

Deconstructing Visual Scenes in Cortex: Gradients of Object and Spatial Layout Information

Assaf Harel, Dwight J. Kravitz and Chris I. Baker

Laboratory of Brain and Cognition, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA

Address correspondence to Assaf Harel. Email: assaf.harel@nih.gov.

Real-world visual scenes are complex cluttered, and heterogeneous stimuli engaging scene- and object-selective cortical regions including parahippocampal place area (PPA), retrosplenial complex (RSC), and lateral occipital complex (LOC). To understand the unique contribution of each region to distributed scene representations, we generated predictions based on a neuroanatomical framework adapted from monkey and tested them using minimal scenes in which we independently manipulated both spatial layout (open, closed, and gradient) and object content (furniture, e.g., bed, dresser). Commensurate with its strong connectivity with posterior parietal cortex, RSC evidenced strong spatial layout information but no object information, and its response was not even modulated by object presence. In contrast, LOC, which lies within the ventral visual pathway, contained strong object information but no background information. Finally, PPA, which is connected with both the dorsal and the ventral visual pathway, showed information about both objects and spatial backgrounds and was sensitive to the presence or absence of either. These results suggest that 1) LOC, PPA, and RSC have distinct representations, emphasizing different aspects of scenes, 2) the specific representations in each region are predictable from their patterns of connectivity, and 3) PPA combines both spatial layout and object information as predicted by connectivity.

Keywords: fMRI, LOC, PPA, RSC, scene recognition

Introduction

Real-world visual scenes are enormously rich stimuli containing many different sources of information including spatial layout, local objects, and semantic associations. Unsurprisingly, given this multidimensionality, functional magnetic resonance imaging (fMRI) studies have revealed a network of cortical regions engaged by visual scene processing including both scene- (parahippocampal place area, PPA; retrosplenial complex, RSC) (Aguirre et al. 1996; Maguire 2001; Epstein 2008; Park and Chun 2009) and object-selective (lateral occipital complex, LOC) (Malach et al. 1995; Grill-Spector et al. 2001; Kourtzi and Kanwisher 2001) areas. The critical question is how each region contributes to scene processing and, ultimately, how they interact. There are 2 major challenges for understanding the differential contributions of these regions to scene representation. First, the regions are located in anatomically distinct locations, and we have limited understanding of the connectivity between them. Second, the complexity of scenes makes it difficult to tease apart the factors underlying the response of a region.

To address the first challenge, we generated predictions from a recent anatomical framework developed in monkey (Kravitz, Saleem, et al. 2011). While the functionally defined LOC and PPA have been extensively studied in humans, the homologous regions in monkey are not firmly established

(but see Bell et al. 2009; Nasr et al. 2011). Assuming a rough anatomical homology, PPA likely corresponds to monkey parahippocampal gyrus, while LOC corresponds to more lateral regions of the ventral temporal cortex (Kravitz, Saleem, et al. 2011). In the monkey, both of these regions receive strong input from V4 (Ungerleider et al. 2008). However, the more medial regions (putatively corresponding to PPA) also receive input from posterior parietal cortex via a pathway including retrosplenial and posterior cingulate cortices (Kravitz, Saleem, et al. 2011), which are both thought to be within the functionally defined RSC in human (Epstein 2008) (To avoid confusion, we use RSC when discussing the functionally defined region human and retrosplenial cortex when discussing the anatomically defined region in monkey.). Importantly, unlike ventral temporal cortex, monkey retrosplenial cortex, which lies within the parietomedial temporal pathway, has little ventral pathway connectivity (Vogt and Pandya 1987; Kobayashi and Amaral 2003, 2007). Translating this knowledge of the connectivity back to humans, we hypothesized that human RSC should be sensitive primarily to spatial and not object information. In contrast, LOC, which lies within the ventral visual pathway, should have object but little spatial information. Finally, PPA, which likely receives input from both ventral visual and the parietomedial temporal pathways, should represent both spatial and object information.

To address the challenge of scene complexity, we systematically controlled the spatial layout and objects in artificial scenes (Fig. 1). Prior studies of visual scene processing have often used naturalistic photographic stimuli (e.g., Bar et al. 2008; Walther et al. 2009; Park et al. 2011), which have high ecological validity but whose features and information are very difficult to control. This complexity may partly explain conflicting results regarding the nature of representations in PPA, with some studies emphasizing spatial representations (Epstein and Kanwisher 1998; Epstein et al. 1999; Epstein and Ward 2010; Park and Chun 2009; Park et al. 2011), and others arguing for coding of scene category (Walther et al. 2009, 2011) or contextual associations (Bar and Aminoff 2003; Aminoff et al. 2007; Bar et al. 2008). Furthermore, PPA contains at least some information about isolated objects (MacEvoy and Epstein 2009), the extent of which is still unknown.

In a prior fMRI study, we used a data-driven approach to determine the type of information primarily represented by PPA. We analyzed the response across PPA to a highly diverse set of naturalistic real-world scene images from multiple semantic categories. The analyses revealed that scene representations in PPA were structured primarily by spatial factors, in particular spatial boundary (open and closed) and distance (near and far) (Kravitz, Peng, et al. 2011; see also Park et al. 2011). Here, we used a complementary approach, using highly controlled minimal scenes to systematically and directly test the

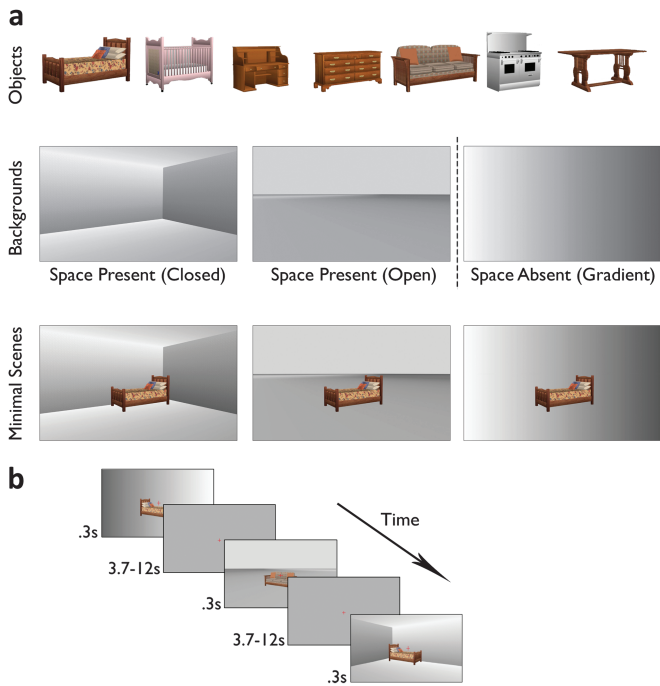


Figure 1. Stimuli and task. (a) Depiction of the stimuli used in the event-related runs. Each of the 7 objects pictured in the top row was superimposed on each of the 3 backgrounds in the second row. Two of the backgrounds implied a spatial layout (open and closed—space present), while one was matched for overall luminance but contained no spatial layout information (gradient—space absent). Images of the backgrounds only formed the object-absent condition (second row). The combined minimal scenes (third row) contained a single object and background, and every combination of the 2 was presented. This fully crossed design allowed the effect of background and object to be evaluated separately in the subsequent response magnitude and response pattern analyses. (b) Depiction of the sequence of events in a typical series of trials during the event-related runs. Each minimal scene was presented for 300 ms. To ensure participants maintained fixation, they performed a shape judgment task on the central fixation cross (see text for details). Intertrial intervals varied from 3700 to 12000 ms, and no response was required from the participants in this period. This variable delay was introduced to allow for efficient deconvolution of the responses.

relative importance of spatial layout and object information in PPA, RSC, and LOC. These minimal scenes contained 1 of 8 possible large-scale objects (furniture items including object absent) on either a spatial (open and closed) or nonspatial background (Fig. 1). This design allowed us to assess the contribution of the mere presence of objects and spaces and the contribution of specific objects and specific spatial backgrounds to both the response magnitude and the information contained in the multivoxel response pattern in PPA, RSC, and LOC.

Consistent with our hypotheses, RSC primarily contained spatial layout information and was not even sensitive to the presence or absence of an object. In contrast, LOC showed the opposite pattern, primarily containing information about object identity and none about spatial layout. Finally, PPA responded more strongly in the presence of objects and contained significant information about both objects and spatial layout. These results suggest that 1) LOC, PPA, and RSC have distinct representations, emphasizing different aspects of visual scenes, 2) the specific representations in each region are predictable from their unique patterns of connectivity (Kravitz, Saleem, et al. 2011), and 3) as predicted by its connectivity, PPA combines both spatial layout and object information, at least for the large-scale objects we tested.

Materials and Methods

Participants

Twenty participants (13 female) aged 19–36 years participated in the experiment. All participants had normal or corrected-to-normal vision and gave written informed consent. The consent and protocol were approved by the National Institutes of Health Institutional Review Board.

Stimuli

The approach taken in this study is complementary to our previous study on scene representations using a large sample of complex naturalistic stimuli (Kravitz, Peng, et al. 2011). To systematically control the information contained in individual scenes, we created 48 minimal scenes using commercial interior design software (Home Design Studio version 14.1.3; Punch! Software). The scenes were images of 1 of 7 different objects superimposed on 1 of 3 types of grayscale background to form minimal scenes (Fig. 1). The 7 objects were common furniture items (bed, crib, desk, dresser, sofa, stove, and table) that were positioned at the center of the image while retaining realistic depth of field. Furniture items were chosen because they can be realistically embedded in a spatial environment (unlike an object “floating” in mid air). Furthermore, furniture items provide strong navigational affordances and elicit strong responses throughout the cortex, including parahippocampal cortex (e.g., Mullally and Maguire 2011). We also included an object-absent condition, in which only the background was present, enabling us to see whether empty spaces are sufficient to drive the response of scene-selective cortex as long as scene geometry is preserved (Epstein and Kanwisher 1998). Note that these object-absent conditions differ from the object-present conditions on multiple dimensions (e.g., achromatic vs. chromatic). In this respect, “object absence” is an operational term used to group these factors together. Based on prior studies highlighting the importance of spatial information in scene representation and in driving the response of parahippocampal and retrosplenial cortices in particular (Aguirre et al. 1996; Epstein 2008; Greene and Oliva 2009; Kravitz, Peng, et al. 2011; Mullally and Maguire 2011), we generated 3 different backgrounds consisting of 2 spatial and 1 nonspatial background. The 2 spatial backgrounds were an empty room (closed scene) or a receding horizon (open scene). The spatial dimension of expanse was chosen based on the results of Kravitz, Peng, et al. (2011), which found that the difference between open and closed scenes was strongly represented in the response of PPA. The nonspatial background (space absent) was a luminance gradient. All backgrounds were equated in their global mean luminance. Finally, since the spatial backgrounds were asymmetric, each one of the 24 scenes (i.e., the full crossing of objects and backgrounds) was mirror reversed to control for potential visual field differences. The manner in which these minimal scenes were constructed allowed us to directly test the differential roles that spatial background and objects play in the representational structure of LOC, PPA, and RSC.

fMRI Localizer Experiment

Three independent block-design scans were collected in each participant to localize LOC, PPA, and RSC (Supplementary Fig. S1). These independent scans were used to avoid any possibility of introducing circularity in the analysis of the later event-related data (Kriegeskorte et al. 2009). Each of these scans was an on/off design with alternating blocks of different stimuli presented while participants performed a one-back task. PPA and RSC were localized by the contrast of scenes and faces, while LOC was defined by the contrast of objects versus retinotopically matched scrambled objects (Kravitz et al. 2010). The contrast of scenes and faces also allowed us to identify the face-selective fusiform face area (FFA). Note that neither the contrast of scenes versus faces, nor the contrast of objects versus scrambled, determines what information from our scene stimuli is represented by the resulting regions of interest (ROIs) response pattern. Scene and object images were grayscale photographs ($5^\circ \times 3^\circ$). There were 2 further on/off scans to help more precisely define the ROIs. There were 2 further on/off scans to help more precisely define the ROIs. First,

a simple contrast of central and peripheral flickering (8 Hz) checkerboards sized less or more than 5° of visual angle, respectively, allowed us to identify and exclude retinotopic voxels from LOC, PPA, and RSC (In LOC, there was greater overlap with central early visual cortex [cEVC] [16%/5% L/R] than with peripheral early visual cortex [pEVC] [6%/1% L/R]. In PPA, there was greater overlap with pEVC [18%/20% L/R] than with cEVC [6%/9% L/R]. No voxels in left or right RSC were found to overlap with cEVC or pEVC.). A second run included the direct contrast of objects and scenes, allowing us to adjudicate whether to place a voxel within LOC or PPA when it was significant in both the contrast of scenes against faces and objects against scrambled.

Event-Related fMRI Experiment

On each trial, within each of 6 event-related runs, a minimal scene was presented for 300 ms, followed by variable length fixation period for 3.7–12 s. Stimulus presentation length was short enough to minimize eye movements. While brief, people are able to extract a great deal of information from scenes at this duration, as demonstrated extensively in prior behavioral work (Potter and Levy 1969; Potter 1976; Thorpe et al. 1996; Joubert et al. 2007). Scenes were presented twice every run in a randomized order, producing 96 trials per run. A 16-s fixation block was added to the beginning and the end of each run, culminating in a total run length of 8 min and 48 s. To ensure participants maintained fixation, they performed a shape judgment task on the central fixation cross. Specifically, simultaneous with the presentation of each scene, one arm of the fixation cross grew slightly longer. Participants reported whether the horizontal or vertical arm lengthened via a button press. Which arm grew was counterbalanced across scenes between runs, such that both arms grew equally often with each scene. There was no change in the fixation cross in the intertrial intervals, and no response was required from the participants in these periods. We used this task, which was orthogonal to scene identity, to investigate scene representations without introducing any confounds or feedback effects produced by task on the specific scene stimuli.

fMRI Scanning Parameters

Participants were scanned on a research dedicated GE 3-T Signa scanner located in the Clinical Research Center on the National Institutes of Health campus in Bethesda. Partial volumes of the temporal and occipital cortices were acquired using an 8-channel head coil (19 slices, 2 × 2 × 3 mm, 0.3 mm interslice gap, time repetition [TR] = 2 s, time echo = 30 ms, matrix size = 96 × 96, field of view = 192 mm). In all scans, oblique slices were oriented approximately parallel to the base of the temporal lobe. Six event-related runs (263 TRs each) and 3 localizer scans (144 TRs) were acquired in each session.

fMRI Preprocessing

Data were analyzed using the AFNI software package (<http://afni.nimh.nih.gov/afni>) and custom Matlab (2007; Mathworks, Natick, MA) scripts. Prior to statistical analysis, all of the images for each participant were motion-corrected to the first image of their first run. Following motion correction, the event-related and the localizer runs were smoothed with a 5-mm full-width at half-maximum Gaussian kernel.

fMRI Statistical Analysis

Functional ROIs were created for each participant from the localizer runs. Here, we focus on PPA, LOC, and RSC, and data from early visual cortex and FFA are presented in Supplementary Data (see also Supplementary Figs S2 and S3). Significance maps of the brain were computed by performing a correlation analysis between the assumed hemodynamic response function and the activation time courses thresholded at $P < 0.0001$ (uncorrected). ROIs were generated from these maps by taking the contiguous clusters of voxels that exceeded threshold and occupied the appropriate anatomical location based on previous work (Epstein and Higgins 2007; Schwarzlose et al. 2008). To more precisely define LOC, PPA, and RSC, we excluded retinotopic voxels from the ROIs and adjudicated between overlapping voxels by assessing the voxel's selectivity in an independent scene-object

localizer (see above). The ROIs were identified in all participants, except for one participant, in which RSC could not be localized. We conducted a standard general linear model (GLM) using the AFNI software package to deconvolve the event-related responses for each voxel within the predefined ROI. It is important to stress that the functional localizer scans were independent from the event-related experimental scans (i.e., separate scans using stimuli that were visually and conceptually distinctive from the stimuli used in the main experiments) to avoid circularity in our analysis (Kriegeskorte et al. 2009).

Beta parameters were extracted for all individual voxels within a given ROI and used as estimates of the magnitude of the ROI's response to the different conditions. The beta estimates were averaged across the ROI and the ten splits of the data (see below) and were later subjected to omnibus analyses of variance (ANOVAs) (all reported P values throughout the manuscript are Greenhouse-Geisser corrected). Separate analyses equating ROI size within each participant did not show qualitative differences in the pattern of results.

Multivoxel response patterns across each ROI from the 6 event-related runs were analyzed using an iterative version of the split-half analysis method (Haxby et al. 2001; Kravitz et al. 2010; Kravitz, Peng, et al. 2011). Specifically, the 6 runs were divided into 2 separate data sets of 3 runs in all 10 possible ways (${}_6C_3/2$). For each half of the data in each of the 10 splits, significance maps were created by performing t -tests between each condition and baseline. In each voxel, the mean t value across conditions was independently removed from each half of the data as in prior studies (Haxby et al. 2001). Finally, the normalized t values for each condition were then extracted from the voxels within each ROI and cross-correlated across the halves of each split. Correlation values were averaged across the 10 splits, yielding similarity matrices that represent the similarity in the spatial pattern of response across the ROI between each pair of conditions. We considered the 2 viewpoints separately and then averaged across the resulting similarity matrices, as a separate analysis of the effect of viewpoint showed no evidence for any effect of viewpoint. This resulted in a single 24 × 24 (8 object conditions × 3 backgrounds) similarity matrix for each ROI, wherein each data point in the matrix represented the correlation of the pattern of response for a pair of scenes across the 2 halves of the data. These correlation values were either “within” a given condition, reflecting the consistency of response across splits of the data (e.g., bed on the closed background vs. bed on the closed background), or “between” a pair of different conditions (e.g., bed on the closed background vs. desk on the open background), reflecting the similarity between the patterns of response to the 2 conditions. Note that the main advantage of our approach (and other multivariate approaches) relative to the standard univariate GLM approach is that it does not involve averaging across voxels within an ROI, thus preserving fine patterns of activity which may not be otherwise visible (for further discussion, see Kriegeskorte et al. 2008; Mur et al. 2009).

Decoding Analyses

To assess whether the response pattern of a given ROI discriminates between the different conditions, we compared the within-condition and the between-condition correlations (see also Haxby et al. 2001; Chan et al. 2010; Kravitz et al. 2010). The ability of an ROI to discriminate a particular condition from the other conditions was indexed by calculating the difference score between the within-condition correlation and the mean of the between-condition correlations (Supplementary Figs S4 and S5). Hence, a discrimination index that is significantly greater than zero indicates the region can reliably decode a specific condition from its counterparts.

Object Information

Our primary measure of object information considered decoding of objects “across” backgrounds as the most rigorous test of object information. Simply put, we considered similarities between objects across backgrounds only, excluding any potential confounds due to low-level image properties available in the backgrounds themselves. To compute the indices, we first averaged the full similarity matrices to create 8 × 8 object matrices, where each value represented the

correlation between a pair of objects across backgrounds (Supplementary Fig. S4). To compute the within-condition object correlations (main diagonal), we averaged all the correlations between an object and itself on different backgrounds. The between-condition object correlations (off-diagonals) were the average of the correlations between that object and other objects on “different” backgrounds. By considering only the correlations between objects on different backgrounds, we hold constant the change in background in both the within-condition and between-condition object correlations.

To establish discrimination indices for object absence decoding, we subtracted the correlation between the object-absent and object-present conditions from the object-absent diagonal. To establish discrimination indices for object identity decoding, we first excluded all correlations involving the object-absent condition and then subtracted the between-condition object correlations from the within-condition object correlations. The discrimination indices resulting from the subtraction of these values represent the amount of object information available across backgrounds.

In addition, we also calculated object identity information within each background separately. We selected from the full matrices, the three 7×7 object matrices (excluding the object-absent condition) for each background where each point represented the correlation between a pair of objects on the same background. To compute the within-condition object correlations (main diagonal), we took the correlation between an object and itself on the same background. The between-condition object correlations (off-diagonal) were the average of the correlations between that object and the other objects on the same background. By only considering the correlations between objects on the same background, we hold background constant in both the within-condition and between-condition object correlations. The resulting discrimination indices represent the amount of object information available present on a “particular” background. To create the average indices for each one of the spatial backgrounds, we averaged across the indices for the open and closed background.

Spatial Background Information

Our measure of spatial background information considered decoding across objects only, so that background information was not confounded with object information. To compute the indices, we first averaged the full similarity matrices into 3×3 background matrices (Supplementary Fig. S5). The within-condition background correlations (main diagonal) were averages of the correlations between different object conditions on the same background. The between-condition background correlations (off-diagonal) were created by averaging the correlations between different objects on different backgrounds. By considering only the correlations between different objects, we hold a change in object constant in both the within-condition and between-condition background correlations. The resulting discrimination indices capture the amount of background information available across objects. To create the average indices for the spatial background, we averaged the indices for the open and closed backgrounds.

Structure of Representations

The decoding analyses were designed to test our specific hypotheses about the information contained within different ROIs. Complementary to these analyses, we also directly compared the structure of representations within each ROI in a manner agnostic to the experimental conditions. The correlation analysis described above yields a similarity matrix for each ROI reflecting the similarity between every pair of scenes in the pattern of response. That similarity matrix represents the structure of the scene representations within that ROI. However, the decoding analyses described above consider only the correlation between a given scene and itself (diagonal value in the similarity matrix) relative to the average correlation with the other scenes and not the relative correlations of the other scenes (off-diagonal values) in detail. In other words, the presence of discrimination in 2 ROIs does not necessarily mean that the structure of representations is similar. Therefore, in addition to the response pattern analysis, we systematically compared the structure of representations across ROIs. Specifically, in each participant, we cross-correlated the full similarity matrices (with the diagonal values excluded) between every

possible pair of the 3 ROIs. The full similarity matrices (excluding the space- and object-absent conditions) capture the structure of each region’s representation across the set of scenes.

Results

Twenty-four minimal scenes ($20^\circ \times 15^\circ$) composed of a single object (or no object) on a background were presented in a rapid event-related fMRI experiment. The objects were all from the same category (furniture: bed, crib, desk, dresser, sofa, stove, table, and object absent) and were superimposed over the center of 1 of 3 different backgrounds, 1 nonspatial (space absent: luminance gradient), and 2 with a spatial layout (space present: open and closed) (Fig. 1), which previous work had suggested would be strongly decoded in PPA (Kravitz, Peng, et al. 2011; Park et al. 2011). The design of these minimal scenes allowed us to test directly and independently the role that spatial layout and object information play in the representations contained in LOC, PPA, and RSC.

We independently localized LOC, PPA, and RSC bilaterally and examined their responses during the event-related experiment (see Materials and Methods). To determine the relative role that spatial backgrounds and objects play in driving the response of the regions, 2 main types of analysis were performed. First, a standard response magnitude analysis was conducted to assess the effects of the presence or absence of spatial and object information. Second, a multivariate response pattern analysis was conducted to reveal the fine structure of background and object representations within PPA, RSC, and LOC.

Presence and Absence of Spatial Layout and Objects

To establish the effects of the presence of each type of information, we averaged the response to the 7 objects (object present) and the 2 spatial backgrounds (space present) and compared them with the response to the object-absent (background only) and space-absent (gradient background) conditions. Differential effects of object and space presence were observed across ROIs. Response magnitudes were entered into a 3-way ANOVA with ROI (LOC/PPA/RSC), space presence (present/absent), and object presence (present/absent) as factors. This ANOVA revealed highly significant interactions between ROI and object presence ($F_{2,36} = 38.76$, $P < 0.0001$) and between ROI and space presence ($F_{2,36} = 11.55$, $P = 0.001$) indicating that the effects of both space and object presence differed between the ROIs. The 3-way interaction of ROI \times space presence \times object presence was not significant ($F_{2,36} = 1.56$), and therefore, we next analyzed the effect of space and object information separately.

Gradient of Object Information across Regions

In both LOC and PPA, but not RSC, responses were higher when an object was present than absent (Fig. 2a). Having already established the interaction between ROI and object presence, we compared response magnitudes in each ROI when an object was present or absent. These difference scores were significantly greater than zero in both LOC and PPA ($t_{19} = 7.67$, $P < 0.001$; $t_{19} = 7.69$, $P < 0.001$, respectively) but not in RSC ($t_{18} = 0.13$, $P = 0.90$). Furthermore, the object presence effect was stronger in LOC than in either PPA ($t_{19} = 2.76$, $P < 0.05$) or RSC ($t_{18} = 8.34$, $P < 0.001$), with significantly stronger response modulation in PPA compared with that observed in RSC ($t_{18} = 7.57$, $P < 0.001$). These results establish a gradient

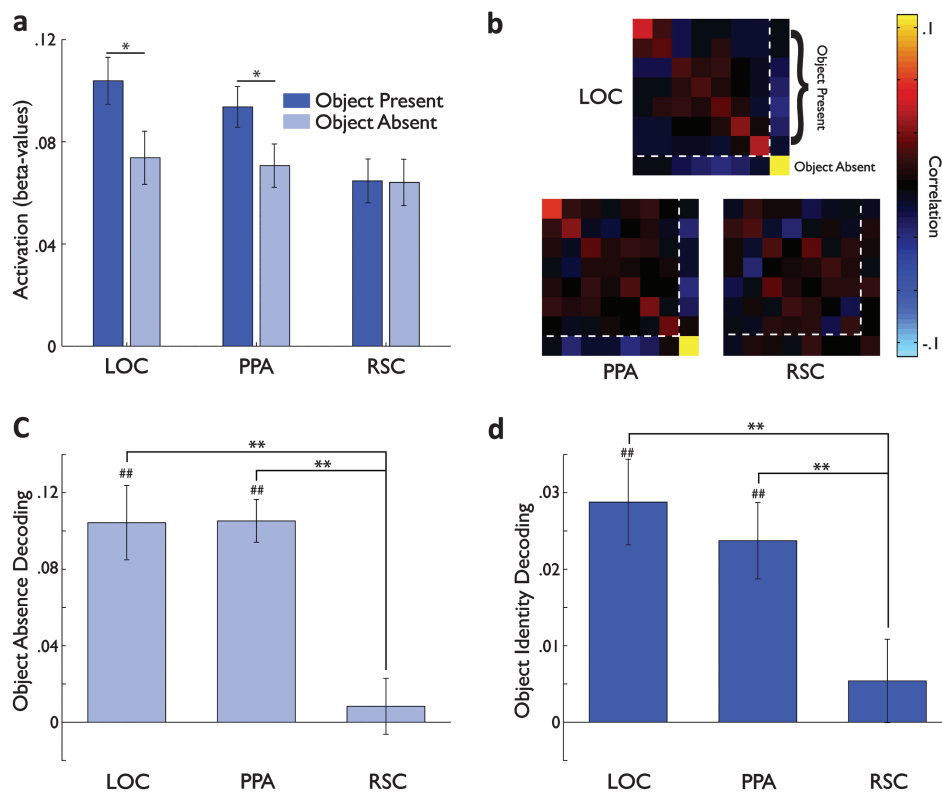


Figure 2. Object information across background. (a) Average response magnitude across the ROIs when an object was presented (dark blue) and when the object was absent (light blue) and only a background was presented. Note the effect of object presence in LOC and PPA (larger in LOC than in PPA) and the complete lack of any effect in RSC. (b) The similarity matrices for each ROI averaged by object (see Materials and Methods). The main diagonal from the top left to bottom right shows the within-object correlation and represents the consistency of the response pattern to a particular object across backgrounds. The off-diagonal values are the correlations between different objects across backgrounds. The order of the objects is the same as in Figure 1a. The first 7 conditions are the object-present conditions (same order as in Fig. 1a) and are bounded by the dotted white lines. The object-absent condition is the eighth and final condition in each plot. Note the strong diagonal value for the object-absent condition in both LOC and PPA but not in RSC. (c) Object absence decoding in LOC, PPA, and RSC. Any number significantly greater than zero indicates that object presence could be decoded in the region. Note the significant decoding in LOC and PPA but not in RSC. (d) Object identity decoding in LOC, PPA, and RSC. Any number significantly greater than zero indicates that object identity (excluding the object-absent condition) could be decoded in the ROI. Note the significant decoding in LOC and PPA but not in RSC and the significant differences in decoding between every pair of ROIs. All error bars in this and every other plot indicate the between-subjects standard error. ## $P < 0.001$ difference from zero; # $P < 0.05$ difference from zero; ** $P < 0.001$ difference between ROIs; and * $P < 0.05$ difference between ROIs.

of sensitivity to the presence of objects across LOC, PPA, and RSC with greatest modulation in LOC, consistent with connectivity of the different regions with the ventral visual pathway.

Having established this gradient in response magnitude, we next examined the object information available in the multivoxel response patterns. We considered both decoding of object absence and decoding of object identity. To avoid confounding object and background information, we conducted these analyses across-background and within-background separately. We begin assessing object information across-background as the most stringent test of object information.

To quantify object decoding across background, the full similarity matrices (see Materials and Methods) for each region were averaged by object to produce an 8 by 8 matrix (Fig. 2b). Next, to quantify whether the absence of an object could be decoded from the pattern of response, we computed a discrimination index for the object-absent condition by subtracting from the correlation between the object-absent condition and itself (Fig. 2b; last diagonal value) and the average correlation between it and the object-present conditions. These object absence discrimination indices were much higher in LOC and PPA than in RSC (Fig. 2c). A one-way ANOVA with ROI (LOC, PPA, and RSC) as a factor revealed a significant main effect of

ROI ($F_{2,36} = 16.02$, $P < 0.001$). Planned comparisons revealed that decoding of object absence was highly significant in LOC ($t_{19} = 5.50$, $P < 0.001$) and PPA ($t_{19} = 9.65$, $P < 0.001$) but not in RSC ($t_{18} = 0.57$). Furthermore, relative to RSC, decoding of object absence was stronger in both LOC ($t_{18} = 4.93$, $P < 0.001$) and PPA ($t_{18} = 5.31$, $P < 0.001$). These results show that even using the more sensitive measure of pattern analysis (Kriegeskorte et al. 2008; Mur et al. 2009), RSC showed no evidence of information about the presence or absence of an object despite the large differences between the stimuli and the multiple dimensions on which they differ.

To quantify object identity decoding, we computed average object identity discrimination indices (excluding the object-absent condition) in each ROI (Fig. 2d). Specifically, for each object in each ROI, we subtracted the average between-condition object correlations from the within-condition object correlations and then averaged these values across objects. Object identity decoding was much stronger in LOC and PPA than in RSC. A one-way ANOVA with ROI (LOC, PPA, and RSC) as a factor revealed a significant main effect of ROI ($F_{2,36} = 14.16$, $P < 0.001$). Planned comparisons showed that object identity decoding was highly significant in LOC ($t_{19} = 7.10$, $P < 0.001$) and PPA ($t_{19} = 7.13$, $P < 0.001$) but not in RSC ($t_{18} = 1.06$, $P = 0.30$). Complementing the object presence effects, object

identity decoding was stronger in both LOC ($t_{18} = 4.36$, $P < 0.001$) and PPA ($t_{18} = 3.75$, $P = 0.001$) than in RSC.

These results demonstrate that object presence and object identity can be decoded across backgrounds in both LOC and PPA but not in RSC. We next evaluated the amount of object information available within each background separately and compared object discrimination indices in the presence and absence of spatial layout information (Fig. 3). A two-way ANOVA with ROI (LOC, PPA, and RSC) and spatial background (present and absent) as factors revealed a significant main effect of ROI ($F_{2,36} = 6.11$, $P < 0.05$), reflecting the same gradient of object information observed across backgrounds: Object identity decoding within background was stronger in LOC than in either PPA ($t_{19} = 2.15$, $P < 0.05$) or RSC ($t_{18} = 2.72$, $P < 0.05$) and was stronger in PPA than in RSC ($t_{18} = 2.22$, $P < 0.05$). Consistent with the object identity decoding across backgrounds, only LOC ($t_{19} = 5.34$, $P < 0.001$) and PPA ($t_{19} = 3.82$, $P < 0.001$) evidenced significant decoding, with none evident in RSC ($t_{18} = -0.257$, $P = 0.40$). However, the interaction between ROI and background did not reach significance ($F_{2,36} = 2.05$, $P = 0.14$). Given that in all analyses, RSC showed no modulation by objects, not even by the presence or absence of objects, we also conducted a direct comparison of object information in PPA and LOC. This analysis revealed a significant interaction between ROI and background ($F_{1,19} = 5.51$, $P < 0.05$). A series of two-tailed t -tests revealed that LOC and PPA differ in their object identity decoding as a function of the type of background against which the objects are presented: Object decoding was stronger in LOC than in PPA in the absence of a spatial background ($t_{19} = 3.39$, $P < 0.01$) but equivalent in the presence of a spatial background ($t_{19} = 0.53$). There was also a trend for better object decoding in the absence of spatial layout within LOC ($t_{19} = 1.64$, $P = 0.11$). These results suggest that object decoding in LOC is negatively impacted by the presence of spatial background information, while object decoding in PPA is not.

In sum, our analyses of object information show a clear distinction between LOC on the one hand and RSC on the other, with PPA in between. Specifically, while LOC was modulated by

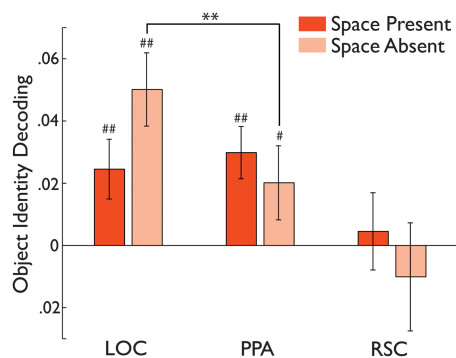


Figure 3. Object information within background. Average object identity decoding within the space-present (dark red) and space-absent (light red) backgrounds (see Materials and Methods). Note that average decoding evidenced a gradient across the ROIs with stronger decoding in LOC than in PPA or RSC and stronger decoding in PPA than in RSC (all $P < 0.05$). Note also the different patterns of decoding across the space-present and space-absent backgrounds in LOC and PPA resulting in a significant interaction between them ($P < 0.05$). All plotting conventions are the same as in Figure 3. # $P < 0.01$ difference from zero; ## $P < 0.05$ difference from zero; and ** $P < 0.01$ difference between ROIs.

the presence or absence of an object and contained information about object identity, RSC showed no evidence for any object information. PPA was very similar in the overall pattern of results to LOC but showed reduced object sensitivity relative to LOC. This gradient in object information—from RSC to PPA to LOC—is consistent with the direct connectivity of LOC and PPA with the ventral visual pathway, whereas RSC has limited if any direct connectivity.

Gradient of Spatial Layout Information across Regions

To understand the source of the interaction between ROI and space presence (see above), we compared response magnitudes in each ROI when a spatial background was present or absent (Fig. 4a). These difference scores were significantly greater than zero in both PPA and RSC ($t_{19} = 2.47$, $P < 0.05$; $t_{18} = 1.75$, $P < 0.05$, respectively) but not LOC ($t_{19} = -0.49$, $P = 0.31$). Thus, both PPA and RSC show an increased response to the presence of spatial layout information, while LOC does not.

Information about background was also reflected in the spatial pattern of response. We performed the most stringent test of background information by considering decoding across objects (i.e., avoiding conditions that share the same object). The full similarity matrices for each ROI (excluding the within-condition object correlations) were averaged by background to produce a 3 by 3 matrix representing background information across objects (Fig. 4b).

First, to establish whether the absence of a spatial background could be decoded from the patterns of response, we calculated discrimination indices for the gradient background (Fig. 4c, see Materials and Methods). A one-way ANOVA with ROI (LOC, PPA, and RSC) as a factor revealed no significant effect of ROI ($F_{2,36} = 0.48$). However, a series of planned comparisons revealed that space absence could be decoded in PPA only ($t_{19} = 2.42$, $P < 0.05$) and not in either LOC ($t_{19} = 0.42$) or RSC ($t_{18} = 0.71$). These results indicate that while PPA had a consistent response to the nonspatial background, neither RSC nor LOC produced reliable responses. In LOC, this likely reflects its general insensitivity to background (see below), while in RSC, this might reflect insensitivity to backgrounds that do not contain any spatial layout information.

To determine whether the 2 spatial backgrounds could be decoded from the patterns of response, we next calculated discrimination indices between the open and closed backgrounds. Open and closed backgrounds were strongly decoded in PPA (Fig. 4d), replicating our previous findings showing grouping by spatial expanse (Kravitz, Peng, et al. 2011). A series of planned comparisons revealed that the spatial backgrounds could be decoded from each other in PPA ($t_{19} = 3.21$, $P < 0.01$) and RSC ($t_{18} = 3.43$, $P < 0.01$), but not in LOC ($t_{19} = 0.75$). These indices were higher for PPA and RSC than LOC (Fig. 4d), though a one-way ANOVA with ROI (LOC, PPA, and RSC) as a factor did not reach significance ($F_{2,36} = 2.06$, $P = 0.15$).

In sum, our analyses of background information show that while PPA and RSC are modulated by the presence of spatial layout and show evidence of spatial background decoding, LOC shows no modulation of its response nor any decoding by background.

Opposite Gradients of Spatial and Object Information across ROIs

Notably, the gradient of object identity information across LOC, PPA, and RSC was in exactly the opposite direction from that

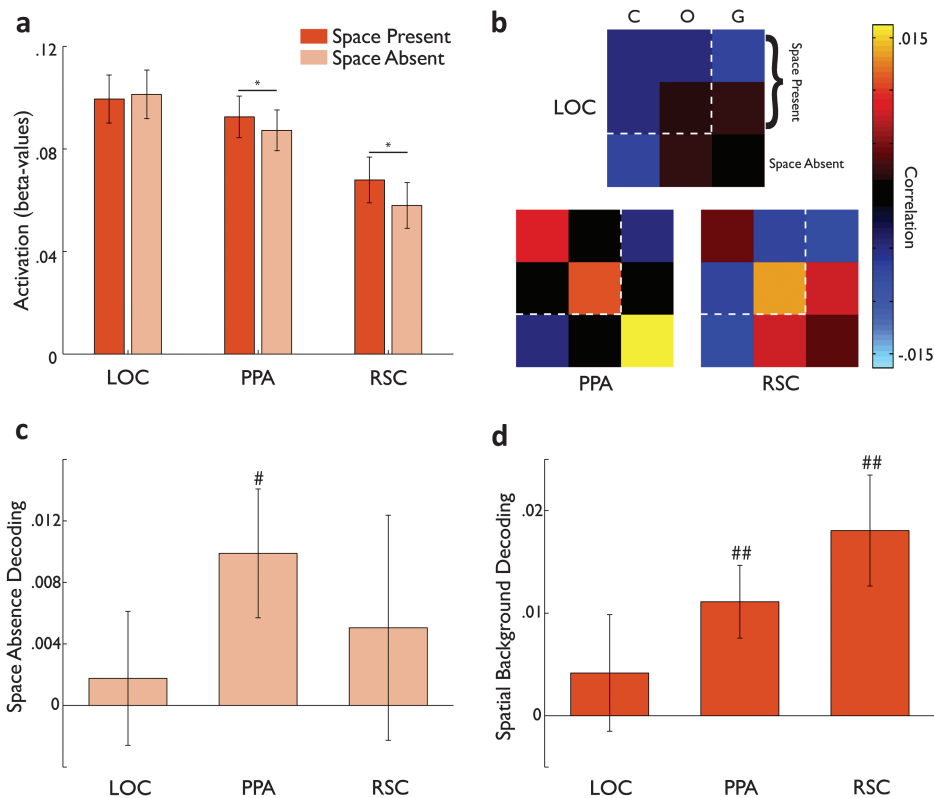


Figure 4. Background information across objects. (a) Average response magnitude across the ROIs for the space-present and space-absent backgrounds. Note the different patterns across backgrounds in LOC compared with PPA and RSC. (b) The similarity matrices for LOC, PPA, and RSC averaged by background (see Materials and Methods). The main diagonal shows the within-background correlation and represents the consistency of the response pattern to a particular background across objects. The off-diagonal values show the average correlation between different objects and different backgrounds. The order of the backgrounds is as in Figure 4a. The first 2 conditions are the space-present backgrounds and are bounded by the dotted white lines. The last condition is the gradient background, which did not contain any spatial layout information (space absent). Note the strong diagonals in PPA and RSC but not in LOC. (c) Space absence decoding in LOC, PPA, and RSC. Note that only PPA evidenced any coding of the absence of spatial layout information. (d) Spatial background decoding in LOC, PPA, and RSC. Note the significant decoding in PPA and RSC but not in LOC. ## $P < 0.01$ difference from zero; # $P < 0.05$ difference from zero; and * $P < 0.05$ difference between conditions.

observed for spatial background information (Figs 2d and 4d, respectively). To formally assess these 2 opposing gradients of representation, we entered the object identity and spatial background discrimination indices into a two-way ANOVA with ROI (LOC, PPA, and RSC) and information type (object and spatial background) as factors, revealing a highly significant interaction ($F_{2,36} = 15.00$, $P < 0.001$). However, the fact that each information type had different numbers of conditions might have affected the absolute magnitude of decoding within a given ROI, potentially limiting the ability to interpret object versus spatial layout differences. Therefore, in order to further test the relative amounts of spatial and object information in each ROI, we transformed the data into rankings across ROIs for each information type (Fig. 5). We then performed pairwise Wilcoxon signed-ranks tests between the information types in each ROI. The relative ranking combined with the nonparametric test allowed us to conservatively compare the 2 sources of information directly, revealing significantly higher object than background decoding rankings in LOC ($z_{19} = 2.28$, $P < 0.05$) and higher background than object decoding rankings in RSC ($z_{19} = 2.46$, $P < 0.05$) with equivalent object and spatial background rankings in PPA ($z_{19} = 0.00$).

The opposite gradients indicate that the variation in information across the ROIs was in line with their connectivity. RSC showed strong background information and very weak

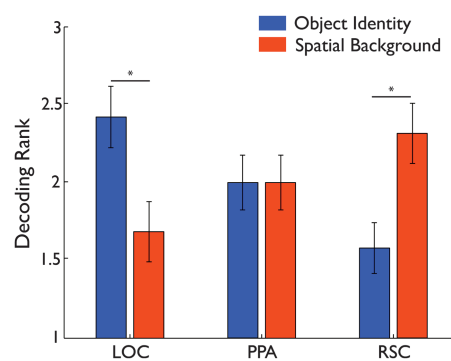


Figure 5. Opposite gradients of object and spatial information. The average object identity (Fig. 2d) and spatial background (Fig. 4d) decoding rankings plotted across LOC, PPA, and RSC. Note the opposing gradients for the 2 types of information. Note also the significantly higher object identity decoding ranks in LOC, spatial background decoding ranks in RSC, and the equivalent ranks for both types of information in PPA. * $P < 0.05$ difference between information types (Wilcoxon signed-ranks test).

object information commensurate with its strong dorsal and weak ventral pathway connections. In contrast, LOC showed the opposite pattern consistent with its opposing pattern of connectivity. PPA, which is connected with both the dorsal and ventral visual pathways, showed equal object and spatial background information.

Structure of Representations

Finally, we performed an analysis designed to evaluate directly the similarity in the representations across LOC, PPA, and RSC. In contrast to the above analyses, which tested particular hypotheses about different types of scene information, the structural similarity analysis was largely agnostic to the particular experimental conditions. We took the full similarity matrices (21×21 conditions, as we excluded the space- and object-absent conditions), which capture the structure of each region's representation across the set of scenes containing both object and spatial layout information. We correlated these matrices between regions within each participant. Based on the anatomical connections, RSC was predicted to be correlated mainly with PPA to which it projects directly as part of the parietomedial temporal pathway, whereas PPA was predicted to be correlated not only with RSC but also with LOC as it receives direct inputs from ventral occipitotemporal regions (Kravitz, Saleem, et al. 2011). Therefore, the key comparison is the relative level of correlation between RSC-PPA and RSC-LOC. If the anatomical connectivity can predict the relationship between the structure of representations across regions, the structure in RSC should be more strongly correlated with that in PPA than in LOC.

It is important to note that the majority of the variability in our stimulus set was between objects (7 instances) rather than spatial backgrounds (2 instances). This imbalance will tend to inflate correlations involving regions with strong object information, as their matrices will contain more structure. However, this effect works against our hypothesis as LOC has stronger object information than does PPA (see above).

In fact, the results of this analysis (Fig. 6) directly support our hypothesis. Stronger correlations were found between RSC and PPA ($r = 0.31$) than between RSC and LOC ($r = 0.18$) ($t_{18} = 3.31, P < 0.01$). LOC was also more strongly connected to PPA ($r = 0.39$) than RSC ($t_{18} = 9.70, P < 0.001$). PPA was more strongly connected to LOC than to RSC ($t_{18} = 2.65, P < 0.05$), though this result might reflect the much stronger object information in LOC.

Discussion

The key finding of the current study is the opposing gradients of object and spatial information across LOC, PPA, and RSC. At one end of the continuum, RSC evidenced almost no information about objects, and its response showed no modulation by the presence or absence of objects. In contrast, LOC contained strong object information but very little information about spatial backgrounds, and its response was not modulated by their presence. Importantly, PPA showed information about “both” objects and spatial backgrounds and was sensitive to the presence or absence of either. Together, these findings are well predicted by the large-scale cortical circuits in which these regions reside and highlight the importance of understanding anatomical connectivity for teasing apart distributed information processing in cortex.

The specific contribution of the regions engaged by visual scenes has been a matter of much debate (for review, see Epstein 2008). On one account, PPA and RSC are part of a network that represents spatial aspects of scenes. PPA responds equally well to empty and furnished rooms (Epstein and Kanwisher 1998), responds more strongly to Lego “scenes” than “objects” (Epstein et al. 1999), even in blind subjects

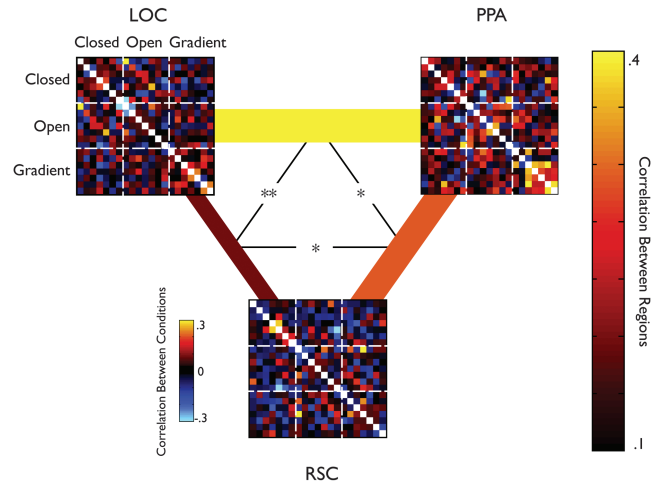


Figure 6. Structure correlations across ROIs. To systematically compare the structure of representations across PPA, RSC, and LOC, we cross-correlated the full similarity matrices between every possible pair of the 3 ROIs, in each participant. Examples of these matrices for a single subject are presented for the 3 ROIs. These similarity matrices (21×21 conditions, by excluding the space-absent and object-absent conditions) capture the structure of representations for each region across the set of scenes. We correlated these matrices between regions within each subject and then averaged the correlation values across subjects. Note that the diagonal values within each matrix were excluded from this analysis and hence are colored white. The lines connecting the similarity matrices for each ROI depict the average correlation between them across subjects. The color and width of each line represent the strength of the correlation between the similarity matrices in each pair of regions. Note the greater correlation between RSC and PPA than between RSC and LOC. ** $P < 0.001$ difference between the correlation between 2 pairs of ROIs; * $P < 0.05$ difference between the correlation between 2 pairs of ROIs.

(Wolbers et al. 2011), and represents spatial factors such as boundary (open versus closed—Kravitz, Peng, et al. 2011; Park et al. 2011), suggesting that PPA encodes gross spatial structure (Epstein and Kanwisher 1998; Epstein et al. 1999; Kravitz, Peng, et al. 2011; Park et al. 2011). RSC is active during mental navigation (Ino et al. 2002), retrieval of spatial information (Epstein, Parker, et al. 2007), and provides scene representations which are view invariant (Epstein, Higgins, et al. 2007; Park and Chun 2009) and extrapolate beyond the border of scenes (Epstein and Higgins 2007; Park et al. 2007). Under this account, LOC might constitute a separate object-based channel for scene recognition (Kim and Biederman 2011; MacEvoy and Epstein 2011). For example, a recent study showed that scene category could be predicted based on the response pattern of LOC to a combination of diagnostic objects from that semantic category (MacEvoy and Epstein 2011).

Alternatively, it has been suggested that PPA and RSC represent contextual associations in general, rather than space per se (Bar et al. 2008; Kveraga et al. 2011). Along these lines, PPA and RSC exhibit higher responses to real-world scenes that have rich contextual associations than those with only weak associations (Bar et al. 2008; but see Epstein and Ward 2010). Furthermore, it was recently proposed that PPA and RSC are distinct scene processing channels, with PPA representing scene category (e.g., beach) and RSC representing navigational information (Dilks et al. 2011).

Our findings do not fit entirely with any of these views, showing strong similarity between LOC and PPA in terms of object information on the one hand and between PPA and RSC in terms of spatial layout information on the other. Consequently, the current findings support the notion that the separation

between scene and object regions is not strictly categorical (see also Haxby et al. 2001). We demonstrate here that the scene region PPA is sensitive to object information, complementing prior studies showing that object regions such as LOC might also be involved in scene recognition (Kim and Biederman 2011; MacEvoy and Epstein 2011). We next consider each of the scene-processing regions in turn and examine how our findings can be integrated with the extant literature.

The role of retrosplenial cortex in spatial cognition and scene processing has been investigated in both humans and nonhumans (Maguire 2001; Byrne et al. 2007; Epstein 2008; Vann et al. 2009). Retrosplenial cortex lies within the parietomedial temporal pathway (Kravitz, Saleem, et al. 2011) that conveys spatial information between the posterior parietal cortex and the medial temporal lobe. Notably, retrosplenial cortex has little direct connectivity with the ventral visual pathway (Vogt and Pandya 1987; Kobayashi and Amaral 2003, 2007). Lesions of retrosplenial cortex lead to heading disorientation (Aguirre and D'Esposito 1999), a disorder that spares object recognition but leaves patients unable to update their egocentric heading relative to the environment (Hashimoto et al. 2010). Consistent with these observations, we found strong spatial layout information but almost no object information in the functionally defined RSC. Together, our results are consistent with theories proposing a spatial role for RSC (Epstein, Parker, et al. 2007; Park and Chun 2009; Dilks et al. 2011; Kravitz, Peng, et al. 2011), but not with suggestions that RSC represents nonspatial contextual associations (Bar et al. 2008), since its response was wholly insensitive to object presence. Furthermore, the data from RSC argue against an attentional account of our results, whereby scenes containing objects draw more attention than empty scenes, since RSC's response was not enhanced in the object-present condition.

Showing an opposing connectivity pattern, LOC, which lies lateral to the parahippocampal cortex within the ventral visual pathway, has little direct connectivity with the parahippocampal cortex or the parietomedial temporal pathway (Kondo et al. 2005). Both lesion (e.g., James et al. 2003) and transcranial magnetic stimulation studies (e.g., Pitcher et al. 2009) have confirmed the necessity of LOC for object recognition with no concomitant spatial deficits. Consistent with these reports, we found strong object information but no information about spatial layout and no modulation of response by background. Interestingly, decoding of object identity within LOC was significantly lower when objects were positioned on a spatial rather than nonspatial background, suggesting sensitivity to spatial background information, but no detailed representation of spatial structure. This lack of detailed spatial representations suggests that LOC's contribution to scene processing is primarily object-based, in line with recent theories (Kim and Biederman 2011; MacEvoy and Epstein 2011).

Relative to RSC and LOC, PPA has the most complex pattern of connectivity, and accordingly, the information it represents reflects its diverse inputs. In monkey, the parahippocampal cortex (areas TF/TH, TFO) receives input from the ventral portions of V4 (Ungerleider et al. 2008). However, it also has direct connectivity with the dorsal visual pathway, retrosplenial, and posterior cingulate cortices, as well as the hippocampus via the parietomedial temporal pathway. Consistent with this connectivity, PPA evidenced both strong object and strong spatial layout information.

We found that PPA was not only sensitive to the presence of space but also it was also sensitive to the specific spatial

background presented, discriminating between open and closed scenes. This supports the notion that its scene representations capture spatial dimensions, particularly spatial expanse, as demonstrated in prior studies employing naturalistic stimuli (Kravitz, Peng, et al. 2011; Park et al. 2011) even with minimal detail. PPA was also the only region to evidence a consistent pattern of response to the nonspatial gradient background, consistent with suggestions that it is also involved in the representation of global statistics and textures (Cant and Goodale 2011; Cavina-Pratesi et al. 2011).

Strikingly, while PPA is usually considered to support large-scale analysis of scene layout rather than analysis of local detail, we found that PPA's response magnitude was modulated by object presence, in contrast to prior studies (Epstein and Kanwisher 1998; Epstein et al. 1999; Wolbers et al. 2011). Furthermore, PPA's response pattern could be used to distinguish between the presence and absence of an object and even to discriminate between highly similar objects (large furniture items). Finally, PPA evidenced object information almost as strong as that observed in LOC in the presence of spatial backgrounds. However, whereas object decoding in LOC improved in the absence of a spatial background, PPA decoding was unaffected or even slightly decreased. This intriguing finding stresses the complex interactions between object and spatial information in PPA and suggests that while both patterns of response in PPA and LOC can be used to decode object identity, their underlying neural representations may be inherently different. To further our understanding of the functional relationship between PPA and LOC, we must establish whether they are directly connected or simply share a common input from ventral V4. Neither our structural analysis nor functional connectivity analyses can distinguish between these possibilities.

Taken together, our current findings are consistent with a role for PPA in representing the spatial aspects of scenes (Epstein 2008; Kravitz, Peng, et al. 2011; Park et al. 2011). Furthermore, the strong object information in PPA reported here implies that spatial and object information are not necessarily dissociated in scene recognition. PPA is likely involved in binding "navigationally relevant" objects with spatial representations to aid in navigation and spatial imagery (Hassabis et al. 2007; Mullally and Maguire 2011). Thus, PPA primarily group scenes by their spatial aspects (Kravitz, Peng, et al. 2011) but still has information about navigationally relevant objects (Janzen and van Turennout 2004).

Another potential source of object information in PPA may be spatial or depth information implied by local objects. Notably, the objects used in the current study are furniture items, which are large fixed objects that occupy a position in space, imply depth and perspective, and provide small-scale navigational affordances. Thus, it may be the case that the objects themselves provide diagnostic information regarding the spatial environment. And whereas RSC primarily encodes the global aspect of space, PPA is sensitive to both sources of spatial information. This raises the question whether our finding that PPA is sensitive to object content is limited to large stationary (i.e., navigationally relevant) objects or whether it can be generalized to other types of objects, particularly smaller manipulable objects (Spiridon and Kanwisher 2002). Mullally and Maguire (2011) recently demonstrated that the object sensitivity in parahippocampal cortex is modulated by how "space-defining" the object perceived (or imagined) is. Furthermore, representations in ventral temporal cortex may

be organized according to real-world object size (Konkle and Oliva 2011). These findings suggest that parahippocampal cortex represents objects in a graded fashion rather than dichotomically. Future research addressing this question will have to take into account not only object size but also other factors such as the depth information implied by the object's viewpoint and orientation as well as the object's contextual relation with its environment (MacEvoy and Epstein 2011). Finally, even if one focuses on the large fixed landmark-suitable objects, it is still striking that PPA can make such fine within-category distinctions. It is important to note that the sensitivity of PPA to object identity may be based on multiple factors, including low/midlevel features (e.g., color, see Steeves et al. 2004) as well as "high-level" factors (e.g., global scene structure, see Walther et al. 2011). Indeed, there must be some visual statistic or combination thereof that is the basis for identity discrimination in PPA because all visual representations, whether high or low level, must reflect some difference in the images.

The current work represents a significant step forward in understanding the neural basis of scene processing as well as a demonstration that anatomical connectivity can be used to predict the functional representations found within any particular cortical region. Scene processing has been very difficult to understand because 1) the complexity and heterogeneity of real-world scenes make it difficult to know the source of any observed effect (e.g., Aminoff et al. 2007; Bar et al. 2008; Epstein and Ward 2010) and 2) the large number of cortical regions involved and their equally complex connectivity make it difficult to understand the contribution of each region. Our general approach to this problem was 2-fold. First, we developed a large-scale anatomical framework, which allowed us to make specific predictions about the functional nature of representations within each region. Second, we used this framework to inform the design of complementary studies that use both naturalistic and artificial scenes to understand the nature of the information being represented in the scene network. Of course, each mode of investigation has its advantages and limitations. Using naturalistic stimuli and data-driven analyses has high ecological validity, but at the same time, it is limited by difficulties in understanding of the underlying sources of the effects. Conversely, using artificially scenes, as in the current study, allow a systematic teasing apart of the different components of scenes, but the experimenter must have a priori hypotheses about the diagnostic dimensions, the stimuli are less rich compared with natural images, and care must be taken in generalizing the results to more complex and realistic visual scenes. The strength of our overall approach is that it integrates these complementary modes of investigation, enabling the generation of new hypotheses guided and constrained by the neuroanatomical data.

In conclusion, using a neuroanatomical framework from monkey and carefully controlled stimuli allowing us to independently manipulate spatial and object information, we have demonstrated the presence of 2 opposing gradients in regions engaged by visual scenes. Spatial layout information increases from LOC to PPA to RSC, while object information decreases. RSC evidenced no information about objects but strong information about spatial layout, commensurate with its prominent dorsal and weak ventral pathway connections. In contrast, LOC contained strong object information but very little spatial layout information. Importantly, PPA, which is connected with both the dorsal and ventral visual pathway, showed information about both objects

and spatial backgrounds. These findings are a demonstration that the functional properties of a region are predictable from its large-scale connectivity, although these predictions could be refined with an understanding of the relative strength of feedforward and feedback processing. The consistency observed in the topography of functions across individuals (e.g., location of category-selective regions) may reflect constraints from connectivity. Importantly, this model undermines a strict separation in function between regions, suggesting that while different regions uniquely contribute to processing, their function is the product of integrating information across large-scale cortical networks.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

Funding

National Institute of Mental Health Intramural Research Program.

Notes

Thanks to Sandra Truong and Viktoria Elks for help with data collection and Saleem Kadharbatcha, Alex Martin, and members of the Laboratory of Brain and Cognition for helpful comments and discussion. *Conflict of Interest:* None declared.

References

- Aguirre GK, D'Esposito M. 1999. Topographical disorientation: a synthesis and taxonomy. *Brain*. 122(Pt 9):1613-1628.
- Aguirre GK, Detre JA, Alsup DC, D'Esposito M. 1996. The parahippocampus subserves topographical learning in man. *Cereb Cortex*. 6:823-829.
- Aminoff E, Gronau N, Bar M. 2007. The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb Cortex*. 17:1493-1503.
- Bar M, Aminoff E. 2003. Cortical analysis of visual context. *Neuron*. 38:347-358.
- Bar M, Aminoff E, Schacter DL. 2008. Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J Neurosci*. 28:8539-8544.
- Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG. 2009. Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J Neurophysiol*. 101:688-700.
- Byrne P, Becker S, Burgess N. 2007. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol Rev*. 114:340-375.
- Cant JS, Goodale MA. 2011. Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. *J Neurosci*. 31:8248-8258.
- Cavina-Pratesi C, Kentridge RW, Heywood CA, Milner AD. 2011. Separate channels for processing form, texture, and color: evidence from fMRI adaptation and visual object agnosia. *Cereb Cortex*. 20:2319-2332.
- Chan AW, Kravitz DJ, Truong S, Arizpe J, Baker CI. 2010. Cortical representations of bodies and faces are strongest in commonly experienced configurations. *Nat Neurosci*. 13:417-418.
- Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N. 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. *J Neurosci*. 31:11305-11312.
- Epstein R, Harris A, Stanley D, Kanwisher N. 1999. The parahippocampal place area: recognition, navigation, or encoding? *Neuron*. 23:115-125.
- Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature*. 392:598-601.

- Epstein RA. 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci*. 12:388–396.
- Epstein RA, Higgins JS. 2007. Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cereb Cortex*. 17:1680–1693.
- Epstein RA, Higgins JS, Jablonski K, Feiler AM. 2007. Visual scene processing in familiar and unfamiliar environments. *J Neurophysiol*. 97:3670–3683.
- Epstein RA, Parker WE, Feiler AM. 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *J Neurosci*. 27:6141–6149.
- Epstein RA, Ward EJ. 2010. How reliable are visual context effects in the parahippocampal place area? *Cereb Cortex*. 20(2):294–303.
- Grill-Spector K, Kourtzi Z, Kanwisher N. 2001. The lateral occipital complex and its role in object recognition. *Vision Res*. 41:1409–1422.
- Hashimoto R, Tanaka Y, Nakano I. 2010. Heading disorientation: a new test and a possible underlying mechanism. *Eur Neurol*. 63:87–93.
- Hassabis D, Kumaran D, Maguire EA. 2007. Using imagination to understand the neural basis of episodic memory. *J Neurosci*. 27:14365–14374.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*. 293:2425–2430.
- Ino T, Inoue Y, Kage M, Hirose S, Kimura T, Fukuyama H. 2002. Mental navigation in humans is processed in the anterior bank of the parieto-occipital sulcus. *Neurosci Lett*. 322:182–186.
- James TW, Culham J, Humphrey GK, Milner AD, Goodale MA. 2003. Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain*. 126:2463–2475.
- Janzen G, van Turenout M. 2004. Selective neural representation of objects relevant for navigation. *Nat Neurosci*. 7:673–677.
- Joubert OR, Rousselet GA, Fize D, Fabre-Thorpe M. 2007. Processing scene context: fast categorization and object interference. *Vision Res*. 47:3286–3297.
- Kim JG, Biederman I. 2011. Where do objects become scenes? *Cereb Cortex*. 21:1738–1746.
- Kobayashi Y, Amaral DG. 2003. Macaque monkey retrosplenial cortex: II. Cortical afferents. *J Comp Neurol*. 466:48–79.
- Kobayashi Y, Amaral DG. 2007. Macaque monkey retrosplenial cortex: III. Cortical efferents. *J Comp Neurol*. 502:810–833.
- Kondo H, Saleem KS, Price JL. 2005. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *J Comp Neurol*. 493:479–509.
- Kourtzi Z, Kanwisher N. 2001. Representation of perceived object shape by the human lateral occipital complex. *Science*. 293:1506–1509.
- Kravitz DJ, Kriegeskorte N, Baker CI. 2010. High-level visual object representations are constrained by position. *Cereb Cortex*. 20:2916–2925.
- Kravitz DJ, Peng CS, Baker CI. 2011. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *J Neurosci*. 31:7322–7333.
- Kravitz DJ, Saleem KS, Baker CI, Mishkin M. 2011. A new neural framework for visuospatial processing. *Nat Rev Neurosci*. 12:217–230.
- Kriegeskorte N, Mur M, Bandettini P. 2008. Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci*. 2:4.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. 12:535–540.
- Kveraga K, Ghuman AS, Kassam KS, Aminoff EA, Hamalainen MS, Chaumon M, Bar M. 2011. Early onset of neural synchronization in the contextual associations network. *Proc Natl Acad Sci U S A*. 108:3389–3394.
- MacEvoy SP, Epstein RA. 2009. Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Curr Biol*. 19:943–947.
- MacEvoy SP, Epstein RA. 2011. Constructing scenes from objects in human occipitotemporal cortex. *Nat Neurosci*. 14:1323–1329.
- Maguire EA. 2001. The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand J Psychol*. 42:225–238.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A*. 92:8135–8139.
- Mullally SL, Maguire EA. 2011. A new role for the parahippocampal cortex in representing space. *J Neurosci*. 31:7441–7449.
- Mur M, Bandettini PA, Kriegeskorte N. 2009. Revealing representational content with pattern-information fMRI—an introductory guide. *Soc Cogn Affect Neurosci*. 4:101–109.
- Nasr S, Liu N, Devaney KJ, Yue X, Rajimehr R, Ungerleider LG, Tootell RBH. 2011. Scene-selective cortical regions in human and nonhuman primates. *J Neurosci*. 31:13771–13785.
- Park S, Brady TF, Greene MR, 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. *J Neurosci*. 31:1333–1340.
- Park S, Chun MM. 2009. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage*. 47:1747–1756.
- Park S, Intraub H, Yi DJ, Widders D, Chun MM. 2007. Beyond the edges of a view: boundary extension in human scene-selective visual cortex. *Neuron*. 54:335–342.
- Pitcher D, Charles L, Devlin JT, Walsh V, Duchaine B. 2009. Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr Biol*. 19:319–324.
- Potter MC. 1976. Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn*. 2:509–522.
- Potter MC, Levy EI. 1969. Recognition memory for a rapid sequence of pictures. *J Exp Psychol*. 81:10–15.
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N. 2008. The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci U S A*. 105:4447–4452.
- Spiridon M, Kanwisher N. 2002. How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron*. 35:1157–1165.
- Steeves JKE, Humphrey GK, Culham JC, Menon RS, Milner AD, Goodale MA. 2004. Behavioral and neuroimaging evidence for a contribution of color and texture information to scene classification in a patient with visual form agnosia. *J Cogn Neurosci*. 16:955–965.
- Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. *Nature*. 381:520–522.
- Ungerleider LG, Galkin TW, Desimone R, Gattass R. 2008. Cortical connections of area V4 in the macaque. *Cereb Cortex*. 18:477–499.
- Vann SD, Aggleton JP, Maguire EA. 2009. What does the retrosplenial cortex do? *Nat Rev Neurosci*. 10:792–802.
- Vogt BA, Pandya DN. 1987. Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J Comp Neurol*. 262:271–289.
- Walther DB, Caddigan E, Fei-Fei L, Beck DM. 2009. Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci*. 29:10573–10581.
- Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L. 2011. Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proc Natl Acad Sci U S A*. 108:9661–9666.
- Wolbers T, Klatzky RL, Loomis JM, Wutte MG, Giudice NA. 2011. Modality-independent coding of spatial layout in the human brain. *Curr Biol*. 21:984–989.