

Dissociable Neural Systems for Recognizing Places and Navigating through Them

 Andrew S. Persichetti and  Daniel D. Dilks

Department of Psychology, Emory University, Atlanta, Georgia 30322

When entering an environment, we can use the present visual information from the scene to either recognize the kind of place it is (e.g., a kitchen or a bedroom) or navigate through it. Here we directly test the hypothesis that these two processes, what we call “scene categorization” and “visually-guided navigation”, are supported by dissociable neural systems. Specifically, we manipulated task demands by asking human participants (male and female) to perform a scene categorization, visually-guided navigation, and baseline task on images of scenes, and measured both the average univariate responses and multivariate spatial pattern of responses within two scene-selective cortical regions, the parahippocampal place area (PPA) and occipital place area (OPA), hypothesized to be separably involved in scene categorization and visually-guided navigation, respectively. As predicted, in the univariate analysis, PPA responded significantly more during the categorization task than during both the navigation and baseline tasks, whereas OPA showed the complete opposite pattern. Similarly, in the multivariate analysis, a linear support vector machine achieved above-chance classification for the categorization task, but not the navigation task in PPA. By contrast, above-chance classification was achieved for both the navigation and categorization tasks in OPA. However, above-chance classification for both tasks was also found in early visual cortex and hence not specific to OPA, suggesting that the spatial patterns of responses in OPA are merely inherited from early vision, and thus may be epiphenomenal to behavior. Together, these results are evidence for dissociable neural systems involved in recognizing places and navigating through them.

Key words: categorization; navigation; occipital place area; parahippocampal place area; retrosplenial complex; scene recognition

Significance Statement

It has been nearly three decades since Goodale and Milner demonstrated that recognizing objects and manipulating them involve distinct neural processes. Today we show the same is true of our interactions with our environment: recognizing places and navigating through them are neurally dissociable. More specifically, we found that a scene-selective region, the parahippocampal place area, is active when participants are asked to categorize a scene, but not when asked to imagine navigating through it, whereas another scene-selective region, the occipital place area, shows the exact opposite pattern. This double dissociation is evidence for dissociable neural systems within scene processing, similar to the bifurcation of object processing described by Goodale and Milner (1992).

Introduction

We can recognize a scene (e.g., a kitchen or a beach) within a fraction of a second, even if we have never seen that particular

place before (Potter, 1976; Thorpe et al., 1996), and almost simultaneously seamlessly navigate through it. Given the ecological importance of scene processing, it is not surprising then that particular regions of the brain respond selectively to visual scene information. Two such regions are the parahippocampal place area (PPA; Epstein and Kanwisher, 1998), and the occipital place area (OPA; Dilks et al., 2013). Although these regions are reliably scene-selective across many studies, the precise function each region plays in scene processing remains unknown.

Several recent studies, however, may provide some clues to their functions. For example, the OPA has been shown to represent information necessary for navigating the immediately visible environment (i.e., “sense” or direction, egocentric distance, first-person perspective motion, boundaries and obstacles, and potential paths for movement in one’s local environment; Dilks et al.,

Received May 11, 2018; revised Sept. 19, 2018; accepted Sept. 24, 2018.

Author contributions: A.S.P. wrote the first draft of the paper; A.S.P. and D.D.D. edited the paper; A.S.P. and D.D.D. designed research; A.S.P. performed research; A.S.P. analyzed data; A.S.P. and D.D.D. wrote the paper.

This work was supported by Emory College, Emory University (D.D.D.), National Eye Institute Grant T32 EY007092 (A.S.P.), and a National Science Foundation Graduate Research Fellowship DGE-1444932 (A.S.P.). We thank the FERN Imaging Center in the Department of Psychology, Emory University, Frederik Kamps for insightful conversation, and Samuel K. Weiller for technical support.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Daniel D. Dilks, Department of Psychology, Emory University, Atlanta, GA 30309. E-mail: ddilks@emory.edu.

<https://doi.org/10.1523/JNEUROSCI.1200-18.2018>

Copyright © 2018 the authors 0270-6474/18/3810295-10\$15.00/0



Figure 1. Example images from the experimental scans. The correct answers for the categorization task (from left to right) are bedroom, kitchen, living room. The correct answers for the navigation task (from left to right) are “left”, “center”, “right”. Note that each task was performed on the exact same image.

2011; Julian et al., 2016; Kamps et al., 2016a,b; Persichetti and Dilks, 2016; Bonner and Epstein, 2017). By contrast, PPA does not represent such navigationally-relevant information, but instead represents the “spatial layout” of scenes, namely relative length and angle information, the geometric properties classically associated with object recognition, but only in the context of large extended surfaces that compose the layout of a scene (Epstein and Kanwisher, 1998; Kamps et al., 2016a; Dillon et al., 2018). Based on these results, we propose that human visual scene processing is composed of at least two distinct systems: one responsible for scene categorization (e.g., recognizing a place as a kitchen or a beach), including PPA, and another responsible for visually-guided navigation (i.e., navigating one’s immediately visible environment), including OPA.

Using fMRI in human adults, we directly test the hypothesis that visual scene processing is composed of dissociable neural systems by using task demands to selectively modulate activity in each system. Specifically, participants viewed images of scenes (i.e., bedrooms, kitchens, and living rooms), and performed a scene categorization, visually-guided navigation, and “baseline” task on each image during separate blocks (Fig. 1). We then analyzed both the average univariate, and the multivariate spatial pattern of responses within each region. We predicted that the average univariate response in PPA would be greater during the categorization task than both the navigation and baseline tasks, and a linear support vector machine would be able to correctly classify (above chance) the categorization task, but not the navigation task. By contrast, we predicted that the average univariate response in OPA would be greater during the navigation task than both the categorization and baseline tasks, and a linear support vector machine would be able to correctly classify the navigation task, but not the categorization task. For completeness, we also examined both the univariate and multivariate responses within another scene-selective region, the retrosplenial complex (RSC; Maguire, 2001), thought to be involved in navigation through the broader environment (Burgess et al., 2001; Wolbers and Büchel, 2005; Vann et al., 2009; Auger et al., 2012, 2015; Marchette et al., 2014), which is different from the kind of navigation tested here. Thus, we predict that there will be no differences in responses between the navigation and categorization tasks in RSC. Crucially, all tasks were completed on the exact same stimuli, task-relevant features for both tasks were contained in the same part of the visual field, behavioral performance was matched between tasks, and fixations while in the scanner were similar between the tasks, ensuring that any neural differences between tasks were not due to differences in low-level visual stimuli, attention, task difficulty, or eye movements.

Materials and Methods

Participants. Twenty participants (Age: 20–36; 8 females) were recruited from the Emory University community. All participants gave informed consent and had normal or corrected-to-normal vision.

Experimental design. For our primary analysis, we used a region-of-interest (ROI) approach in which we localized scene-selective regions (Localizer runs) and then used an independent set of runs to investigate both the univariate and multivariate responses of these regions during blocks of a categorization, a navigation, and a baseline task (Experimental runs). As a secondary analysis, we also conducted both a univariate and a multivariate analysis at the group-level to explore responses to the Experimental runs across the entire slice prescription; described further in the Data analysis section.

For the Localizer runs, we used a blocked design in which participants viewed images of faces, objects, scenes, and scrambled objects, as previously described (Epstein and Kanwisher, 1998). Each participant completed two runs. Each run was 336 s long and consisted of four blocks per stimulus category. The order of the stimulus category blocks in each run was randomized across runs. Each block contained 20 images randomly drawn from the same category. Each image was presented for 300 ms, followed by a 500 ms interstimulus interval (ISI) for a total of 16 s per block. We also included five 16 s fixation blocks: one at the beginning, three in the middle interleaved between each set of stimulus blocks, and one at the end of each run. Participants performed a one-back task, responding every time the same image was presented twice in a row.

For the Experimental runs, we used a block design in which participants completed 10 runs that included blocks of scene images that were divided by two tasks: categorization and navigation. A subset of 11 participants also completed blocks of a one-back task (i.e., the baseline task), in which they simply responded whether the current image was the same or different from the previous image. Each run included three blocks of each task presented in a random order. Each block was 24 s long and consisted of 16 images presented on a neutral gray background. Each image was presented for 500 ms followed by a 1 s ISI. Participants could respond at any time during the stimulus presentation or the subsequent ISI.

During the categorization task, participants were asked to imagine they were standing in the room, and indicated via button press whether the current room was a bedroom, kitchen, or living room. During the navigation task, participants were asked to imagine they were walking on a continuous path through the room, and indicated via a button press whether they could leave through the door on the left, center, or right wall. The navigation task was designed in this way to simulate real-world navigation through the environment, similar to walking on a sidewalk, and not the grass, or on a clear path through a cluttered space. Furthermore, although our navigation task did not actually require participants to navigate through the environment, we were confident that having participants simply imagine navigating through the rooms would be sufficient for activation in regions involved in visually-guided navigation because we designed the task based on object processing experiments that used similarly “passive” tasks to elicit activation in dorsal regions respon-

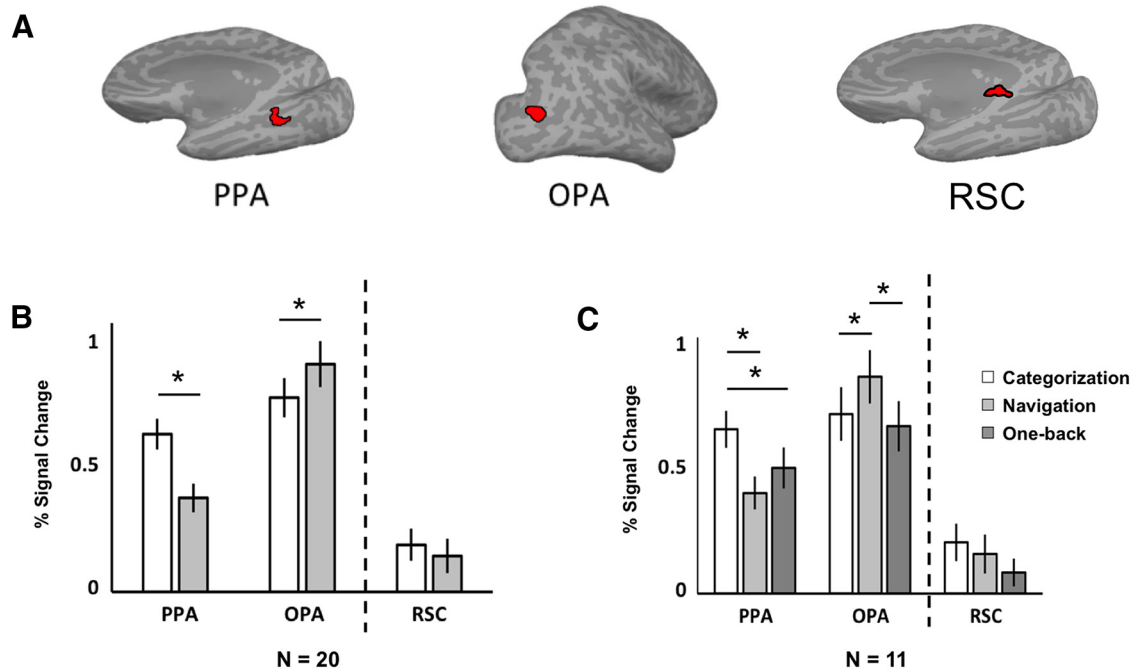


Figure 2. *A*, Example ROIs, labeled accordingly, from one participant, displayed on the cortical surface. Specifically, using independent data from the localizer runs, PPA, OPA, and RSC were localized bilaterally as regions that responded more strongly to scenes than objects ($p < 10^{-4}$, uncorrected). *B*, The average neural response (percentage signal change) to the categorization and navigation tasks for each independently defined scene-selective ROI, labeled accordingly. *C*, The average neural response (percentage signal change) to the categorization, navigation, and one-back tasks for each independently defined scene-selective ROI in 11 of 20 participants, labeled accordingly. Error bars represent ± 1 SEM. Asterisks indicate a significant difference in responses between conditions ($p < 0.05$).

sible for the control of actions directed at objects (Chao and Martin, 2000; Okada et al., 2000; Johnson-Frey et al., 2005). During the baseline task, participants indicated via button press whether the presented image was the exact same or different from the previous image. Fixation blocks were interleaved between each experimental block, and one appeared at the beginning of each run, and another at the end, for a total of 10 fixation blocks per run. During fixation blocks, a white central fixation cross, subtending 0.5° of visual angle, was displayed on a neutral gray background for 6 s. Before each experimental block, the white fixation cross was replaced by instructions indicating which task was about to start and the appropriate responses for that task (white letters on the same neutral gray background) for 4 s. Participants were instructed to remain fixated on the central fixation cross that was presented on the screen at all times other than during the instruction screen. Eye tracking was used to ensure participants remained fixated on the central fixation cross throughout the experiment.

The stimuli were made using the Sims 3 software (Electronic Arts). Each stimulus was a computer-generated image, $12^\circ \times 9^\circ$ of visual angle in size, of a room that included furniture, a door on each visible wall, a continuous path that started at the “front” of the room and led to only one of the doors, and a white central fixation cross that subtended 0.5° of visual angle. Given the known retinotopic biases across the scene-selective regions (Silson et al., 2015), and thus to ensure that any neural or behavioral differences across tasks cannot be attributed to a difference in attention to different areas of the image, all of the navigationally-relevant information (i.e., the breaks in the path, the connections to the doors, and the doors) and the categorization-relevant information (i.e., the furniture) were contained between 3.25° – 7.5° (both calculated from the bottom of the image).

Before data collection, participants watched videos of first-person perspective motion through a room to each of the three doors, to ensure that they understood how to imagine themselves walking through the rooms. Next, participants completed training sessions in which they practiced both the navigation and categorization tasks until they were at least 80% accurate on both tasks. Following training, all participants reported being able to imagine walking along a path toward a door on each trial.

MRI scan parameters. Scanning was done on a 3T Siemens Trio scanner at the Facility for Education and Research in Neuroscience (FERN) at Emory University. Functional images were acquired using a 32-channel head matrix coil and a gradient echo single-shot echoplanar imaging sequence. Twenty-eight slices were acquired for both the Localizer runs, and the Experimental runs. For all runs: repetition time = 2 s; echo time = 30 ms; flip angle 90° ; voxel size = $1.5 \times 1.5 \times 2.5$ mm with a 0.2 mm interslice gap; and slices were oriented approximately between perpendicular and parallel to the calcarine sulcus, covering the occipital and temporal lobes, and most of the parietal and frontal lobes. Whole-brain, high-resolution T1-weighted anatomical images (repetition time = 1900 ms; echo time = 2.27 ms; inversion time = 900 ms; voxel size = $1 \times 1 \times 1$ mm) were also acquired for each participant for registration of the functional images.

Data analysis. Analysis of the fMRI data were conducted using the FSL software (Smith et al., 2004) and custom MATLAB code. The FreeSurfer software (Dale et al., 1999) was used to register the volumetric data from the group analysis to the surface space for visualization. Before statistical analysis, images were skull-stripped (Smith, 2002), and registered to the participants’ T1 weighted anatomical image (Jenkinson et al., 2002). Additionally, data were spatially smoothed (5 mm kernel) for the univariate analysis, but not the multivariate analysis, de-trended, and then fit with a general linear model that contained covariates that were convolved with a double-gamma function to approximate the hemodynamic response function. After preprocessing, scene-selective ROIs were bilaterally defined in each participant using data from the independent Localizer runs as those regions that responded more strongly to scenes than objects ($p < 10^{-4}$, uncorrected), following the method of Epstein and Kanwisher (1998). For the univariate analysis, both PPA and OPA were defined in in at least one hemisphere in all 20 participants (Fig. 2A). Next, for each ROI of each participant, the average response across voxels for each task was extracted and converted to percentage signal change relative to a fixation baseline. Finally, repeated-measures ANOVAs were performed on the neural responses for each ROI. A 2 (ROI: PPA, OPA) \times 2 (hemisphere: Left, Right) \times 2 (task: Categorization, Navigation) repeated-measures ANOVA did not reveal a significant ROI \times hemi-

sphere \times task interaction ($F_{(1,18)} = 3.46, p = 0.08, \eta_p^2 = 0.16$). Thus, data from each hemisphere were collapsed for all further analyses. However, given the marginally significant p value from this test ($p = 0.08$), we also report the results for each hemisphere separately for both the OPA and PPA. For completeness, we also functionally defined RSC in at least one hemisphere in 19 of 20 participants. Finally, we anatomically defined the entirety of V1 (fovea and periphery) using the Jülich histological atlas in FSL (Amunts et al., 2000). We then used the calcarine sulcus as an anatomical landmark to bisect V1 into ventral and dorsal ROIs (V1v and V1d).

For the multivariate analysis, we first identified the 100 most scene-selective voxels in algorithmically identified regions corresponding to the three scene-selective regions of cortex (Julian et al., 2012) in each hemisphere of the unsmoothed data, using the contrast of scenes $>$ objects from the Localizer runs in each of the 11 participants that completed all three tasks. We then extracted the pattern of responses for all three tasks from the GLM that was run on the Experimental data (i.e., the pattern of β weights from the GLM run on the fMRI data), and then normalized the pattern corresponding to each task by subtracting out the grand mean (i.e., the mean response to all tasks) from each voxel in each ROI. Next, in separate regression models, we fit the voxelwise pattern of responses to the navigation and categorization tasks with the voxelwise pattern of responses to the baseline task within each ROI to obtain the residual values to both the navigation and categorization tasks at each voxel, such that the spatial patterns of responses in each ROI are now the residual patterns for the navigation and categorization tasks after accounting for the response to the baseline task. We then trained a linear support vector machine (SVM) to classify the type of task based on the spatial pattern of these residuals. Specifically, the SVM was trained on nine runs of the data and then tested on the tenth run for all possible combinations of runs, and then the classification results across tests were averaged together. The classification accuracies for both tasks were then directly compared with chance performance (i.e., 50% accuracy) using one-sample t tests.

In addition to the ROI analyses described above, we performed both a univariate and multivariate group-level analysis to explore responses to the Experimental runs across the entire slice prescription. These analyses were conducted using the same parameters as were used in the ROI analyses, except that the data from each participant were registered to standard stereotaxic (MNI) space. For the univariate analysis, we averaged the data for each task separately across participants, and then looked for voxels that were more active for one task than the other (i.e., Navigation $>$ Categorization, Categorization $>$ Navigation). For the multivariate analysis, we moved a cubic searchlight comprised of 125 voxels (5^3) across the brain (Kriegeskorte et al., 2006), so that it centered on each voxel in turn. At each voxel, we first trained a linear SVM to classify each task based on the voxelwise pattern of residuals for the navigation and categorization tasks after accounting for the response to the baseline task. Next, we obtained the difference in classification ability across the searchlight voxels between the two tasks (i.e., Navigation $>$ Categorization, Categorization $>$ Navigation). The SVM searchlight analysis was performed in the native space of each participant and then the results were registered to MNI space and all participants' data were averaged together. For each contrast (Navigation $>$ Categorization, Categorization $>$ Navigation) in both the univariate and multivariate analyses, we performed a nonparametric one-sample t test using the FSL randomize program (Smith and Nichols, 2009) with default variance smoothing of 5 mm, which tests the t value at each voxel against a null distribution generated from randomly flipping the sign of participants' t statistic maps 5000 random permutations of group membership in the univariate analysis, and 2048 random permutations of group membership in the multivariate analysis (i.e., the total number of possible permutations across the 11 participants included in the multivariate analysis). The resultant statistical maps were then corrected for multiple comparisons ($p < 0.01$, FWE) using threshold-free cluster enhancement (TFCE; Smith and Nichols, 2009).

Eye tracking. Eye movements were recorded monocularly using a scanner-mounted Eyelink 1000 Plus eye tracker (SR Research) with a sampling rate of 500 Hz. The default nine-point calibration and validation sequences were repeated before each run during the experiment. We

were unable to collect data from the eyetracker in 4 of the 20 participants who completed the navigation and categorization tasks only, and in 2 of the 11 participants who completed all three tasks due to technical difficulties. However, the fMRI and behavioral data from these participants did not differ qualitatively from the rest of the participants, and thus were included in the analysis. We created interest areas for each degree of visual angle away from the central fixation cross, and then calculated the total duration time of fixations within each interest area for each task. We then directly compared these duration values using a repeated-measures ANOVA.

Results

Univariate differences between categorization and navigation in scene-selective cortex

As predicted, a 2 (ROI: PPA, OPA) \times 2 (task: Categorization, Navigation) repeated-measures ANOVA revealed a significant interaction ($F_{(1,19)} = 110.54, p < 10^{-5}, \eta_p^2 = 0.85$), and paired-sample t tests showed that the PPA responded significantly more during the categorization task compared with the navigation task ($t_{(19)} = 8.31, p < 10^{-5}, d = 0.96$), whereas OPA responded significantly more during the navigation task compared with the categorization task ($t_{(19)} = 2.94, p < 0.01, d = 0.35$; Fig. 2*A,B*). The PPA responded more to the categorization task compared with the navigation task in all 20 participants, and the OPA responded more to the navigation task compared with the categorization task in 17 of 20 participants. These results reveal a double dissociation between the responses in PPA and OPA, and thus strongly suggest distinct neural systems selectively involved in scene categorization and visually-guided navigation.

In addition to the PPA and OPA, we defined another well known scene-selective region, the RSC (Fig. 2*A*), which is thought to support navigational processes that require the integration of representations of the local scene with representations of the broader environment (Burgess et al., 2001; Wolbers and Büchel, 2005; Vann et al., 2009; Auger et al., 2012, 2015; Marchette et al., 2014); a different type of navigation from the visually-guided navigation tested here. Thus, because neither the categorization nor the navigation task reflects the primary role of RSC in navigation through the broader environment, we predicted that RSC would respond similarly to both tasks. A paired-samples t test confirmed this prediction ($t_{(18)} = 0.92, p = 0.37, d = 0.13$; Fig. 2*B*). Note, however, these results should be interpreted with caution given this null result. Next, directly comparing across ROIs, a 3 (ROI: PPA, OPA, RSC) \times 2 (task: Categorization, Navigation) repeated-measures ANOVA revealed a significant ROI \times task interaction ($F_{(2,36)} = 67.05, p < 10^{-12}, \eta_p^2 = 0.79$), with RSC responding similarly to both tasks relative to PPA and OPA (interaction contrasts, both p values $< 10^{-4}$, both η_p^2 values > 0.61), and PPA and OPA still showing the opposite patterns of response (interaction contrast, $p < 10^{-4}, \eta_p^2 = 0.85$).

Next, investigating each hemisphere separately (for details, see Materials and Methods), we again found a double dissociation between the responses in the right PPA and right OPA, yet a single dissociation between the responses in the left PPA and left OPA. Specifically, in the right hemisphere, a 2 (ROI: right PPA, right OPA) \times 2 (task: Categorization, Navigation) repeated-measures ANOVA revealed a significant interaction ($F_{(1,19)} = 99.59, p < 10^{-5}, \eta_p^2 = 0.84$), with right PPA responding significantly more during the categorization task compared with the navigation task ($t_{(19)} = 7.98, p < 10^{-5}, d = 0.82$), and right OPA responding significantly more during the navigation task compared with the categorization task ($t_{(19)} = 4.85, p < 10^{-3}, d = 0.47$). In the left hemisphere, a 2 (ROI: left PPA, left OPA) \times 2 (task: Categoriza-

tion, Navigation) repeated-measures ANOVA revealed a significant interaction ($F_{(1,18)} = 49.28, p < 10^{-5}, \eta_p^2 = 0.73$), with left PPA responding significantly more during the categorization task compared with the navigation task ($t_{(19)} = 7.20, p < 10^{-5}, d = 1.01$), and left OPA responding similarly during both tasks ($t_{(18)} = 1.02, p = 0.32, d = 0.16$). Interestingly, this finding of right lateralization in the OPA during the navigation task is consistent with prior lesion and neuroimaging studies that provided evidence that some navigational processes might be right lateralized (for review, see Maguire, 2001). Note, however, the lack of a significant difference between the responses to the categorization and navigation tasks in the left OPA (a null result) needs to be interpreted with caution, and numerically, at least, the left OPA responded more to the navigation task compared with the categorization task in 11 of 19 participants. Furthermore, in the group analysis (see below), we found both right and left OPA responding more to the navigation task compared with the categorization task. Thus, it is not clear whether the left OPA is or is not involved in visually-guided navigation, or in both visually-guided navigation and categorization, and further research is needed to directly test this hypothesis.

But might attention to different areas of the scene images during the different tasks explain our results. For example, if the different types of information used during the navigation and categorization tasks were located in different areas of the scene images, then participants could simply attend to different areas of the scene during each task. This concern is especially relevant because a recent study reported that the PPA has an upper visual field bias, and the OPA has a lower visual field bias (Silson et al., 2015). Although we do not think that these differences in visual field biases between PPA and OPA can explain our results, because we designed the scene images so that both the categorization-relevant information (i.e., the furniture) and the navigationally-relevant information (i.e., the breaks in the path, the connections to the doors, and the doors) were evenly distributed across the images (for more details, see Materials and Methods), we ran a further test to confirm that differences in visual field biases between PPA and OPA cannot explain our results. Specifically, because activation in V1 is modulated by changes to attention (Brefczynski and DeYoe, 1999; Serences and Boynton, 2007), we split visual area V1 about the calcarine sulcus into a ventral ROI representing the upper visual field (V1v) and a dorsal ROI representing the lower visual field (V1d) in each participant, and then compared responses to the tasks in V1v and PPA (which both have an upper visual field bias), and V1d and OPA (which both have a lower visual field bias). If the results in scene-selective cortex are due simply to differences in attention to the lower and upper visual fields between the two tasks, then responses in PPA should be similar to responses in V1v, whereas responses in OPA should be similar to V1d. Contrary to this prediction, a 2 (ROI: PPA, V1v) \times 2 (task: Categorization, Navigation) repeated-measures ANOVA revealed a significant interaction ($F_{(1,19)} = 14.04, p < 10^{-3}, \eta_p^2 = 0.43$), with V1v responding similarly to both tasks relative to PPA, whereas a separate 2 (ROI: OPA, V1d) \times 2 (task: Categorization, Navigation) repeated-measures ANOVA also revealed a significant interaction ($F_{(1,19)} = 4.82, p < 0.05, \eta_p^2 = 0.20$), with V1d responding similarly to both tasks relative to OPA. Thus, the univariate results in the PPA and OPA are not because of differences in spatial attention between the two tasks. However, an interesting direction for future research is to investigate how the differential task modulation between PPA and OPA found in the current study interact with known differences in the retinotopic biases between these two regions; for

example, how OPA and PPA would respond to task-specific stimuli presented in only the upper or lower visual fields.

Although this analysis suggests that the PPA is involved in scene categorization, but not visually-guided navigation, and the OPA shows the opposite pattern, both regions respond well above a fixation baseline to the non-preferred task. To ensure then that each region does not play a role in the non-preferred task, we need to show that the response during the non-preferred task is similar to the response when the participant is simply viewing a scene in each ROI. To this end, we compared the responses to the categorization and navigation tasks to the response during the one-back task (i.e., a baseline task) in each ROI in the 11 participants that completed all three tasks. We consider this “one-back task” a reasonable baseline comparison for both the categorization task and visually-guided navigation task because the response to this task should be similar to the response when the participant is simply “seeing” a scene (i.e., it is effectively task neutral). Thus, we predict that the preferred task in each region (i.e., the categorization task in PPA, and the visually-guided navigation task in OPA) should elicit a response over and above the response to the task-neutral baseline task. By contrast, the non-preferred task (i.e., the visually-guided navigation task in PPA, and the categorization task in OPA) should elicit a response similar to the baseline task, because that region is not involved in either task. As predicted, a 2 (ROI: PPA, OPA) \times 3 (task: Categorization, Navigation, Baseline) repeated-measures ANOVA revealed a significant interaction ($F_{(2,20)} = 68.57, p < 10^{-5}, \eta_p^2 = 0.87$), with a significant difference between the categorization and navigation tasks between the PPA and OPA (interaction contrast, $p < 10^{-5}, \eta_p^2 = 0.91$), as well as significant differences between the baseline task and both the categorization and navigation tasks (interaction contrasts, both p values < 0.01 , both η_p^2 values > 0.57 ; Fig. 2C), demonstrating the complete functional dissociation between the PPA and OPA. Furthermore, for PPA, a three-level (task: Categorization, Navigation, Baseline) repeated-measures ANOVA revealed a significant main effect of task ($F_{(2,20)} = 8.97, p < 0.01, \eta_p^2 = 0.47$), with a significantly greater response during the categorization task relative to both the navigation and baseline tasks (main effect contrasts: both p values < 0.05 , both Cohen's d values > 0.60), and crucially, no difference between the navigation and baseline tasks (main effect contrast: $p = 0.17, d = 0.41$; Fig. 2C). For OPA, a three-level (task: Categorization, Navigation, Baseline) repeated-measures ANOVA revealed a significant main effect of task ($F_{(2,20)} = 5.57, p < 0.05, \eta_p^2 = 0.36$), with a significantly greater response during the navigation task relative to both the categorization and baseline tasks (main effect contrasts: both p values < 0.05 , both Cohen's d values > 0.37), and crucially, no difference between the categorization and baseline tasks (main effect contrast: $p = 0.44, d = 0.10$; Fig. 2C). These results are strong evidence that the PPA is involved in scene categorization, but not visually-guided navigation, and the OPA is involved in visually-guided navigation, but not scene categorization.

For completeness, in RSC, a three-level (task: Categorization, Navigation, Baseline) repeated-measures ANOVA did not reveal a significant main effect of task ($F_{(2,18)} = 2.35, p = 0.12, \eta_p^2 = 0.21$), thus finding no significant differences between the tasks in RSC (Fig. 2C). Finally, directly comparing all three ROIs, a 3 (ROI: PPA, OPA, RSC) \times 3 (task: Categorization, Navigation, Baseline) repeated-measures ANOVA revealed a significant interaction ($F_{(4,36)} = 24.87, p < 10^{-5}, \eta_p^2 = 0.73$), with a significant difference between the categorization and navigation tasks be-

tween the PPA and OPA (interaction contrast: $p < 10^{-5}$, $\eta_p^2 = 0.91$), as well as significant differences between the baseline task and both the categorization and navigation tasks (interaction contrasts: all p values < 0.01 , all η_p^2 values > 0.54). Further, the difference between the categorization and navigation tasks significantly differed between the RSC and both PPA and OPA (interaction contrasts: both p values < 0.01 , both η_p^2 values > 0.60), suggesting that the RSC is not selectively involved in either task (Fig. 2C).

Multivariate differences between categorization and navigation in scene-selective cortex

Although the results from the univariate analysis were consistent with our prediction that PPA would respond selectively to the categorization task and OPA would respond selectively to the navigation task, it could still be the case that each region represents its non-preferred task at the finer-grained multivoxel pattern level. Further, we also examined the spatial pattern of responses in the RSC, because despite finding no significant differences between the tasks in the univariate analysis, it could still be the case that RSC can discriminate the tasks at a finer-grained multivoxel pattern level. To test this potential finer-grained level of representation, we trained a linear SVM to classify each task based on the voxelwise pattern of residuals for the navigation and categorization tasks after accounting for the response to the baseline task (see Materials and Methods for more details). As predicted, in the PPA the SVM correctly classified the categorization task above chance ($t_{(10)} = 2.41$, $p < 0.05$, $d = 0.73$), but not the navigation task ($t_{(10)} = 0.41$, $p = 0.69$, $d = 0.12$; Fig. 3), revealing that the spatial pattern of responses to the categorization task was more stable and reliable than the spatial pattern of responses to the navigation task. By contrast, and not exactly as predicted, in the OPA, the SVM correctly classified not only the navigation task above chance, but also the categorization task (both p values < 0.01 , both d values > 0.95 ; Fig. 3). Finally, the SVM did not classify either task above chance in the RSC (both p values > 0.70 , d values < 0.12).

Next, we analyzed the results from the linear SVM in V1v and V1d to test whether the spatial patterns of responses in PPA and OPA are merely inherited from upstream, early visual regions. The SVM did not classify either task above chance in V1v (both p values > 0.25 , d values < 0.33), which is a qualitatively different pattern than was found in both PPA and OPA. Interestingly, the SVM correctly classified both tasks above chance in V1d (both p values < 0.05 , d values > 0.84), which is a qualitatively different pattern than was found in PPA, but similar to the results in OPA. These results suggest that the spatial patterns of responses in PPA are not simply inherited from early visual cortex, whereas the spatial patterns of responses in OPA may be, thus it is possible that the spatial patterns of responses (but, crucially, not the univariate responses) in OPA are simply inherited from adjacent early visual regions and are not specific to the OPA, and thus may be epiphenomenal to behavior (Schalk et al., 2017).

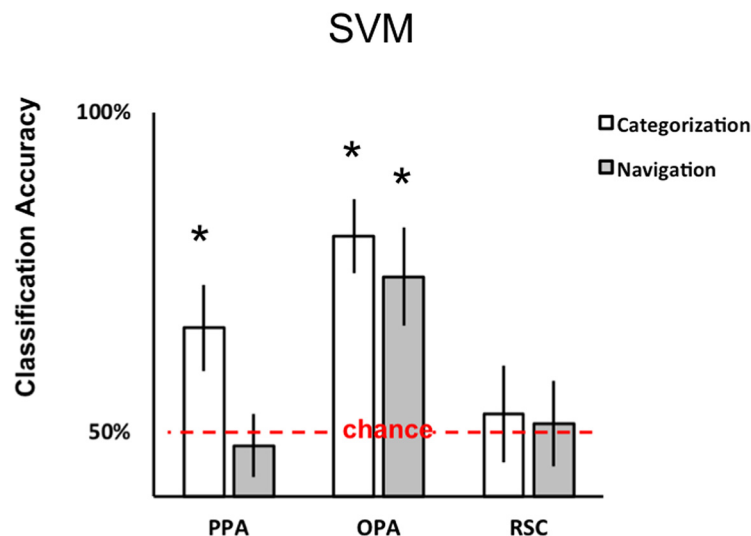


Figure 3. Classification accuracy from the linear SVM in scene-selective ROIs. In PPA, classification accuracy was above chance for the categorization task, but not the navigation task. By contrast, in OPA, classification accuracy was above chance for both tasks. In RSC, classification accuracy was not above chance for either task. Error bars represent ± 1 SEM. Asterisks indicate classification accuracy significantly greater than chance ($p < 0.05$).

Categorization versus navigation in other brain regions?

To investigate whether cortical regions beyond our functionally defined ROIs might be involved in scene categorization or visually-guided navigation, we performed a group-level analysis across the entire slice prescription for both the univariate and multivariate analyses, using a nonparametric permutation test (Smith and Nichols, 2009). For the univariate analysis, we identified regions that responded more during the categorization task than during the navigation task, and regions that responded more during the navigation task than during the categorization task. Consistent with the results from the univariate ROI analysis, at a threshold of $p < 0.01$ (FWE corrected), we found regions that responded more during the categorization task than the navigation task overlapping the functionally defined left and right PPA (defined in a comparable group-level contrast of scene $>$ objects from the Localizer runs) that extended from the posterior parahippocampal gyrus to the lingual and fusiform gyri (Fig. 4). No other brain regions emerged from this contrast. We also found regions that responded more during the navigation task than during the categorization task overlapping the functionally defined left and right OPA (again, defined in a comparable group-level contrast of scene $>$ objects from the Localizer runs) that were adjacent to the parieto-occipital and transverse occipital sulci (Fig. 4). In the same contrast, we also found a swath of cortex in both hemispheres that extended from the inferior to superior parietal lobe, which overlapped with a functionally defined region in the superior parietal lobe (SPL, again, defined in a comparable group-level contrast of scene $>$ objects from the Localizer runs) that has been implicated in navigation (Spiers and Maguire, 2007; Marchette et al., 2014; Fig. 4). We propose that this region may be another scene-selective region in the dorsal stream that is involved in navigating one's immediately visible environment. Importantly, we also found regions that responded more during the navigation task than during the categorization task in both the left and right middle temporal gyrus and lateral occipital sulcus that overlapped with the functionally defined motion area MT (defined using a group-level activation map derived from a separate group of participants; Fig. 4). Visual area MT has been implicated in motion perception (Tootell et al., 1995), illusory

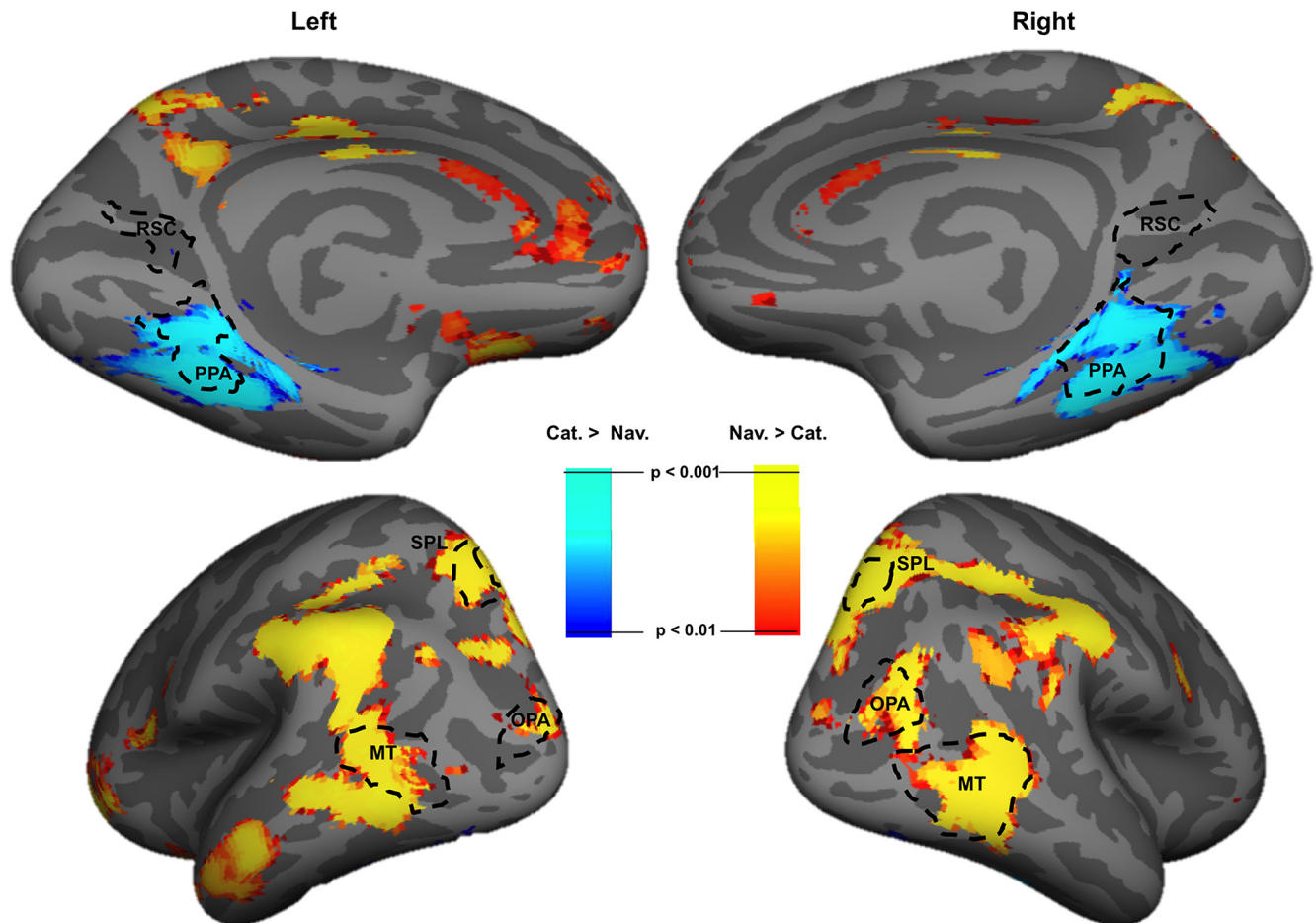


Figure 4. A group-level analysis across the entire slice prescription identified regions that responded more during the categorization task than during the navigation task (Cat > Nav), and regions that responded more during the navigation task than during the categorization task (Nav > Cat). At a threshold of $p < 0.01$ (FWE corrected) we found (1) a bilateral region that responded more to the categorization task than during the navigation task overlapping the group-level functionally defined PPA, and (2) regions that responded more during the navigation task than during the categorization task overlapping the group-level functionally defined OPA. Note the lack of activation to both contrasts in RSC.

self-motion (Kovács et al., 2008), and motion imagery (Kourtzi and Kanwisher, 2000), thus more activation during the navigation task relative to the categorization task suggests that participants were indeed imagining walking along the path toward a door during the navigation task. Neither contrast elicited activation overlapping the functionally defined RSC.

For the multivariate analysis, we used a searchlight technique (Kriegeskorte et al., 2006), in the 11 participants that completed all three tasks. Specifically, at each voxel, we first trained a linear SVM to classify each task based on the voxelwise pattern of residuals for the navigation and categorization tasks after accounting for the response to the baseline task (see Materials and Methods for more details). We then compared the classification accuracy for the categorization task to the classification accuracy for the navigation task. Consistent with the results from the multivariate ROI analysis, we found clusters of voxels that were better at classifying the categorization task than the navigation task within the functionally defined left and right PPA. We also found better classification for the categorization task than the navigation task in lateral occipital cortex, which suggests that, in addition to the PPA, object-selective cortex might represent the furniture in the rooms during the categorization task. Also consistent with the results from the multivariate ROI analysis, we found clusters of voxels that were better at classifying the navigation task than the categorization task within the functionally defined left and right OPA, as well as

clusters that were better at classifying the categorization task than the navigation task. We also found clusters of voxels that were better at classifying the navigation task relative to the categorization task within the functionally defined motion area MT, which is more evidence that participants were indeed imagining walking along the path toward a door during the navigation task. Finally, the SVM searchlight did not identify any voxels able to classify either task in the functionally defined RSC.

Non-task factors cannot explain our results

But might our results be explained by other factors than the task modulation itself (e.g., differences in low-level visual stimuli, task difficulty, or eye movements)? We do not think so for several reasons. First, differences in low-level visual stimuli cannot explain our results because participants completed all three tasks on the exact same stimuli. Second, our findings cannot be because of one task being more difficult than another. During the scan, we found no differences between the navigation and categorization tasks across all 20 participants ($t_{(19)} = 0.26$, $p = 0.80$, $d = 0.07$), or in the subset of 11 participants that completed all three tasks ($t_{(10)} = 1.34$, $p = 0.21$, $d = 0.45$). Further, the subset of 11 participants that completed all three tasks performed nearly identically on all three tasks (89.9% correct on the categorization task, 93.0% correct on the navigation task, and 88.3% on the baseline task). A three-level (task: Categorization, Navigation, Baseline)

repeated-measures ANOVA on the accuracy scores of each participant revealed no significant main effect of condition ($F_{(2,20)} = 2.78, p = 0.09, \eta_p^2 = 0.22$), confirming that there were no differences between any of the tasks. Third, different patterns of eye movements across the three tasks cannot explain our findings. During the scan, participants remained fixated within 2° of the fixation cross 90.7% of the total trial time during the categorization task, 90.6% of the total trial time during the navigation task, and 91.9% of the total trial time during the baseline task. A three-level (task: Categorization, Navigation, Baseline) repeated-measures ANOVA confirmed that fixations did not differ between tasks ($F_{(2,16)} = 0.54, p = 0.59, \eta_p^2 = 0.06$).

Discussion

The current study aimed to selectively modulate separable neural systems for recognizing places and navigating through them by changing task demands to reflect the functions of those systems. We demonstrated in a univariate analysis that the PPA responded selectively to an image of a scene when a viewer was performing a categorization task on the image relative to when the same viewer was performing both a navigation and baseline task on the exact same image, whereas OPA showed the opposite pattern of results. Even at the fine-grained level of multivoxel patterns of responses, we showed that spatial patterns of responses to the categorization task in PPA were classified with above-chance accuracy by a linear SVM, whereas the pattern of responses to the navigation task were not. By contrast, the SVM was able to classify both the navigation and categorization tasks above chance in OPA. However, the SVM was also able to classify both tasks above baseline in V1d, which suggests that the spatial patterns of responses (but, not the univariate responses) in OPA are simply inherited from adjacent early visual regions and are not specific to the OPA, and thus may be epiphenomenal to behavior (Schalk et al., 2017). Together, these univariate and multivariate analyses are evidence for distinct neural systems underpinning our diverse abilities to recognize places and navigate through them. Finally, our findings cannot be explained by differences across tasks with regard to low-level visual properties of the stimuli, attention, task difficulty, or eye movements.

Our finding that PPA responds selectively to the categorization task, and not to the navigation task is evidence against the hypotheses that this region is directly involved in navigating the local environment (Ghaem et al., 1997; Epstein and Kanwisher, 1998; Rosenbaum et al., 2004; Cheng and Newcombe, 2005; Rauchs et al., 2008; Spelke et al., 2010). Although these studies argue that the representation of spatial layout information in PPA is necessary for reorientation and navigation, we contend that such information need not be used for navigation only, and most certainly can be used for the recognition of scene category. Indeed, several behavioral and computer vision studies have found that spatial layout information can be used to categorize scenes, and at times is even necessary for doing so (Oliva and Schyns, 1997; Oliva and Torralba, 2001; Greene and Oliva, 2009; Walther et al., 2011). For example, Walther and Shen (2014) found that participants' ability to categorize scenes was significantly impaired when contour junctions (i.e., angle), and to a lesser extent, contour lengths were removed. The results from these studies are consistent with our finding that PPA represents the local scene for the purpose of recognizing that scene as a member of a particular category, and is distinct from a system solely devoted to visually-guided navigation.

Our finding that OPA responds selectively to the navigation task, and not to the categorization task dovetails with several

recent fMRI studies showing that OPA represents information necessary for navigating one's immediately visible environment (e.g., egocentric information; Dilks et al., 2011; Julian et al., 2016; Kamps et al., 2016a,b; Persichetti and Dilks, 2016; Bonner and Epstein, 2017; Dillon et al., 2018). However, two transcranial magnetic stimulation (TMS) studies (Dilks et al., 2013; Ganaden et al., 2013) found that TMS over OPA impaired categorization accuracy of scenes but not objects, whereas TMS over an object-selective region (the lateral occipital complex) impaired categorization accuracy of objects but not scenes. At first glance, these findings appear inconsistent with the proposed role of OPA in visually-guided navigation, insofar as TMS disrupted a scene categorization task. (Note that these studies never tested a navigation task.) However, we propose that although TMS to OPA disrupted performance on a categorization task, it did so not as a result of the task, but rather as a result of OPA simply seeing a scene, similar to the finding that there was no difference in OPA's average univariate response between the categorization and baseline tasks here. Thus, given the current findings, with TMS over OPA, we would predict significantly worse performance on a navigation task compared with both a categorization task and a baseline task, and no difference between the categorization and baseline tasks. Furthermore, such a result would provide strong evidence that the spatial pattern of responses in OPA are not causally involved in either task, whereas the univariate responses are causally involved in the visually-guided navigation task, as described above.

Together, our findings that PPA responds selectively to the categorization task relative to the navigation task, whereas OPA shows the opposite pattern, suggest that the human visual scene processing system is similar to the well characterized ventral/dorsal distinction for object processing (Trevarthen, 1968; Schneider, 1969; Ingle, 1973; Perenin and Jeannerod, 1979; Ungerleider and Mishkin, 1982; Goodale and Milner, 1992), with one system responsible for recognition (a "what" system) and another for visually-guided action (a "how" system). Note that our finding of two distinct systems for scene processing does not mean that the two systems cannot and do not interact. Indeed, two recent studies (Baldassano et al., 2013; Nasr et al., 2013) found functional correlations between OPA and a posterior portion of PPA, suggesting that these regions are connected (at least functionally, maybe anatomically), thereby enabling crosstalk between them. However, no matter if, or how, these two systems communicate, our results demonstrate that they are functionally dissociable.

Although we have demonstrated that PPA is not involved in navigating through the local environment, it still may be the case that PPA is involved in navigational processes requiring the integration of representations of the local scene with representations of the broader environment. Indeed, several studies found that PPA represents stable and permanent features in the local scene (e.g., specific buildings) as particular landmarks that can then be used to orient oneself within the broader, not immediately visible environment (Epstein, 2008, 2014 for review). Interestingly, in many of these same studies, RSC was also involved in landmark recognition (Morgan et al., 2011; Troiani et al., 2014; Marchette et al., 2015), whereas still other studies found landmark sensitivity in the RSC only, not PPA (Auger et al., 2012, 2015; Marchette et al., 2014). Because the current experiment did not probe landmark recognition, we cannot definitively rule out the possibility that PPA was encoding the scenes as unique places (e.g., "Lauren's kitchen" versus "Kim's kitchen") instead of as members of a general scene category (i.e., "a kitchen") during the categoriza-

tion task. However, we do not think landmark sensitivity can explain the selective response to the categorization task in PPA because, as mentioned above, RSC is also (or exclusively) sensitive to landmarks, and thus would show the same pattern as PPA, but it does not. Future work is needed to tease apart sensitivity to category membership from sensitivity to landmarks in PPA, but either way, it is clear from the current study that PPA is not involved in visually-guided navigation.

Furthermore, our finding that RSC responded similarly during both the categorization and navigation tasks (although essentially a null result) is consistent with the evidence that RSC is involved in navigation through the broader environment (Burgess et al., 2001; Wolbers and Büchel, 2005; Vann et al., 2009; Auger et al., 2012, 2015; Marchette et al., 2014), which is different from the kind of navigation tested here. For instance, our navigation task asked participants to simply imagine walking through a single room on each trial, indicating from which door (left, center, or right) they could leave the room, which did not require situating the local scene within the broader environment. Thus, it is not surprising that RSC responded similarly to both tasks. By contrast, OPA showed a different pattern of responses than RSC, responding more to the navigation task than the categorization task. As such, these data suggest that there are at least two navigation systems: one responsible for visually-guided navigation through the immediate environment (as measured by our task), including OPA, and the other for orienting to the broader environment, including RSC (Kamps et al., 2016a).

In conclusion, we have shown a double dissociation between the univariate responses in the scene-selective PPA and OPA, with PPA responding more when participants were asked to categorize a scene than when asked to navigate through it, whereas OPA shows the opposite pattern. Further, this division of labor was supported by the multivariate spatial pattern of responses in each region. Together, these results provide strong evidence for dissociable neural systems involved in recognizing places and navigating through them.

References

- Amunts K, Malikovic A, Mohlberg H, Schormann T, Zilles K (2000) Brodmann's areas 17 and 18 brought into stereotaxic space—where and how variable? *Neuroimage* 11:66–84. [CrossRef Medline](#)
- Auger SD, Mullally SL, Maguire EA (2012) Retrosplenial cortex codes for permanent landmarks (Baker CI, ed). *PLoS One* 7:e43620. [CrossRef Medline](#)
- Auger SD, Zeidman P, Maguire EA (2015) A central role for the retrosplenial cortex in de novo environmental learning. *eLife* 4:e09031. [CrossRef Medline](#)
- Baldassano C, Beck DM, Fei-Fei L (2013) Differential connectivity within the parahippocampal place area. *Neuroimage* 75:228–237. [CrossRef Medline](#)
- Bonner MF, Epstein RA (2017) Coding of navigational affordances in the human visual system. *Proc Natl Acad Sci U S A* 114:4793–4798. [CrossRef Medline](#)
- Brefczynski JA, DeYoe EA (1999) A physiological correlate of the “spotlight” of visual attention. *Nat Neurosci* 2:370–374. [CrossRef Medline](#)
- Burgess N, Maguire EA, Spiers HJ, O'Keefe J (2001) A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage* 14:439–453. [CrossRef Medline](#)
- Chao LL, Martin A (2000) Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12:478–484. [CrossRef Medline](#)
- Cheng K, Newcombe NS (2005) Is there a geometric module for spatial orientation? squaring theory and evidence. *Psychon Bull Rev* 12:1–23. [CrossRef Medline](#)
- Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. segmentation and surface reconstruction. *Neuroimage* 9:179–194. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Dilks DD, Julian JB, Paunov AM, Kanwisher N (2013) The occipital place area is causally and selectively involved in scene perception. *J Neurosci* 33:1331–1336a. [CrossRef Medline](#)
- Dillon MR, Persichetti AS, Spelke ES, Dilks DD (2018) Places in the brain: bridging layout and object geometry in scene-selective cortex. *Cereb Cortex* 28:2365–2374. [CrossRef Medline](#)
- Epstein RA (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci* 12:388–396. [CrossRef Medline](#)
- Epstein RA (2014) Neural systems for visual scene recognition. In: *Scene vision: making sense of what we see*. Cambridge, MA: MIT.
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601. [CrossRef Medline](#)
- Ganaden RE, Mullin CR, Steeves JK (2013) Transcranial magnetic stimulation to the transverse occipital sulcus affects scene but not object processing. *J Cogn Neurosci* 25:961–968. [CrossRef Medline](#)
- Ghaem O, Mellet E, Crivello F, Tzourio N, Mazoyer B, Berthoz A, Denis M (1997) Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport* 8:739–744. [CrossRef Medline](#)
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [CrossRef Medline](#)
- Greene MR, Oliva A (2009) Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cognit Psychol* 58:137–176. [CrossRef Medline](#)
- Ingle D (1973) Two visual systems in the frog. *Science* 181:1053–1055. [CrossRef Medline](#)
- Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17:825–841. [CrossRef Medline](#)
- Johnson-Frey SH, Newman-Norlund R, Grafton ST (2005) A distributed left hemisphere network active during planning of everyday tool use skills. *Cereb Cortex* 15:681–695. [CrossRef Medline](#)
- Julian JB, Fedorenko E, Webster J, Kanwisher N (2012) An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage* 60:2357–2364. [CrossRef Medline](#)
- Julian JB, Ryan J, Hamilton RH, Epstein RA (2016) The occipital place area is causally involved in representing environmental boundaries during navigation. *Curr Biol CB* 26:1104–1109. [CrossRef Medline](#)
- Kamps FS, Julian JB, Kubilius J, Kanwisher N, Dilks DD (2016a) The occipital place area represents the local elements of scenes. *Neuroimage* 132:417–424. [CrossRef Medline](#)
- Kamps FS, Lall V, Dilks DD (2016b) The occipital place area represents first-person perspective motion information through scenes. *Cortex* 83:17–26. [CrossRef Medline](#)
- Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. *J Cogn Neurosci* 12:48–55. [CrossRef Medline](#)
- Kovács G, Raabe M, Greenlee MW (2008) Neural correlates of visually induced self-motion illusion in depth. *Cereb Cortex* 18:1779–1787. [CrossRef Medline](#)
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [CrossRef Medline](#)
- Maguire E (2001) The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand J Psychol* 42:225–238. [CrossRef Medline](#)
- Marchette SA, Vass LK, Ryan J, Epstein RA (2014) Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nat Neurosci* 17:1598–1606. [CrossRef Medline](#)
- Marchette SA, Vass LK, Ryan J, Epstein RA (2015) Outside looking in: landmark generalization in the human navigational system. *J Neurosci* 35:14896–14908. [CrossRef Medline](#)
- Morgan LK, Macevoy SP, Aguirre GK, Epstein RA (2011) Distances between real-world locations are represented in the human hippocampus. *J Neurosci* 31:1238–1245. [CrossRef Medline](#)
- Nasr S, Devaney KJ, Tootell RB (2013) Spatial encoding and underlying circuitry in scene-selective cortex. *Neuroimage* 83:892–900. [CrossRef Medline](#)
- Okada T, Tanaka S, Nakai T, Nishizawa S, Inui T, Sadato N, Yonekura Y, Konishi J (2000) Naming of animals and tools: a functional magnetic

- resonance imaging study of categorical differences in the human brain areas commonly used for naming visually presented objects. *Neurosci Lett* 296:33–36. [CrossRef Medline](#)
- Oliva A, Schyns PG (1997) Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognit Psychol* 34:72–107. [CrossRef Medline](#)
- Oliva A, Torralba A (2001) Modeling the shape of a scene: a holistic representation of the spatial envelope. *Int J Comput Vis* 42:145–175. [CrossRef](#)
- Perenin MT, Jeannerod M (1979) Subcortical vision in man. *Trends Neurosci* 2:204–207. [CrossRef](#)
- Persichetti AS, Dilks DD (2016) Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex* 77:155–163. [CrossRef Medline](#)
- Potter MC (1976) Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn* 2:509–522. [CrossRef Medline](#)
- Rauchs G, Orban P, Balteau E, Schmidt C, Degueldre C, Luxen A, Maquet P, Peigneux P (2008) Partially segregated neural networks for spatial and contextual memory in virtual navigation. *Hippocampus* 18:503–518. [CrossRef Medline](#)
- Rosenbaum RS, Ziegler M, Winocur G, Grady CL, 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. [CrossRef Medline](#)
- Schalk G, Kapeller C, Guger C, Ogawa H, Hiroshima S, Lafer-Sousa R, Saygin ZM, Kamada K, Kanwisher N (2017) Facephenes and rainbows: causal evidence for functional and anatomical specificity of face and color processing in the human brain. *Proc Natl Acad Sci U S A* 114:12285–12290. [CrossRef Medline](#)
- Schneider GE (1969) Two visual systems. *Science* 163:895–902. [CrossRef Medline](#)
- Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55:301–312. [CrossRef Medline](#)
- Silson EH, Chan AW, Reynolds RC, Kravitz DJ, Baker CI (2015) A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *J Neurosci* 35:11921–11935. [CrossRef Medline](#)
- Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44:83–98. [CrossRef Medline](#)
- Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17:143–155. [CrossRef Medline](#)
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23:S208–S219. [CrossRef Medline](#)
- Spelke E, Lee SA, Izard V (2010) Beyond core knowledge: natural geometry. *Cogn Sci* 34:863–884. [CrossRef Medline](#)
- Spiers HJ, Maguire EA (2007) A navigational guidance system in the human brain. *Hippocampus* 17:618–626. [CrossRef Medline](#)
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520–522. [CrossRef Medline](#)
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15:3215–3230. [CrossRef Medline](#)
- Trevarthen CB (1968) Two mechanisms of vision in primates. *Psychol Forsch* 31:299–337. [CrossRef Medline](#)
- Troiani V, Stigliani A, Smith ME, Epstein RA (2014) Multiple object properties drive scene-selective regions. *Cereb Cortex* 24:883–897. [CrossRef Medline](#)
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Ingle D, Goodale M A, Mansfield R, eds), pp 549–587. Cambridge, MA: MIT.
- Vann SD, Aggleton JP, Maguire EA (2009) What does the retrosplenial cortex do? *Nat Rev Neurosci* 10:792–802. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Walther DB, Shen D (2014) Nonaccidental properties underlie human categorization of complex natural scenes. *Psychol Sci* 25:851–860. [CrossRef Medline](#)
- Wolbers T, Büchel C (2005) Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J Neurosci* 25:3333–3340. [CrossRef Medline](#)