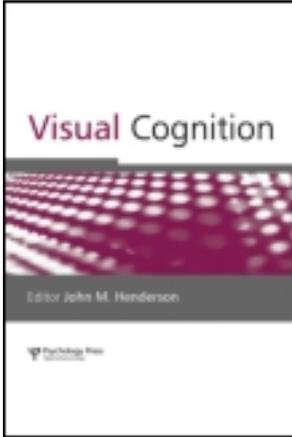


This article was downloaded by: [University of South Carolina]

On: 18 April 2013, At: 12:58

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



Visual Cognition

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pvis20>

Functions of parahippocampal place area and retrosplenial cortex in real-world scene analysis: An fMRI study

John M. Henderson ^a, David C. Zhu ^b & Christine L. Larson ^c

^a University of South Carolina, USA

^b Michigan State University, USA

^c University of Wisconsin, USA

Version of record first published: 25 Aug 2011.

To cite this article: John M. Henderson, David C. Zhu & Christine L. Larson (2011): Functions of parahippocampal place area and retrosplenial cortex in real-world scene analysis: An fMRI study, *Visual Cognition*, 19:7, 910-927

To link to this article: <http://dx.doi.org/10.1080/13506285.2011.596852>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages

whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Functions of parahippocampal place area and retrosplenial cortex in real-world scene analysis: An fMRI study

John M. Henderson¹, David C. Zhu²,
and Christine L. Larson³

¹University of South Carolina, USA

²Michigan State University, USA

³University of Wisconsin, USA

We used functional MRI to investigate several hypotheses concerning the functions of posterior parahippocampal cortex and retrosplenial cortex, two regions that preferentially activate to images of real-world scenes compared to images of other meaningful visual stimuli such as objects and faces. We compared activation resulting from photographs of rooms, city streets, cityscapes, and landscapes against activation to a control condition of objects. Activation in posterior parahippocampal cortex, including parahippocampal place area, was greater for all scene types than objects, and greater for scenes that clearly convey information about local three-dimensional (3-D) structure (city streets and rooms) than scenes that do not (cityscapes and landscapes). Similar differences were observed in retrosplenial cortex, though activation was also greater for city streets than rooms. These results suggest that activation in both cortical areas is primarily related to analysis or representation of local 3-D space. The results are not consistent with hypotheses that these areas reflect panoramic spatial volume, an artificial versus natural category distinction, an indoor versus outdoor distinction, or the number of explicit objects depicted in a scene image.

Keywords: Functional MRI; High-level vision; Parahippocampal place area; Real-world scenes; Retrosplenial cortex.

Please address all correspondence to John M. Henderson, Department of Psychology, and McCausland Center for Brain Imaging, University of South Carolina, Columbia, SC 29208, USA. E-mail: john.henderson@sc.edu

We thank Michigan State University's Departments of Radiology and Psychology for their support, and David McFarlane, Jie Huang, Jeremy Grounds, Shawn Katterman, Stephen Kemsley, and Jordan Robinson in East Lansing, and Krista Ehinger in Edinburgh, for their assistance on the project.

Real-world scenes can be defined as views from a particular vantage point of a real-world environment (Henderson, 2005, 2007). There is growing evidence that perceptual and cognitive processing of real-world scenes differs in important ways from other types of meaningful visual stimuli such as objects and faces. For example, real-world scenes contain unique regularities in spatial structure and semantics that are not present in other types of visual stimuli (Biederman, Mezzanotte, & Rabinowitz, 1982; Henderson & Hollingworth, 1999). From a processing perspective, scenes are recognized very quickly using a route that does not seem to require the recognition of individual objects and their spatial relationships, suggesting a direct pathway to scene categorization (Intraub, 1981; Li, van Rullen, Koch, & Perona, 2005; Potter, 1976; Rousselet, Joubert, & Fabre-Thorpe, 2005; Schyns & Oliva, 1994; Thorpe, Fize, & Marlot, 1996).

Is the special status of visual scenes supported by domain-specific processing systems in the brain? Data from functional MRI (fMRI) point to two candidate cortical scene regions. One region, the parahippocampal place area (PPA; Epstein, Harris, Stanley, & Kanwisher, 1999; Epstein & Kanwisher, 1998; see Epstein, 2005, for review) is located in the posterior region of parahippocampal cortex (pPHC) and is selectively activated by scenes over faces, single objects, and object arrays (Epstein, Graham, & Downing, 2003; Epstein et al., 1999; Epstein, Higgins, & Thompson-Schill, 2005; Epstein & Kanwisher, 1998; Epstein, Parker, & Feiler, 2007; Spiridon & Kanwisher, 2002; see also Aguirre & D'Esposito, 1997; Bar & Aminoff, 2003; Henderson, Larson, & Zhu, 2007; Kohler, Crane, & Milner, 2002; Maguire, Frith, & Cipolotti, 2001; Walther, Caddigan, Fei-Fei, & Beck, 2009). A second area, retrosplenial cortex (RSC), and particularly Brodmann's Area 30 (BA 30), similarly preferentially activates to real-world scenes over faces and objects (Bar & Aminoff, 2003; Epstein & Higgins, 2006; Epstein et al., 2007; Henderson, Larson, & Zhu, 2007, 2008; Walther et al., 2009). Although pPHC and RSC both appear to be involved in real-world scene analysis, it remains unclear what specific functions these areas support and whether they differ.

The purpose of the present study was to investigate further the functions of PPA and RSC in scene processing. One difficulty with studying real-world scenes is that they can differ along a number of dimensions that are often correlated. For example, indoor scenes tend to have local three-dimensional (3-D) structure (i.e., the volumetric geometry of the local space is explicitly defined), whereas outdoor scenes tend not to explicitly depict local volumetric structure. Similarly, scenes of natural environments do not typically depict local volumetric structure (i.e., they are panoramic in that they provide an open view of great spatial extent), whereas scenes of artificial environments tend to be closed spaces with explicitly defined local 3-D structure. Therefore, it can be difficult to disentangle effects of different sets

of correlated scene features (e.g., indoor/outdoor, artificial/natural, open/closed), particularly when relatively small and homogenous sets of scenes are used as stimuli.

Furthermore, there are reasons to think that some of these variables might influence the degree of activation observed in cortical regions sensitive to scenes. First, natural and artificial scenes have different Fourier power spectra, with artificial scenes possessing more high-frequency power in horizontal and vertical dimensions (e.g., Torralba & Oliva, 2003; Vailaya, Jain, & Jiang Shang, 1998). Recent evidence suggests that PPA may be more generally sensitive to high spatial frequencies (Rajimehr, Devaney, Young, Postelnicu & Tootell, 2008). Second, many objects tend to be associated with particular environments, and object–scene associations appear to activate PPA (Bar, 2004; Bar & Aminoff, 2003; Bar, Aminoff, & Ishai, 2008; Bar, Aminoff, & Schacter, 2008; see also Eichenbaum, Yonelinas, & Ranganath, 2007; but see Epstein & Ward, 2010). Greater numbers of objects might generate more associations and so more activation (Bar, Aminoff, & Ishai, 2008; Bar, Aminoff, & Schacter, 2008).

In the present study we included conditions that allowed us to begin to tease apart the degree to which PPA and RSC preferentially activate to scenes with and without explicit local 3-D volumetric structure, smaller versus larger degrees of spatial volume, artificial versus natural scene categories, and larger versus smaller numbers of individuated objects. These features were contrasted in a large set of photographs of rooms, city streets, cityscapes (skylines), and landscapes (Figure 1). The rooms and city streets both clearly depicted local enclosed 3-D volumetric structure, with at least three surfaces enclosing the depicted space, whereas the cityscapes and landscapes did not. The rooms were indoor environments, whereas the streets, cityscapes, and landscapes were outdoor environments. The rooms, streets, and cityscapes were artificial environments, whereas the landscapes were natural environments. The cityscapes and landscapes were both outdoor panoramic environments, but the cityscapes were artificial whereas the landscapes were natural. The rooms and city streets were both artificial environments with local enclosed 3-D structure, but the rooms included large numbers of individuated human-scaled objects that could be manipulated within the context of the scenes (Henderson & Hollingworth, 1999), whereas the streets did not.

With these conditions, then, we can draw predictions for several hypotheses concerning PPA and RSC function (Figure 2). First, if PPA or RSC are involved in analysing explicit local 3-D volumetric structure, then activation should be less for cityscapes and landscapes (which are panoramic) than for rooms and city streets (which depict the local enclosed 3-D environment surrounding the viewer). Second, if PPA or RSC preferentially active to scenes with less local volumetric structure but more navigable space (i.e., more open

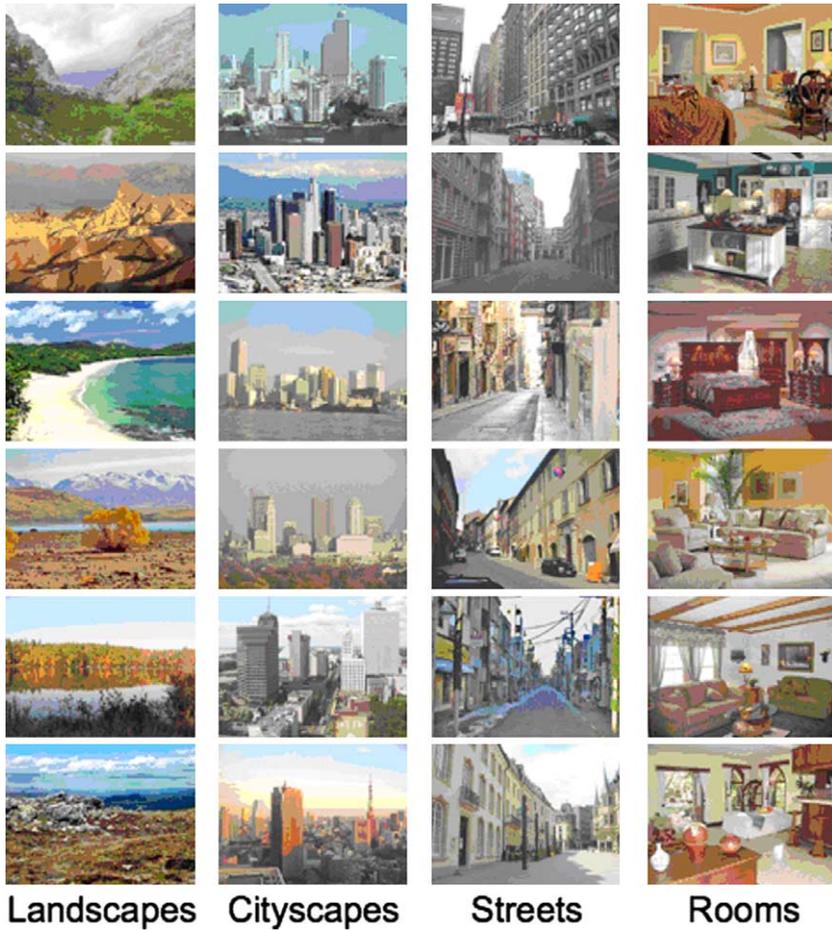


Figure 1. Examples of the scene images used in the experiment. A set of 100 unique scenes were presented without repetition to each subject in each condition, along with an equal number of object control stimuli. To view this figure in colour, please see the online issue of the Journal.

space), then we predict activation in the following order from most to least: Landscapes, cityscapes, city streets, and rooms. Third, if PPA or RSC are involved in processing features related to artificial versus natural scenes, consistent with results showing greater activation to indoor than outdoor scenes (Henderson et al., 2007; see also the outdoor versus landscape comparison in Epstein & Kanwisher, 1998), then rooms, city streets, and cityscapes should produce more activation than landscapes. The cityscape versus landscape contrast provides a more controlled test of the artificial versus natural contrast holding spatial volume of the scene constant. Fourth,

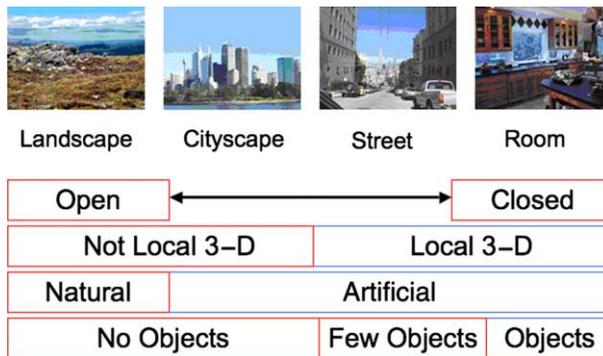


Figure 2. Predictions of the spatial volume, local 3-D structure, natural versus artificial, and object-association hypotheses for the four scene conditions used in the experiment. To view this figure in colour, please see the online issue of the Journal.

if processing in PPA or RSC reflects associations to explicitly depicted objects, then we predict less activation for cityscapes and landscapes (which explicitly depict no discrete objects) compared to rooms (which depict many). The city streets we chose for the experiment also depicted very few human-scaled individuated objects, but there were arguably some (e.g., parked cars in a street), so if objects are important then streets would be expected to produce more activation than cityscapes and landscapes.

To test these hypotheses, we presented a large number of unique scenes in each of the image conditions in a block-design fMRI experiment (Henderson et al., 2007, 2008). The use of large unrepeated image sets circumvents reduction in activation associated with scene repetition (Epstein et al., 2003, 1999, 2005; Goh et al., 2004; Grill-Spector et al., 1999; Grill-Spector & Malach, 2001). The use of a larger n also reduces potential sampling bias problems, and better supports generalization of the results to the population of scenes sampled in each condition. As a control condition, we included an equal number of object photographs. Because we were interested in scene perception and interpretation rather than intentional memory encoding or retrieval, participants passively viewed the images (Goh et al., 2004). To ensure that they attended to the display, participants were asked to press a button once at the beginning of each block of pictures. Finally, because we were interested in relatively complete scene interpretation rather than initial gist processing, scenes were displayed for 2.5 s each. Although the initial gist of a scene can be generated from as little as 50 ms of scene presentation, participants must scan a scene over extended time to fully encode and understand the objects and spatial relationships present in them (Henderson, 2003, 2007). Because several of the hypotheses under consideration here require complete encoding of spatial and object information, the use of the

longer display duration was necessary. There is good evidence that longer scene display durations result in activation of the same scene regions as shorter display durations (e.g., Walther et al., 2009).

MATERIALS AND METHOD

Subjects

Twenty-three right-handed, healthy college students from Michigan State University volunteered to participate in this study and signed consent forms approved by the Michigan State University Institutional Review Board. Data from five subjects were discarded: One due to lack of activation in primary visual areas suggesting lack of attention during the study, two due to irregular anatomical structures, one because the functionally defined PPA was not found, and one due to excess motion. Eighteen subjects (eight male, mean age 20.4, range 18–25) were included in the data analysis.

Stimuli

Stimuli were 500 full-colour digitized photographs selected from a variety of sources, with 100 unique pictures for each of the five conditions: Rooms, City Streets, Cityscapes, Landscapes, and Objects (Figure 1). A variety of picture types were used in each category within the constraints of the hypotheses to be tested (e.g., landscapes did not include artificial regions; city streets depicted a street environment with clear local enclosed 3-D volume). Stimuli were displayed in colour on a 640×480 LCD monitor mounted on top of the RF head coil. The LCD subtended $12^\circ \times 16^\circ$ of visual angle.

Procedure

A block-design paradigm was controlled by an IFIS-SA system (Invivo Corp., Gainesville, FL). The experiment was divided into five functional runs each lasting 6 min and 55 s. In each run, subjects were presented with 10 blocks of visual stimulation after an initial 15 s “resting” period. In each block, 10 unique pictures from one condition were presented. Within a block, each picture was presented for 2.5 s with no interstimulus interval. A 15 s baseline condition (a white screen with a black fixation cross at the centre) followed each block. Each condition was shown in two blocks per run. Both the order of conditions within each run and the order of pictures within a block were initially randomly determined. The five functional runs were presented to half of the subjects in a forward order and half in a reverse order.

Image acquisition

The experiment was conducted on a 3T GE Signa EXCITE scanner (GE Healthcare, Milwaukee, WI) with an eight-channel head coil. During each session, images were first acquired for the purpose of localization, and first and higher order shimming procedures were then carried out to improve magnetic field homogeneity (Kim, Adalsteinsson, Glover, & Spielman, 2002). To study brain function, echo planar images, starting from the most inferior regions of the brain, were then acquired with the following parameters: 34 contiguous 3 mm axial slices in an interleaved order, TE = 27.7 ms, TR = 2500 ms, flip angle = 80°, FOV = 22 cm, matrix size = 64 × 64, ramp sampling, and with the first four data points discarded. Each volume of slices was acquired 162 times during each of the five functional runs while subjects viewed the pictures, resulting in a total of 810 volumes of images over the course of the entire experiment. After functional data acquisition, high-resolution volumetric T_1 weighted spoiled gradient-recalled (SPGR) images with cerebrospinal fluid suppressed were obtained to cover the whole brain with 124 1.5 mm sagittal slices, 8° flip angle, and 24 cm FOV. These images were used to identify anatomical locations.

fMRI data preprocessing and analysis

All fMRI data preprocessing and analysis was conducted with AFNI software (Cox, 1996). For each subject, acquisition timing difference was first corrected for different slice locations. With the first functional image as the reference, rigid-body motion correction was done in three translational and three rotational directions. The amount of motion in these directions was estimated and then the estimations were used in data analysis. For each subject, spatial blurring with a full width half maximum of 4 mm was applied to reduce random noise (Parrish, Gitelman, LaBar, & Mesulam, 2000), and also to reduce the issue of intersubject anatomical variation and Talairach transformation variation during group analysis. For the group analysis, all images were converted to Talairach coordinate space (Talairach & Tournoux, 1988) with an interpolation to 1 mm³ voxels.

For the data analysis of each individual subject, the reference function throughout all functional runs for each picture category was generated based on the convolution of the stimulus input and a gamma function (Cox, 1996), which was modelled as the impulse response when each picture was presented. The functional image data acquired was compared with the reference functions using the 3-DDeconvolve software for multiple linear regression analysis and general linear tests (Ward, 2002). Multiple linear regressions were applied on a voxel-wise basis for *t*-statistic tests and to find the magnitude change when each picture condition was presented,

compared to the reference functions. The equivalent BOLD percentage signal change relative to the baseline state was then calculated. General linear tests were also applied on a voxel-wise basis to find the statistical significance of pairwise comparisons for all the picture conditions. For the previous analysis, in addition to applying the reference functions for the five picture conditions, MRI signal modelling also included the subject motion estimations in the three translational and the three rotational directions, and the constant, linear, and quadratic trends for each of the five functional runs.

Monte Carlo simulation of the effect of matrix and voxel sizes of the imaging volume, spatial correlation of voxels, voxel intensity thresholding, masking, and cluster identification was applied to estimate overall statistical significance with respect to the whole brain (Ward, 2000).

Parahippocampal place area (PPA) and retrosplenial cortex (RSC) region of interest (ROI) analyses

In Talairach coordinate space, a PPA ROI was identified in each hemisphere for each subject. Regions showing preferential activation to scenes over objects based on the contrast of “Rooms+City Street+Cityscape+Landscape – 4 × Object” (an 18-subject median whole-brain corrected p -value $< 1.38 \times 10^{-3}$) in the parahippocampal gyrus were defined as the PPA. The whole-brain corrected p -values were the results of the following active voxel selection criteria: A voxel-wise p -value $< 10^{-4}$ based on the t -test and active voxels that were nearest-neighbour and within a cluster size of 107 mm^3 . This contrast resulted in clear PPA ROIs in each hemisphere for each subject. We computed the average BOLD percentage signal change for all voxels in the PPA ROI individually for each subject as a function of condition and hemisphere. These data were entered into a Condition (Object, room, city street, cityscape, landscape) × Hemisphere (left, right) ANOVA. An analogous Condition × Hemisphere ANOVA was calculated for the mean percentage signal change by subject for all voxels in the anatomically defined BA 30 ROI. This RSC ROI was limited to BA 30 (defined using the AFNI Talairach atlas) due to a lack of differential activation in BA 29 to any type of scenes in the present study, as well as previous work indicating that RSC activation to scenes was most pronounced in BA 30 (Henderson et al., 2007; see also Maguire, 2001).

Whole brain analysis

After the percentage signal change was estimated with respect to each picture condition for each subject, an ANOVA was performed over the 18 subject data sets for group analysis with a mixed-effect two-factor model. Picture condition (five levels) was the first factor and was modelled to provide a fixed effect.

Subject was the second factor and was modelled as a random effect. The ANOVA results were used to extract the activated voxels for all pairwise condition contrasts (voxel-based p -value $< 10^{-3}$ and whole-brain corrected p -value $< .016$). The active voxel selection criteria required that the voxels were nearest-neighbour and within a cluster size of 142 mm^3 . Based on application of these criteria to the whole brain, the voxel-based p -value $< 1 \times 10^{-3}$ was corrected to be an equivalent of whole-brain corrected p -value $< .016$.

RESULTS

ROI analyses

PPA. All included subjects showed well-defined PPA ROIs, with an average right PPA ROI volume of 3967 mm^3 (range 1577 – 6239 mm^3), and average left PPA ROI volume of 3185 mm^3 (range 523 – 4732 mm^3). The centroids of the right and left PPA ROIs with respect to the t -values in Talairach coordinates based on the group analysis were $(20, -40, -2)$ and $(-19, -41, -2)$, respectively.

Average percentage signal change for all voxels in the PPA ROI was computed for each subject as a function of condition and hemisphere. These data were entered into a Condition \times Hemisphere ANOVA (Huynh-Feldt corrected). The analysis revealed a significant effect of condition, $F(4, 68) = 64.21$, $p < .001$, $\eta^2 = .791$, hemisphere, $F(1, 17) = 6.40$, $p < .02$, $\eta^2 = .274$, and a Condition \times Hemisphere interaction, $F(4, 68) = 4.41$, $p = .004$, $\eta^2 = .206$ (Figure 3). Post hoc comparisons (Bonferroni-adjusted) indicated that objects were less robust activators of PPA than all types of scenes ($ps < .001$). Among the scenes there was no difference in PPA activation between rooms and city streets ($p > .99$) or between cityscapes and landscapes ($p > .99$); however, both rooms and streets elicited greater

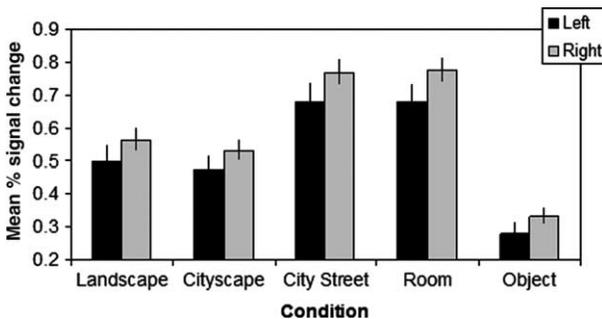


Figure 3. Mean percentage signal change (and SEM) across all voxels in the PPA ROI for 18 subjects as a function of image condition and hemisphere.

PPA activation than both cityscapes and landscapes ($ps < .004$). Following up on the Condition \times Hemisphere interaction, post hoc comparisons indicated greater activation in the right than left hemisphere for rooms and city streets ($p < .05$), but not objects or cityscapes ($ps > .05$). There was a trend for greater right hemisphere activation for landscapes ($p < .03$), but this contrast did not survive the Bonferroni-correction threshold.

RSC (BA 30). In an analogous Condition \times Hemisphere ANOVA on RSC BA 30, we found a significant main effect of condition, $F(4, 68) = 42.95$, $p < .001$, $\eta^2 = .716$, along with a main effect of hemisphere, $F(1, 17) = 5.20$, $p < .04$, $\eta^2 = .234$, which indicated greater right hemisphere activation. These main effects were qualified by a Condition \times Hemisphere interaction, $F(4, 68) = 6.82$, $p < .001$, $\eta^2 = .286$ (see Figure 4). Bonferroni-corrected post hoc contrasts for the condition effect indicated that all four scene types elicited greater activation than objects ($ps < .001$). Whereas rooms and city streets activated RSC similarly ($p = .18$), as did cityscapes and landscapes ($p > .99$), rooms and city streets tended to be the strongest elicitors of RSC activation. City streets led to greater percent signal change than did cityscapes or landscapes ($ps < .001$). Rooms also resulted in greater activation than cityscapes ($ps < .02$) and marginally greater activation than landscapes ($p < .09$). Considering hemispheric differences, Bonferroni-corrected post hoc contrasts revealed a right hemisphere activation advantage for city streets ($p < .05$), but the trends for greater right hemisphere activation for rooms, cityscapes, and landscapes did not survive Bonferroni correction ($ts < 2.65$).

Whole brain analyses

pPHC. The whole brain analyses confirmed the PPA ROI findings. As seen in Table 1 and Figure 5, similar to the ROI results, rooms and city streets produced equivalent activation, as did cityscapes and landscapes.

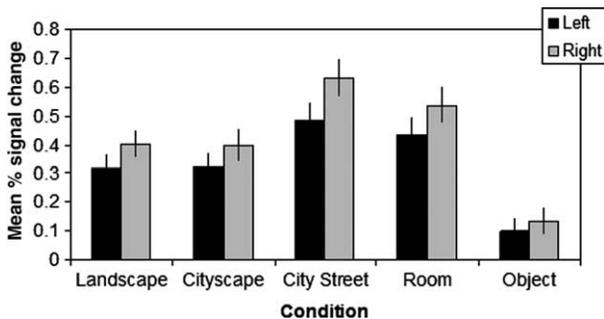


Figure 4. Mean percentage signal change (and SEM) across all voxels in the BA 30 ROI for 18 subjects as a function of image condition and hemisphere.

TABLE 1
Group results of the whole-brain analysis for right (R) and left (L) parahippocampal gyrus and Brodmann's Area 30

	<i>Parahippocampal gyrus</i>				<i>BA 30</i>			
	<i>R</i>		<i>L</i>		<i>R</i>		<i>L</i>	
	<i>Mean t-value</i>	<i>Activation volume (mm³)</i>	<i>Mean t-value</i>	<i>Activation volume (mm³)</i>	<i>Mean t-value</i>	<i>Activation volume (mm³)</i>	<i>Mean t-value</i>	<i>Activation volume (mm³)</i>
Rooms – City Street					–4.461	375		
Rooms – Cityscape	5.854	5007	5.162	2114	5.768	574	4.677	354
Rooms – Landscape	4.822	2111	4.505	1073				
City Street – Cityscape	5.537	2786	4.868	1496	5.193	900	4.55	659
City Street – Landscape	4.759	2440	4.532	1449	4.694	1169	4.408	705
Cityscape – Landscape								
(Rooms + City Street) – (Cityscape + Landscape)	5.621	4555	5.058	2998	5.367	1115	4.870	942
All Scene – Object	5.182	3929	3.823	3115	6.785	1985	6.027	1894

For each contrast, mean *t*-value and cluster size is reported. The last row of the table presents a contrast between scenes depicting more local 3-D geometry (rooms, city streets) and those with less (cityscapes, landscapes). Positive *t*-statistic values indicate greater activation for the first condition listed compared to the second. Nonsignificant volumes (< 142 mm³) are not shown (blank cell).

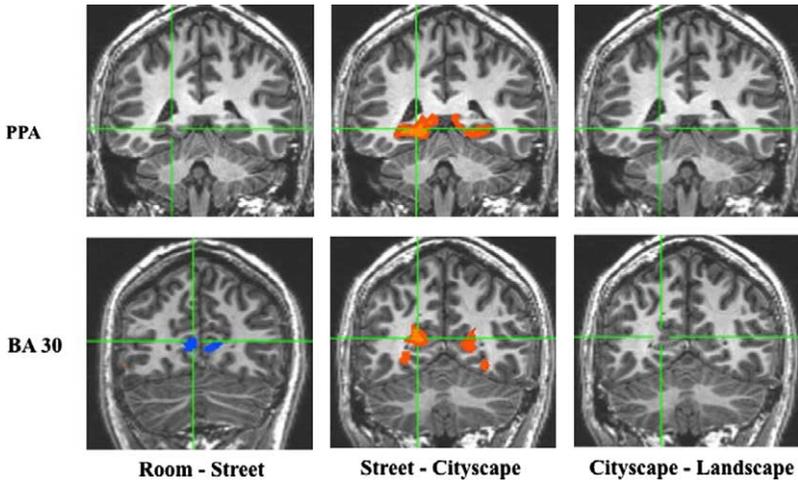


Figure 5. The whole brain contrasts for PPA (middle row) and BA 30 (bottom row). The green crosshairs are centred on the centroid of the right PPA (20, -40, -2) and the peak t -value locations for the corresponding contrasts at the right BA 30 regions for street-cityscape (18, -42, -2) and for room-street (4, -65, 12). The cityscapes and landscapes differed markedly in their power spectra yet produced similar fMRI activation in PPA and RSC, whereas the streets and cityscapes were more similar in the power spectra yet produced markedly different activation levels. To view this figure in colour, please see the online issue of the Journal.

Also in keeping with the PPA ROI data, pPHC was more strongly activated to rooms and city streets than cityscapes or landscapes.

RSC BA 30. As with the ROI data, city streets activated BA 30 more than landscapes and cityscapes, whereas the rooms versus cityscapes comparison reached significance but the rooms versus landscapes comparison did not. There was a small cluster of significant activation for rooms compared to landscapes, but it did not meet the cluster size threshold of 142 mm^3 . Interestingly, there was also a small but significant cluster indicating greater activation for city streets than rooms in this region (Table 1, Figure 5). As in our prior study (Henderson et al., 2007; see also Maguire, 2001), whereas BA 30 showed clear differential activation across several scene types, BA 29 (also associated with RSC) produced only a general scene effect, with greater activation for all scene types compared to objects, but no differences among scene categories.

DISCUSSION

The purpose of this study was to investigate the nature of real-world scene processing in the brain. We were specifically interested in investigating the

scene properties that activate two previously identified cortical areas relevant to scene analysis: PPA and RSC. We compared activation to scenes that did and did not explicitly depict local 3-D volume, scenes that varied in their degree of openness, and artificial versus natural scenes. To implement these comparisons, we contrasted activation in a single experiment to a large sample of photographs of rooms, city streets, cityscapes, and landscapes, using single objects as a control condition. We drew predictions for each hypothesis concerning the function of PPA and RSC for this set of conditions (Figure 2).

The results for both PPA and RSC (more specifically, BA 30, the region of RSC showing differential activation to the scene conditions) align most closely with the predictions of the 3-D local spatial geometry hypothesis. This hypothesis predicts activation to be least for cityscapes and landscapes and greatest for rooms and city streets, because only the latter two conditions explicitly depict the volumetric structure of local enclosed 3-D space. The results were consistent with this prediction. Furthermore, in a contrast that controlled for both the indoor versus outdoor feature (both were outdoor) and general semantic class (both were cities), street scenes produced a clear activation advantage over cityscapes in both PPA and RSC, as predicted by the 3-D geometry hypothesis.

The results of this study are not as easily accommodated by several alternative hypotheses. First, if scene processing regions preferentially activate to scenes depicting more open environments with greater spatial volume, as they might on the theory that these regions play a role in placing a viewer within larger scale space for the purposes of spatial navigation, then we might have expected a ranked ordering of activation with least activation for rooms, followed by streets, then cityscapes, and greatest activation for landscapes. Opposite to this prediction, the more global views of cityscapes and landscapes produced the least activation.

Second, scene regions might preferentially activate to artificial over natural scenes. There is some evidence in the literature hinting that this feature may be relevant for scene processing. For example, our prior results showed greater activation for indoor than outdoor scenes in PPA (Henderson et al., 2007; see also Epstein & Kanwisher, 1998). Those results could have been due to preference for local 3-D geometric structure, but they are also consistent with a preference for artificial over natural scenes. On this hypothesis, rooms, streets, and cityscapes should all have produced more activation than landscapes. In the present study, the comparison of cityscapes and landscapes provided a direct test of this hypothesis. This contrast is particularly strong because the cityscape and landscape conditions controlled a number of other factors including indoor/outdoor (both are outdoor) and general volume (both are panoramic). Furthermore, both of these scene

categories differed in their Fourier power spectra (Figure 6), with the artificial scenes showing more power in horizontal and vertical orientations as has typically been found for artificial versus natural scenes (Torralba & Oliva, 2003; Vailaya et al., 1998). Contrary to the hypothesis that artificial/natural is an important dimension for either PPA or RSC, cityscapes and landscapes produced equivalent activation in both regions. Furthermore, the finding that rooms and city streets produced more activation than cityscapes (which did not differ from landscapes) suggests that it is not simply the presence of greater Fourier power in horizontal and vertical orientations in the images (Figure 6) that drives activation. The cityscapes and landscapes differed markedly in their power spectra yet produced similar fMRI activation in PPA and RSC, whereas the streets and cityscapes were more similar in the power spectra yet produced markedly different activation levels. It seems that something about the structure of higher order features, such as the fact that they create a local surrounding 3-D space, is critical.

Third, if PPA reflects the number of explicitly depicted objects, then rooms should have produced more activation than streets, cityscapes, and landscapes. The finding that rooms, with their large numbers of explicitly depicted objects, produced equivalent activation to city streets (with few if any explicit objects) argues against this hypothesis. In fact, the BA 30 results were opposite to this prediction, with the whole brain analysis showing greater activation to streets than to rooms.

Together then, the complete pattern of data is most consistent with the hypothesis that PPA and RSC are involved in analysing the geometry of local volumetric space defined by surfaces that enclose that space. Furthermore, in the present study a very similar pattern of activation was observed in PPA and BA 30, with rooms and streets producing more activation than cityscapes and landscapes in both regions. However, in the whole brain analysis of BA 30, there was also a significant activation advantage for streets over rooms. (This trend was not reliable in the ROI analysis, which had less power to detect the difference because it averaged over both active and inactive voxels in the anatomically defined ROI region.) The advantage

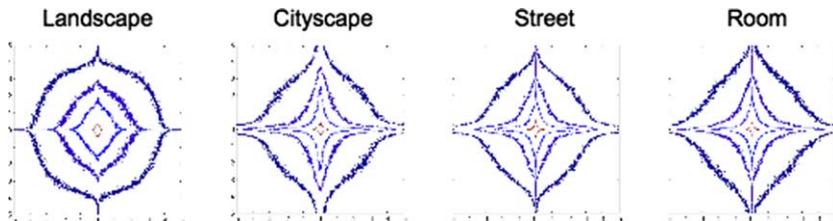


Figure 6. Fourier power spectra for the scenes presented in each scene category.

for streets over rooms suggests that BA 30 may preferentially activate to more open local 3-D volume than PPA, a result that is consistent with the hypothesis that RSC is more involved in analysis of the global environment related to navigation (Aguirre & D'Esposito, 1997; Burgess, Becker, King, & O'Keefe, 2001; Maguire, 2001; Rosenbaum, Ziegler, Winocur, Grady, & Moscovitch, 2004).

In a related study we observed an interaction between region (PPA vs. BA 30) and scene type (indoor vs. outdoor), with greater activation to indoor than outdoor scenes in PPA but not in RSC (Henderson et al., 2007). The indoor and outdoor scenes in that study were similar to the rooms and landscapes in the present study. To investigate whether a similar pattern was produced here, we conducted an ANOVA on the subject data focusing on the room and landscape scenes across the PPA and BA 30 ROIs. This analysis revealed a significant Condition \times ROI interaction ($p < .03$) with a larger room–landscape advantage in PPA than in BA 30. This result also suggests that local 3-D structure may play more of a role in defining PPA activation than BA 30 activation. Epstein and Higgins (2006) also reported activation differences across PPA and RSC consistent with this distinction.

One potential alternative explanation for the activation advantage in PPA and BA 30 for scenes with explicit local 3-D structure is that these scenes might preferentially engage the subject's attention. Contrary to this hypothesis, to the extent that different scene categories generate differential levels of interest, past evidence suggests that more open natural outdoor scenes are more engaging than the more closed artificial scenes that typically depict 3-D structure (Biederman & Vessel, 2006; Vessel & Biederman, 2002), potentially because of the value of the former in evolutionary terms (Kaplan, 1992). Therefore, on the attention hypothesis, we would expect the landscape scenes to produce the greatest activation, opposite to the results obtained here. Furthermore, without a theoretically motivated reason to propose a priori that scenes with 3-D geometry preferentially engage attention, the attention explanation seems circular.

CONCLUSION

The present study demonstrates that two regions of cortex involved in high-level scene processing preferentially activate to room and city street scenes compared with cityscapes and landscapes. These results are most compatible with the hypothesis that these dedicated scene regions support processing of the local 3-D environment. This interpretation in turn suggests that an important way in which scenes differ from other visual categories is that they are views of the local 3-D environment within which a viewer is often embedded. The results also provide additional support for some degree of

dissociation in function between two scene processing areas, with RSC responding more to scenes with larger depicted 3-D volume.

REFERENCES

- Aguirre, G., & D'Esposito, M. (1997). Environmental knowledge is subserved by separate dorsal/ventral areas. *Journal of Neuroscience*, *17*, 2512–2518.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*, 617–629.
- Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, *38*, 347–358.
- Bar, M., Aminoff, E., & Ishai, A. (2008). Famous faces activate contextual associations in the parahippocampal cortex. *Cerebral Cortex*, *18*(6), 1233–1238.
- Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *Journal of Neuroscience*, *28*, 8539–8544.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.
- Biederman, I., & Vessel, E. (2006). Perceptual pleasure and the brain. *The American Scientist*, *94*, 248–255.
- Burgess, N., Becker, S., King, J. A., & O'Keefe, J. (2001). Memory for events and their spatial context: Models and experiments. *Philosophical Transactions of the Royal Society of London*, *356B*, 1493–1503.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance images. *Computers and Biomedical Research*, *29*, 162–173.
- Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152.
- Epstein, R., Graham, K., & Downing, P. (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., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Epstein, R., & Ward, E. J. (2010). How reliable are visual context effects in the parahippocampal place area? *Cerebral Cortex*, *20*, 294–303.
- Epstein, R. A. (2005). The cortical basis of visual scene processing. *Visual Cognition*, *12*, 954–978.
- Epstein, R. A., & Higgins, J. S. (2006). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cerebral Cortex*, *17*, 1680–1693.
- 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. A., Parker, W. E., & Feiler, A. M. (2007). Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *Journal of Neuroscience*, *27*, 6141–6149.
- Goh, J. O. S., Soon, C. S., Park, D., Gutchess, A., Hebrank, A., & Chee, M. W. L. (2004). Cortical areas involved in object, background, and object-background processing revealed with functional magnetic resonance adaptation. *Journal of Neuroscience*, *24*, 10223–10228.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203.

- Grill-Spector, K., & Malach, R. (2001). fMR-adaption: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321.
- Henderson, J. M. (2003). Human gaze control in real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504.
- Henderson, J. M. (2005). *Real-world scene perception*. New York, NY: Psychology Press.
- Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, *16*, 219–222.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.
- Henderson, J. M., Larson, C. L., & Zhu, D. C. (2007). Cortical activation to indoor versus outdoor scenes: An fMRI study. *Experimental Brain Research*, *179*, 75–84.
- Henderson, J. M., Larson, C. L., & Zhu, D. C. (2008). Full scenes produce more activation than close-up scenes and scene-diagnostic objects in parahippocampal and retrosplenial cortex: An fMRI study. *Brain and Cognition*, *66*, 40–49.
- Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 604–610.
- Kaplan, S. (1992). Environmental preference in a knowledge-seeking, knowledge-using organism. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 581–598). New York, NY: Oxford University Press.
- Kim, D., Adalsteinsson, E., Glover, G., & Spielman, D. (2002). Regularized higher-order in vivo shimming. *Magnetic Resonance in Medicine*, *48*, 715–722.
- Kohler, S., Crane, J., & Milner, B. (2002). Differential contributions of the parahippocampal place area and the anterior hippocampus to human memory for scenes. *Hippocampus*, *12*, 718–723.
- Li, F. F., van Rullen, R., Koch, C., & Perona, P. (2002). Why does natural scene categorization require little attention? *Visual Cognition*, *12*, 893–924.
- Maguire, E. A. (2001). The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology*, *42*, 225–238.
- Maguire, E. A., Frith, C. D., & Cipolotti, L. (2001). Distinct neural systems for the encoding and recognition of topography and faces. *Neuroimage*, *13*, 743–750.
- Parrish, T., Gitelman, D., LaBar, K., & Mesulam, M. (2000). Impact of signal-to-noise on functional MRI. *Magnetic Resonance in Medicine*, *44*, 925–932.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 509–522.
- Rajimehr, R., Devaney, K., Young, J., Postelnicu, G., & Tootell, R. (2008). The “parahippocampal place area” responds selectively to high spatial frequencies in humans and monkeys. *Journal of Vision*, *8*(6), 85, 85a.
- Rosenbaum, R. S., Ziegler, M., Winocur, G., Grady, C. L., & Moscovitch, M. (2004). “I have often walked down this street before”: fMRI studies on the hippocampus and other structures during mental navigation of an old environment. *Hippocampus*, *14*, 826–835.
- Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get the “gist” of real-world natural scenes? *Visual Cognition*, *12*, 852–877.
- 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.
- Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipital-temporal cortex? An fMRI study. *Neuron*, *35*(6), 1157–1165.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3D proportional system*. New York, NY: Georg Thieme Verlag.
- Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.

- Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems*, *14*, 391–412.
- Vailaya, A., Jain, A., & Jiang Shang, H. (1998). On image classification: City images vs. landscapes. *Pattern Recognition*, *31*, 1921–1935.
- Vessel, E. A., & Biederman, I. (2002, May). *An fMRI investigation of preference habituation*. Paper presented at the Annual Meeting of the Vision Science Society, Sarasota, FL.
- 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. *Journal of Neuroscience*, *29*, 10573–10581.
- Ward, B. (2000). *Simultaneous inference for fMRI data*. Milwaukee, WI: Biophysics Research Institute, Medical College of Wisconsin.
- Ward, B. (2002). *Deconvolution analysis of fMRI time series data*. Milwaukee, WI: Biophysics Research Institute, Medical College of Wisconsin.

Manuscript received June 2011

Manuscript accepted June 2011