological evidence has suggested differences in the processing of edge-based and surface information for the purposes of natural scene perception (Fu et al., 2016).

Interactions between neural and behavioral results

We observed consistent neural evidence that the representation of scene shape and scene texture are weighted equally in PPA (all 3 Experiments), yet we also found evidence that attention to these features elicited different behavioral responses (for Experiments 1 and 2, but not 3; see Supplementary Materials for all behavioral results). How can these differences be reconciled? First, behavioral differences in scene texture discriminations in Experiments 1 and 2 did not translate into greater activation for the scene texture condition (compared with the scene shape condition) in these Experiments, which suggests that the allocation of attention was not disproportionately unbalanced across these conditions. Indeed, if attentional demands had been greater in the scene texture condition, then we would have expected to find some evidence of

higher activation in this condition compared with the scene shape condition (i.e., Murray and Wojciulik, 2004), but we did not. Second, the pattern of neural results across all three experiments was very similar (e.g., no univariate difference in response amplitude, or multivariate difference in decoding accuracy, between the processing of scene shape vs. scene texture in PPA), despite differences in behavioral results (a difference in processing scene shape vs. scene texture in Experiments 1 and 2, but not 3). Third, previous research exploring the impact of task difficulty on neuronal representations has found evidence for a dissociation between manipulations of task difficulty and processing in ventral visual areas, including PPA (Xu et al., 2007).

However, given that previous research has revealed that the type of behavioral task used can impact neuronal representations in ventral occipito-temporal cortex (Harel et al., 2014; Lowe et al., 2016), as well as empirical evidence that neuronal activity is causally related to behavioural perception (e.g., Mégevand et al., 2014), it is not prudent to completely dismiss potential interactions between neural and behavioral results in the present study. Of course, at some level of representation, neuronal processing leads to changes in behavioral responses (and vice versa). Given this, and the importance of having an equal allocation of attentional resources when comparing different types of feature-selective processing, future research should examine the processing of shape and texture in both object and scene perception (both separately and in combination) in greater detail. Together with the present results, these new studies will increase our understanding of relationships between neural and behavioral processing in scene- and object-selective cortex.

Investigating the scene-processing network

Our results have also revealed functional dissociations within the scene-processing network, warranting consideration of these regions. Recent investigations have demonstrated the causal involvement of OPA in boundary perception (Julian et al., 2016), and have suggested that this region may be involved in representing the local elements of a scene, and first-person perspective for visually-guided navigation (Kamps et al., 2016). In the present study, we see evidence for dissociations between PPA and OPA across experiments: OPA showed greater sensitivity to the processing of scene geometry (Experiment 1), and was able to discriminate scene geometry from surface properties across the first two experiments. Together, these results suggest that scene geometry and texture are processed distinctly within OPA, and that this region may be particularly sensitive to the processing of scene geometry. The dissociable processing of these features with PPA may reflect the involvement of OPA in representing geometry for the purposes of local boundary perception, but further research is necessary to test these ideas directly.

While our results indicate that RSC was more similar to PPA overall, some dissociations and interactions across these regions highlight their differences. For example, across experiments, while both regions could not discriminate scene features, an inverse relationship was observed in object perception wherein PPA was able to discriminate objects features, yet this was not the case in RSC. How these results translate to the involvement of these regions in scene perception, recognition, and navigation warrant further investigation. Moreover, these results highlight the differences across these regions in feature-processing, and provide stronger evidence for the role of PPA in processing surface properties: While PPA may be more sensitive to processing the surface properties of an object over its shape, RSC does not show this sensitivity, but may process these features more generally. Together, these results emphasize the

importance of examining these three scene-selective regions in conjunction to explore their contributions to visual perception.

Investigating the face-processing network

Our results reveal an interesting dissociation within the face-processing network: While FFA could not discriminate between scene features in either of our first two experiments, we found the opposite was true for OFA. In addition, neither of these regions showed a significant release from adaptation to changes in scene geometry in Experiment 3, but FFA showed a release from adaptation to changes in scene texture. This finding in FFA was unexpected given the stimuli were scenes, yet previous work has indicated that FFA may show sensitivity to both shape and texture/material properties (Cant and Goodale, 2007; 2011), which is consistent with studies that have demonstrated sensitivity to shape (Merigan, 1996) and texture (Merigan, 2000) in regions along the fusiform gyrus. Moreover, the sensitivity to shape observed along the fusiform gyrus reported previously may have fallen outside of the functional borders of FFA (moving in to territory occupied by LOC, specifically, the posterior fusiform gyrus) and thus we did not observe this result in the present study.

With regard to OFA, this region has shown sensitivity to changes in the physical appearance of a face without changes in identity (Rotshtein et al., 2005), and has been proposed to act as the first stage of a cortical network for face processing by extracting feature and part-based information (Pitcher et al., 2011). Together, these findings may help to explain sensitivity of this region to visual feature processing. When we compare LO and OFA, we see both similarities (e.g., greater sensitivity to objects compared with scenes) and dissimilarities (e.g., release from adaptation for shape in LO but not OFA). Dissociations are expected given their functional roles in object- and face-processing, respectively, but their similarities could be due to

their physical proximity within the ventral stream, and their roles in extracting local visual feature information. Indeed, these regions may utilize similar information (i.e., shape and texture) for different purposes, but future research should explore these similarities and dissociations in greater detail.

Conclusions

In summary, our findings demonstrate that PPA responds just as strongly to changes in the surface properties of a scene as it does to changes in spatial structure. Moreover, neural responses to these scene features could not be discriminated from one another in PPA, despite significant discrimination of these features in object perception. We further observed greater responses to scene-texture compared with object-texture in PPA, regardless of whether scenes and objects were presented separately or together. Interestingly, while PPA showed greater sensitivity to processing texture compared with shape in object perception when objects were viewed independently of scenes, an interaction across object and scene perception altered this relationship, wherein we observed equal sensitivity to these properties when objects were perceived in the context of a scene. We conclude that texture processing in PPA (and scene-selective cortex more generally) may be mediated by domain-specific, rather than domain-general mechanisms, and the representations of scene geometry and surface properties are weighted equally in PPA, with their processing potentially mediated by similar underlying neuronal mechanisms.

References

Aminoff, E., Gronau, N., & Bar, M. (2007). The parahippocampal cortex mediates spatial and nonspatial associations. *Cerebral Cortex*, **17**, 1493-1503

Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *The Journal of Neuroscience*, **28**, 8539-8544

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive psychology*, **20**, 38-64

Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in a material world: how visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of Neurophysiology*, **102**, 3111-3118

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

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

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

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

Cant, J. S., & Xu, Y. (2017). The contribution of object shape and surface properties to object ensemble representation in anterior-medial ventral visual cortex. *Journal of Cognitive Neuroscience,29*, 398 – 412.

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on the perception of scene gist. *Journal of Experimental Psychology: Human perception and performance*, **34**, 660

Diana, R. A. (2016). Parahippocampal Cortex Processes the Nonspatial Context of an Event. *Cerebral Cortex*, bhw014

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

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

Duda, R.O., Hart, P.E., Stork, D.G. (1995) Pattern Classification, Second ed. Wiley Interscience, Chapter 3.

Epstein, R. A. (2014). Neural systems for visual scene recognition. *Scene vision: making sense of what we see. MIT Press, Cambridge, MA*, 105-134

Epstein, R., Graham, K. S., & Downing, P. E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron*, ***37***, 865-876

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: Recognition, navigation, or encoding? *Neuron*, **23**, 115-125

Epstein, R. A., Higgins, J. S., & Thompson-Schill, S. L. (2005). Learning places from views: variation in scene processing as a function of experience and navigational ability. *Journal of Cognitive Neuroscience*, **17**, 73-83

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

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, **2**, 189-210

Fu, Q., Liu, Y. J., Dienes, Z., Wu, J., Chen, W., & Fu, X. (2016). The role of edge-based and surface-based information in natural scene categorization: Evidence from behavior and event-related potentials. *Consciousness and Cognition*, *43*, 152-166.

Gallivan, J. P., Cant, J. S., Goodale, M. A., & Flanagan, J. R. (2014). Representation of object weight in human ventral visual cortex. *Current Biology*, **24**, 1866-1873

Gallivan, J. P., McLean, D. A., Flanagan, J. R., & Culham, J. C. (2013). Where one hand meets the other: limb-specific and action-dependent movement plans decoded from preparatory signals in single human frontoparietal brain areas. *The Journal of Neuroscience*, **33**, 1991-2008

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495-504.

Goda, N., Tachibana, A., Okazawa, G., & Komatsu, H. (2014). Representation of the material properties of objects in the visual cortex of nonhuman primates. *The Journal of Neuroscience*, **34**, 2660-2673

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence. *Visual Cognition*, **12**, 878-892

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

Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences*, **111**, E962-E971

Hsu, C. W., & Lin, C. J. (2002). sA comparison of methods for multiclass support vector machines. *Neural Networks, IEEE Transactions on*, *13*(2), 415-425

Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision research*, **47**, 3286-3297

Julian, J. B., Ryan, J., Hamilton, R. H., & Epstein, R. A. (2016). The occipital place area is causally involved in representing environmental boundaries during navigation. *Current Biology*, **26**, 1104-1109

Kamps, F. S., Julian, J. B., Kubilius, J., Kanwisher, N., & Dilks, D. D. (2016). The occipital place area represents the local elements of scenes. *NeuroImage*, **132**, 417-424

Kamps, F. S., Lall, V., & Dilks, D. D. (2016). The occipital place area represents first-person perspective motion information through scenes. *Cortex*, **83**, 17-26

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

Kornblith, S., Cheng, X., Ohayon, S., & Tsao, D. Y. (2013). A network for scene processing in the macaque temporal lobe. *Neuron*, **79**, 766-781

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

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

Loschky, L. C., Hansen, B. C., Sethi, A., & Pydimarri, T. N. (2010). The role of higher order image statistics in masking scene gist recognition. *Attention, Perception, & Psychophysics*, **72**, 427-444

Lowe, M. X., Ferber, S., & Cant, J. S. (2015). Processing context: Asymmetric interference of visual form and texture in object and scene interactions. *Vision research*, **117**, 34-40

Lowe, M. X., Gallivan, J. P., Ferber, S., & Cant, J. S. (2016). Feature diagnosticity and task context shape activity in human scene-selective cortex. *NeuroImage*, **125**, 681-692

Malcolm, G. L., Groen, I. I., & Baker, C. I. (2016). Making Sense of Real-World Scenes. *Trends in Cognitive Sciences*, **20**, 843-856

Mégevand, P., Groppe, D. M., Goldfinger, M. S., Hwang, S. T., Kingsley, P. B., Davidesco, I., & Mehta, A. D. (2014). Seeing scenes: topographic visual hallucinations evoked by direct electrical stimulation of the parahippocampal place area. Journal of Neuroscience, **34**(16), 5399-5405

Merigan, W.H. (1996). Basic visual capacities and shape discrimination after lesions of extrastriate area V4 in macaques. Visual Neuroscience, **13**, 51–60

Merigan, W.H. (2000). Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. Visual Neuroscience, **17**, 949 –958

Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage*, **53**, 103-118

Mohan, R., & Nevatia, R. (1992). Perceptual organization for scene segmentation and description. *IEEE transactions on pattern analysis and machine intelligence*, *14*(6), 616-635

Mullin, C. R., & Steeves, J. K. (2013). Consecutive TMS-fMRI reveals an inverse relationship in BOLD signal between object and scene processing. *The Journal of Neuroscience*, **33**, 19243-19249

Murray, S. O., & Wojciulik, E. (2004). Attention increases neural selectivity in the human lateral occipital complex. Nature neuroscience, **7**(1), 70-74.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive psychology*, **34**, 72-107

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive psychology*, **41**, 176-210

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, **42**, 145-175

Park, S., Brady, T. F., Greene, M. R., & Oliva, A. (2011). Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *The Journal of Neuroscience*, **31**, 1333-1340

Peuskens, H., Claeys, K. G., Todd, J. T., Norman, J. F., Van Hecke, P., & Orban, G. A. (2004). Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *Journal of cognitive neuroscience*, **16**, 665-682

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental brain research*, *209*(4), 481-493.

Rafique, S. A., Solomon-Harris, L. M., & Steeves, J. K. (2015). TMS to object cortex affects both object and scene remote networks while TMS to scene cortex only affects scene networks. *Neuropsychologia*, *79*, 86-96.

Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., & Tootell, R. B. (2011). The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biol*, **9**, e1000608

Rao, A. R., & Lohse, G. L. (1993). Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, **55**, 218-233

Renninger, L. W., & Malik, J. (2004). When is scene identification just texture recognition? *Vision research*, 44, 2301-2311

Robin, J., Lowe, M.X., Pishdadian, S., Rivest, J., Cant., J.S., & Moscovitch, M. (2017). Selective scene perception deficits in a case of topographical disorientation. *Cortex*

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature neuroscience*, *8*(1), 107-113.

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *Neuroimage*, **30**, 1088-1096

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological science*, **5**, 195-200

Steeves, J. K., Humphrey, G. K., Culham, J. C., Menon, R. S., Milner, A. D., & Goodale, M. A. (2004). Behavioral and neuroimaging evidence for a contribution of color and texture

information to scene classification in a patient with visual form agnosia. *Journal of Cognitive Neuroscience*, **16**, 955-965

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: computation in neural systems*, **14**, 391-412

Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *The Journal of Neuroscience*, **29**, 10573-10581

Watson, D. M., Hymers, M., Hartley, T., & Andrews, T. J. (2016). Patterns of neural response in scene-selective regions of the human brain are affected by low-level manipulations of spatial frequency. *NeuroImage*, *124*, 107-117

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior research methods*, **42**, 671-684

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

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

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