# Distinct representations of spatial and categorical relationships across human scene-selective cortex

**Andrew S. Persichetti[a,b] and Daniel D. Dilks[b,1]**

[a]Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, MD 20892; and [b]Department of Psychology, Emory University, Atlanta, GA 30322

Edited by Nancy Kanwisher, Massachusetts Institute of Technology, Cambridge, MA, and approved September 5, 2019 (received for review February 22, 2019)

**We represent the locations of places (e.g., the coffee shop on 10th Street vs. the coffee shop on Peachtree Street) so that we can use them as landmarks to orient ourselves while navigating large-scale environments. While several neuroimaging studies have argued that the parahippocampal place area (PPA) represents such navigationally relevant information, evidence from other studies suggests otherwise, leaving this issue unresolved. Here we hypothesize that the PPA is, in fact, not well suited to recognize specific landmarks in the environment (e.g., the coffee shop on 10th Street), but rather is involved in recognizing the general category membership of places (e.g., a coffee shop, regardless of its location). Using fMRI multivoxel pattern analysis, we directly test this hypothesis. If the PPA represents landmark information, then it must be able to discriminate between 2 places of the same category, but in different locations. Instead, if the PPA represents general category information (as hypothesized here), then it will not represent the location of a particular place, but only the category of the place. As predicted, we found that the PPA represents 2 buildings from the same category, but in different locations, as more similar than 2 buildings from different categories, but in the same location. In contrast, another scene-selective region of cortex, the retrosplenial complex (RSC), showed the exact opposite pattern of results. Such a double dissociation suggests distinct neural systems involved in categorizing and navigating our environment, including the PPA and RSC, respectively.**

retrosplenial complex (RSC) | parrahippocampal place area (PPA) | navigation | categorization | scene processing

Recognizing places (or "scenes") in the world as landmarks and orienting ourselves to them is crucial for our ability to navigate large-scale environments. Given the ecological importance of landmark recognition, it is perhaps not surprising then that we may have dedicated cortex for processing landmark information (1–3). Indeed, several studies have argued that a scene-selective region, the parahippocampal place area (PPA) (4), is implicated in such landmark recognition (2). However, the role of the PPA in landmark recognition, and navigation more generally, is controversial for 4 reasons. First, many of the papers that claim the PPA is involved in landmark recognition are methodologically flawed by not independently defining a region of interest in an individual subject analysis, not correcting for multiple comparisons in a group analysis, or both (5–8). Such flaws in methodology lead to major issues, such as double dipping and inflated false positives, thus rendering the resultant findings difficult to interpret (9–11). Second, several studies have found that the PPA is insensitive to many kinds of navigationally relevant information, including sense (i.e., left/right), egocentric distance (i.e., near/far), first-person perspective motion, and the number and location of objects in a local space (refs. 12–15, but see refs. 16 and 17), making it ill-suited to support navigational processes. Third, still another study has recently found that the PPA is involved in scene categorization (e.g., recognizing a place as a kitchen or a bedroom), not visually guided navigation (i.e., our ability to navigate the immediately visible environment) (18), directly revealing the PPA's lack of involvement in at least 1 kind of navigation. Fourth, several other studies have shown that another scene-selective region, the retrosplenial complex (RSC), represents the spatial locations and facing directions of landmarks, situates them in relation to other places in the environment, and is involved in orienting oneself to landmarks while navigating (19–21), making the RSC, not the PPA, a better candidate cortical region involved in processing landmark information (22). For these reasons, it seems unlikely that the PPA is involved in landmark-based navigation. Therefore, we hypothesize that the PPA is involved in scene categorization, and not involved in any kind of navigation (i.e., visually guided or landmark based).

Here, using fMRI multivoxel pattern analysis, we directly test the hypothesis that the PPA is not involved in landmark-based navigation, but rather is involved in recognizing the category membership of scenes. Specifically, we scanned participants after they learned the layout of a virtual town that consisted of a park square surrounded by 8 buildings (Fig. 1*A*). There were 2 buildings on each corner of the town. Each building belonged to a particular category: 2 coffee shops, 2 hardware stores, 2 gyms, and 2 dentist offices. Importantly, the locations and facing directions of any 2 buildings belonging to the same category were dissociable from the category information (e.g., 1 gym was in the northeast corner of the town, while the other was in the southwest corner; Fig. 1*A*). If the PPA represents landmark information, then it must be able to discriminate between 2 places of the same category, and represent them as particular places in distinct locations and facing directions. In contrast, if the PPA represents general category information, then it will not represent the location or facing direction of a particular place, but only the category membership of the place, as predicted here.

## Significance

**Humans are capable of many sophisticated behaviors. One such behavior is using landmarks to navigate from one place to another, distant place. This type of navigation, known as landmark-based navigation, requires the navigator to extract spatial information from places in the environment. Another sophisticated behavior is recognizing the type of place one is in and acting appropriately in that environment (e.g., recognizing that you are in a kitchen, and thus making a cup of coffee). This type of behavior does not rely on spatial information so much as it relies on extracting cues related to the category membership of different places. Here we show that these complex interactions with our environment, navigating between places and recognizing them, are neurally dissociable.**

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

The authors declare no competing interest.

This article is a PNAS Direct Submission.



[1]To whom correspondence may be addressed. Email: dilks@emory.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1903057116/-/DCSupplemental.

First published September 30, 2019.

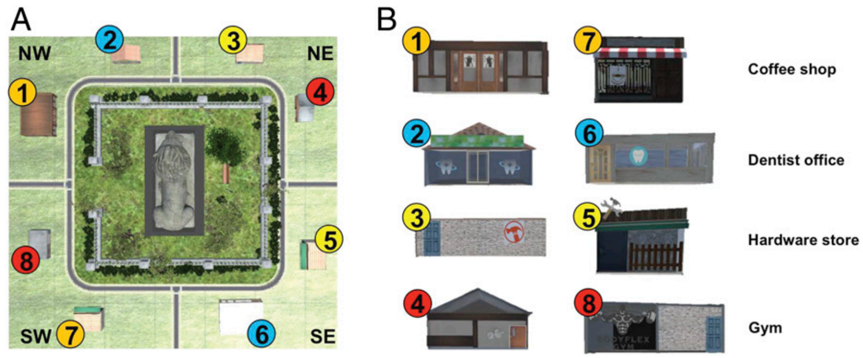


**Fig. 1.** (*A*) A bird's eye view of the virtual town. (Note, participants never saw this map-like representation of the town.) The overlaid numbers correspond to the buildings' positions in space (e.g., buildings 1 and 2 are located in the northwest corner of the town, while buildings 2 and 3 both face south). The colors of the numbered circles correspond to the category membership of the buildings, such that shared colors denote shared category membership. Each building could be paired by location, facing direction, or category. (*B*) Frontal views of the buildings along with their corresponding numbers from Fig. 1*A*. Note that none of the building pairings shared visual features, such as shape, size, or texture.



We contrast the patterns of responses within the PPA with the patterns of responses in the RSC, a region established to be involved in landmark-based navigation, and predict that the RSC will show the opposite pattern of responses from the PPA.

## Results

As predicted, the PPA contained information about the category of buildings, not their locations or facing directions. Specifically, a 3-level (Category, Location, Facing Direction) repeated-measures ANOVA revealed a significant main effect of condition [$F(2,26) = 3.73$; $P < 0.05$; $\eta_P^2 = 0.22$; Fig. 2*C*], with the Category condition significantly greater than both the Location and Facing Direction conditions (main effect contrasts, both *P*'s < 0.05, d's > 0.88), and no difference between Location and Facing Direction (main effect contrast, $P = 0.79$; $d = 0.10$; $BF_{10} < 1$). In contrast, the RSC contained information about the locations and facing directions of the buildings, but not the category. Specifically, a 3-level (Category, Location, Facing Direction) repeated-measures ANOVA revealed a significant main effect of condition [$F(2,26) = 12.08$; $P < 10^{-3}$; $\eta_P^2 = 0.48$; Fig. 2*C*], with both the Location and Facing Direction conditions significantly greater than the Category condition (main effect contrasts, both *P* values < 0.01, *d*s > 1.45), and no difference between Location and Facing Direction (main effect contrast, $P = 0.07$; $d = 0.53$; $BF_{10} < 1$). Next, since these results suggest a functional dissociation between the PPA and RSC, we directly tested this possibility by conducting a 2 (region of interest [ROI]: PPA, RSC) × 3 (Condition: Category, Location, Facing Direction) repeated-measures ANOVA. We found a significant ROI × Condition interaction [$F(2,26) = 15.47$; $P < 10^{-4}$; $\eta_P^2 = 0.54$], with the PPA containing information about the category membership of buildings, but not their locations or facing directions, and the RSC showing the opposite pattern of response. This result reveals a double dissociation between the responses in the PPA and RSC, and thus suggests distinct neural systems involved in categorizing scenes (including PPA) and representing navigationally relevant information in the broader environment (including RSC).

In addition to the PPA and RSC, we also asked whether a third scene-selective region of cortex, the occipital place area (OPA) (23), contains information about the categories, locations, or facing directions of buildings. We did not expect OPA to contain information related to scene categorization or landmark-based navigation, since it is thought to instead support visually guided navigation through the local environment, which is a kind of navigation not tested here (12–15, 24, 25). As predicted, the OPA does not contain information about the categories, locations, or facing directions of the buildings. Specifically, a 3-level (Category, Location, Facing Direction) repeated-measures ANOVA did not reveal a significant main effect of condition in the OPA [$F(2,26) = 0.19$; $P = 0.83$; $\eta_P^2 = 0.01$; Fig. 2*C*]. Finally, we conducted a 3 (ROI: OPA, PPA, RSC) × 3 (Condition: Category, Location, Facing Direction) repeated-measures ANOVA to test for the complete functional dissociation among all 3 of the scene-selective regions. We found a significant ROI × Condition interaction [$F(4,52) = 4.80$; $P < 0.01$; $\eta_P^2 = 0.27$], with PPA containing information about the category membership of buildings, but not their locations or facing directions; RSC showing the opposite pattern of response; and the OPA absent any of the information tested. This result reveals that the OPA does not contain information relevant for scene categorization or landmark-based navigation.

But might it be the case that the representation of scene category information in the PPA is a result of the presence of common objects on the facades of buildings from the same category? While we do not think so because buildings from the same category do not always share common objects (e.g., 1 coffee shop features a cup, while the other features a coffee pot), and when the objects are of the same type, they are not visually the same (e.g., the dumbbells common across the 2 gyms are visually quite different), we nonetheless directly tested this possibility by analyzing the pattern similarity in 2 regions of object-selective cortex: the lateral occipital sulcus and posterior fusiform sulcus. If the building categories are defined by the presence of objects, then these object-selective regions should represent category information in a similar way to the PPA. However, a 3-level (Category, Location, Facing Direction) repeated-measures ANOVA did not reveal a significant main effect of condition in either ROI (both *P* values > 0.25, $\eta_P^2$s < 0.09; *SI Appendix*, Fig. S1), thus confirming that the representation of scene category information in the PPA is not simply a result of the presence of common objects on the facades of buildings from the same category.

**Visualization of the Representational Spaces in the PPA and RSC.** In addition to the analysis presented here, in which we found evidence that the PPA represents the category membership of the buildings in a virtual town, but not their locations or facing directions, while the RSC showed the opposite pattern of results, we asked whether the representations of the buildings resembled the actual map of the town in each region of cortex. First, we applied multidimensional scaling to the dissimilarity matrix from each region of cortex to derive 2-dimensional layouts of the buildings from the virtual town (Fig. 3). We then correlated the dissimilarity matrix from each ROI with the Euclidean distance matrices derived from the actual layout of the buildings in the town, and the layout of the buildings if they were clustered by

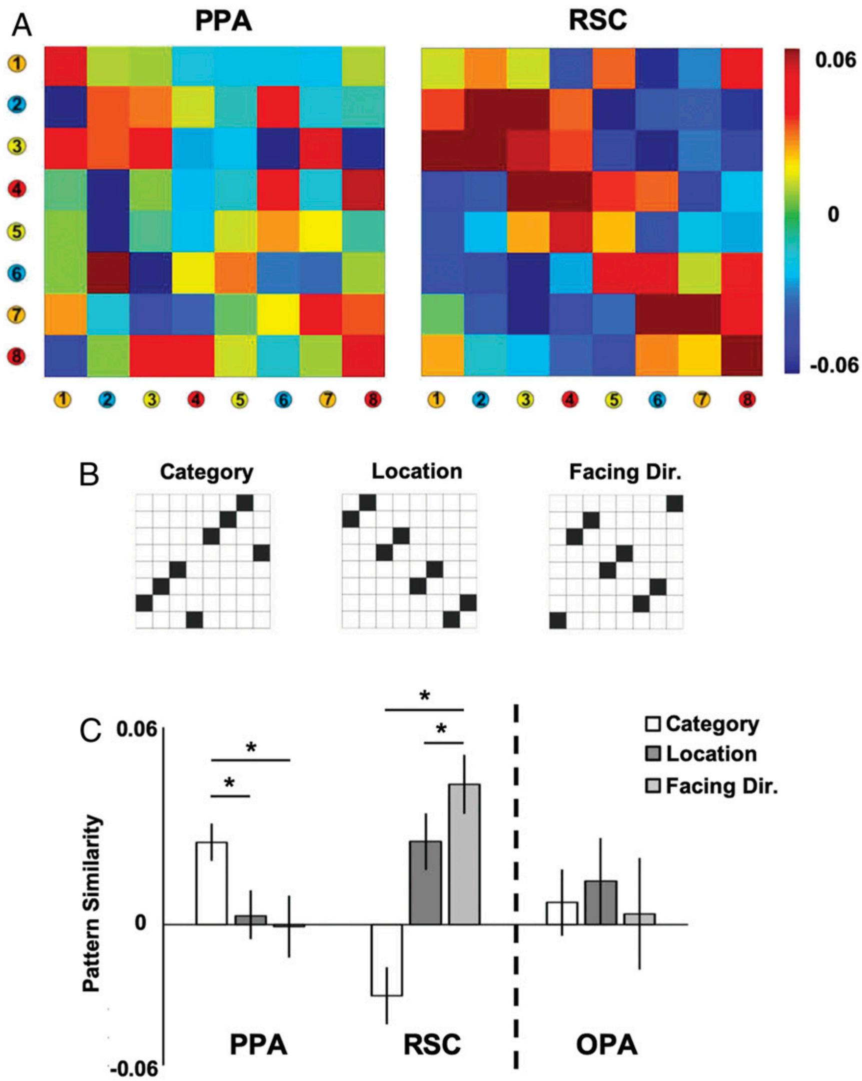


**Fig. 2.** (*A*) The 8 × 8 z-transformed correlation matrices for PPA and RSC, averaged across participants. (*B*) The idealized matrices for each condition of interest (i.e., Category, Location, and Facing Direction). (*C*) The beta weights obtained from each regressor, averaged across participants (referred to here as Pattern Similarity). Asterisks indicate differences between conditions. Error bars are ± 1 SEM.

category, respectively, and used permutation testing to quantify the similarity between the ROI matrices and the Euclidean distance matrices (see *Methods* for more details). If the PPA represents the category membership of the buildings, but not information about their locations or facing directions, then the layout derived from the neural pattern of responses in the PPA should not recreate the map of the town, but instead should cluster the buildings based on their category membership. In contrast, if the RSC represents the locations and facing directions of the buildings, then the layout derived from the neural pattern of responses in the RSC should resemble the map of the town. As predicted, we found that the building layout derived from the PPA grouped the buildings by category membership ($r^2 = 0.39$; $P < 0.05$), and not by the actual map of the town ($r^2 = 0.08$; $P = 0.33$). In contrast, the building layout from the RSC was similar to the map of the virtual town ($r^2 = 0.68$; $P < 10^{-3}$), and did not group the buildings based on their category membership ($r^2 = -0.03$; $P = 0.53$; Fig. 3). Finally, we found that the layout of the distances between the buildings in the PPA was significantly different from the distances between buildings in the RSC ($z = 4.22$; $P < 10^{-4}$). The results from this analysis demonstrate that the information represented in the PPA can be used to distinguish between the kinds of places in an environment, while the information represented in the RSC is adequate for constructing a navigationally useful map of the broader environment.

## Discussion

We had participants learn the locations, facing directions, and categories of different types of buildings in an unfamiliar virtual town, and then used multivoxel pattern analysis to ask which scene-selective regions of cortex contain navigationally relevant information or category information. We found that the PPA does not contain navigationally relevant information about the locations or facing directions of the buildings, but instead contains information about their category membership. In contrast, another scene-selective region, the RSC, contains information about the locations and facing directions of the buildings, but not their category membership. This double dissociation suggests that cortical scene processing can be divided into at least 2 systems: 1 involved in scene categorization, which includes the PPA, and 1 for landmark-based navigation, including the RSC.

Our finding that the PPA contains information about the categories of buildings, but not information necessary for landmark-based navigation, is consistent with 2 groups of prior studies. First,

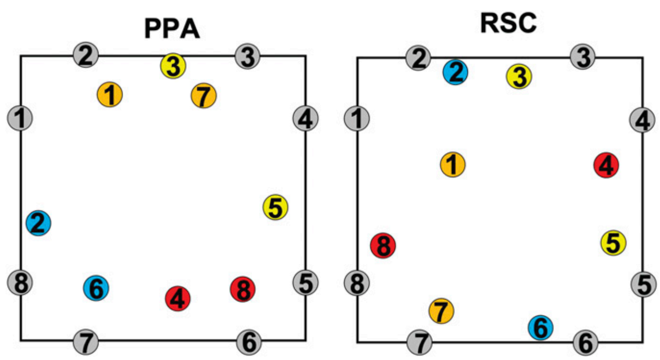


**Fig. 3.** The results of the multidimensional scaling analysis in each ROI. The gray circles are in the positions of the buildings in the virtual town, while the positions of the colored circles are the result of the multidimensional scaling analysis. In the PPA, the colored circles do not correspond to the gray circles, but rather are grouped more by category, while the colored circles are close to the corresponding gray circles in the RSC map.

our findings are consistent with reports that the PPA does not process information necessary for yet another kind of navigation (i.e., visually guided navigation), such as sense, egocentric distance, first-person perspective motion, and locations of obstacles and boundaries (refs. 12–15, but see ref. 16 and 17), but instead processes information that could be used for recognizing and categorizing scenes (e.g., the relative lengths and angles that make up the large extended surfaces in scenes) (26). Second, our findings are also consistent with other studies in which the PPA was found to be directly involved in scene categorization (e.g., recognizing a place as a forest, not a beach) (18, 27, 28), but not visually guided navigation (18). For example, Persichetti and Dilks (18) found that the PPA was active when participants were asked to categorize rooms as bedrooms, kitchens, or living rooms, but not when asked to imagine navigating through those same rooms. The results from Persichetti and Dilks (18), taken together with the results from the current study, suggest that the PPA is not involved in visually guided navigation through the immediately visible environment, nor in landmark-based navigation through the broader environment, but is instead involved in recognizing the categories of places (e.g., a city, coffee shop, or bedroom).

In contrast, our finding that PPA is involved in scene categorization is seemingly inconsistent with several other studies. First, 2 fMRI studies found the PPA is not involved in scene categorization (29, 30). In these studies, however, the authors used distinctions such as natural and manmade, or ice caves, which might not be the proper level of representation of category information in the PPA. Thus, the question is not whether PPA is involved in scene categorization but, instead, what is the level of category information represented in the PPA. Second, our hypothesis that the PPA is involved in scene categorization, but not navigation, may seem inconsistent with neuropsychological studies that report lesions to the parahippocampal cortex lead to impaired navigational abilities (31, 32). However, there are no reports of a lesion restricted to the PPA leading to such impairments: all of the reports are from patients with lesions of the parahippocampal cortex and/or retrosplenial cortex and/or hippocampus. Third, one might argue that our hypothesis that the scene processing system comprises distinct cognitive systems is at odds with reports of connectivity among the PPA, RSC, OPA, and hippocampus (33–35). However, our findings demonstrate that the scene categorization system and navigation systems are functionally dissociable, but that does not mean that they cannot and do not interact. More work is needed to explore the functional relevance of the connections between these brain regions.

Furthermore, the current findings are inconsistent with a group of studies (5–8) claiming to show that the PPA responds more strongly to objects that were previously encountered at intersections (and thus navigationally relevant) than to equivalent objects that were previously encountered in positions that were not navigationally relevant. While these studies employed clever experimental designs, they were also methodologically flawed by not independently defining a region of interest in an individual-subject analysis, not correcting for multiple comparisons in a group analysis, or both, thus leaving open the question of whether (or not) the PPA represents objects positioned at navigationally relevant decision points as landmarks. That said, a similar study, which was not methodologically flawed, found that the PPA responded more strongly to navigationally relevant buildings than to buildings that were not navigationally relevant (36). While this study both independently defined the PPA in an individual-subject analysis and corrected for multiple comparisons in a group analysis, the inferences drawn from their results are nonetheless still limited for 2 reasons. First, the authors reported that in addition to the PPA, several other regions that are often reported as responding to spatial information (e.g., the RSC, frontal eye fields, and superior parietal lobule) were also potentially sensitive to the navigational relevance of the buildings along the route. Therefore, it could be the case that the PPA and these other regions responded more to navigationally relevant buildings simply because there is more visuospatial information connected to buildings placed at intersections than those placed where the route goes only 1 way. Second, it is possible that the greater response to navigationally relevant buildings in these regions is due to participants' paying more attention to buildings that are located in places where they are actively making a decision versus when they are simply passively "walking" by them. Thus, the interpretation that the PPA is involved in landmark recognition because it responds to navigationally relevant buildings is tenuous, and it seems to be born out of the assumption that if a region is scene-selective, then it must be involved in navigation, which is unfounded.

Our finding that the RSC contains information about the locations and facing directions of the buildings is consistent with several prior studies showing that the RSC is involved in navigation through the broader environment (1, 19, 20, 37–39). In these studies, the RSC was found to support several functions, from recognizing landmarks and situating them in relation to other places in the environment to orienting oneself to landmarks while navigating. In the current study, the RSC contains information about the locations and facing directions of buildings, both of which are represented in an allocentric frame of reference. Prior studies have also shown that the RSC represents scenes in an egocentric frame of reference. For example, the RSC is sensitive to sense and egocentric distance information in scenes (12, 15). Taken together, these results are consistent with the hypothesis that the RSC integrates egocentric and allocentric spatial information while learning the layout of a new environment, in the service of constructing a cognitive map of the environment and navigating large-scale environments (19, 20, 40).

Moreover, our findings that the RSC contains information about locations and facing directions of buildings in a virtual environment is a replication and extension of findings reported by Marchette and colleagues (41). They reported that the RSC represents locations and facing directions relative to artifacts placed in corners of virtual museums. In their study, the layout of artifacts within a museum was the same as the layout of the buildings in our virtual town (i.e., 8 total artifacts placed on either side of all 4 corners of the museum). Thus, we replicate their findings in a broader environment (i.e., a virtual town) than the environments used in their experiment (i.e., the inside of museums).

Finally, our finding that the OPA does not contain information about the locations, facing directions, or categories of buildings is consistent with the evidence that it is involved in visually guided navigation (12–15, 18, 24, 25), but not scene categorization (18). Furthermore, visually guided navigation does not require representations of the locations of buildings relative to one another, or their facing directions. Thus, the type of navigation supported by OPA is dissociable from the type of navigation supported by the RSC, which suggests that there are at least 2 dissociable navigation systems: 1 system responsible for orienting to the broader environment, including the RSC, and another for visually guided navigation through the immediate environment, including the OPA.

In conclusion, we found a double dissociation between the multivoxel patterns of response in the PPA and RSC, with the PPA containing information about the categories of buildings, but not their locations or facing directions, while the RSC showed the opposite pattern of responses. Our results suggest that the PPA does not contain the spatial information that is crucial for recognizing landmarks in the environment and orienting to them, but instead contains information necessary for categorizing places. Furthermore, our results suggest that cortical scene processing can be divided into at least 2 systems, with 1 involved in scene categorization, including the PPA, and the other for landmark-based navigation, including the RSC.

## Methods

**Participants.** Fourteen healthy adults (aged 20 to 38 y; 8 women) were recruited from the Emory University community. All participants gave informed consent and had normal or corrected-to-normal vision.

**Design.** We used a ROI approach, in which we localized scene-selective cortical regions (Localizer runs) and then used an independent set of data (Experimental runs) to investigate the multivoxel pattern in each ROI to each building from our virtual town. Specifically, we correlated the patterns of responses to buildings from the same location of town (e.g., 2 buildings in the Northwest corner of the town), facing direction (e.g., 2 buildings facing South), and category (e.g., 2 coffee shops) in each ROI.

For the Localizer runs, participants saw pictures of faces, objects, scenes, and scrambled objects in a blocked design, as previously described (26). Each run was 336 s long and consisted of 4 blocks per stimulus category. The order of the stimulus category blocks in each run was pseudorandomized across runs. Each block contained 20 pictures from the same category, with each picture presented for 300 ms, followed by a 500-ms interstimulus interval, for a total of 16-s blocks. We also included five 16-s fixation blocks: 1 at the beginning; 3 in the middle, interleaved between each set of stimulus category blocks; and 1 at the end of each run. Participants performed a 1-back task, responding with a button press every time the same picture was presented twice in a row.

For the Experimental runs, participants completed 8 runs, each with an average of 83 experimental trials and 18 trials in which the screen was a blank neutral gray screen. Each run ended with 8 s of a white central fixation cross on a blank gray screen. Each run was 412 s long. The trial sequence was generated using a de Bruijn sequence that was designed to order the trials such that every stimulus preceded every other stimulus an equal number of times (42, 43). On each experimental trial, an image of a building subtending $4° \times 6°$ of visual angle appeared on the top half of a neutral gray screen below the words "Imagine facing:" printed in white. After 1 s, another building appeared below the first building, underneath the words "LEFT or RIGHT?" Both buildings remained on the screen for 2 s, followed by 1 s of a blank screen, for a total of 4 s per trial. On each trial, participants performed a judgment of relative direction task, in which they imagined themselves facing the top building, and then judged whether the bottom building would be to the left or right side of their body, given their current facing position. Two buildings that were directly across from one another in the town were never paired during the experiment, since there would not be a correct "left" or "right" answer on such trials. The judgment of relative direction task required participants to mentally reorient themselves to a particular building in each trial (41).

We used the Unity software (https://unity.com, Unity Technologies) to create a virtual town that consisted of a town square surrounded on all sides by city streets. There was 1 building on each outer edge of all 4 corners of the streets, for a total of 8 buildings in the town (Fig. 1*A*). Each building was of a particular type: 2 coffee shops, 2 hardware stores, 2 gyms, and 2 dentist offices. Crucially, any 2 buildings that shared the same category, location, or facing direction did not also share similar visual features, such as shape, size, or texture (Fig. 1*B*). Before entering the MRI scanner, each participant explored the virtual town from a first-person perspective (*SI Appendix*, Fig. S2*A*), learning the types of buildings and their locations. First, participants freely navigated around the town by using the arrow keys on the computer keyboard for 5 min. Next, they were dropped at a random point in the town and were directed to navigate to a particular building (e.g., the northeast gym). This was repeated 3 times for each building and took each participant roughly the same amount of time to complete. Finally, each participant was given a posttest in which they saw a frontal view of each building individually during each trial and was asked to identify either the category of the building or its location (*SI Appendix*, Fig. S2*B*). Each run of the posttest included 4 blocks: 2 blocks asking about category and 2 asking about location; the "category" and "location" blocks alternated, and all buildings were shown once during each block. All participants completed 2 runs of this posttest, and all were 100% accurate on the second posttest.

**fMRI Scanning.** Scanning was completed on a 3T Siemens Trio scanner at the Facility for Education and Research in Neuroscience (FERN) at Emory University (Atlanta, GA). Functional images were acquired using a 32-channel head matrix coil and a gradient echo single-shot echo planar imaging sequence. Twenty-eight image slices were acquired for both the Localizer and Experimental runs (repetition time = 2 s). These slices were oriented approximately between perpendicular and parallel to the calcarine sulcus, covering the occipital and temporal lobes and the lower portion of the parietal lobe (as in refs. 18 and 26). For all scans: echo time was 30 ms, and voxel size was 1.5 mm × 1.5 mm × 2.5 mm, with a 0.25-mm interslice gap. For each participant, whole-brain, high-resolution T1 weighted anatomical images were also acquired for anatomical localization.

**Data Analysis.** fMRI data analysis was conducted using the FSL software (44) and custom MATLAB code. Before statistical analysis, data from both the Localizer and Experimental runs were skull-stripped (45), motion corrected using FSL's MCFLIRT tool (46), and registered to a separate functional volume with the same slice prescription. Data from the Localizer runs were spatially smoothed with a 6-mm kernel, whereas the Experimental runs were not spatially smoothed. All data were fit with a general linear model (GLM) that was high-pass filtered to remove low temporal frequencies. For the Localizer runs, the GLM contained a covariate for each stimulus type (i.e., faces, objects, scenes, and scrambled objects). For the Experimental runs, the GLM contained a covariate for each of the 8 buildings, and another for the trials in which there was a blank screen. These covariates were convolved with a double-gamma function to approximate the hemodynamic response function, and the temporal derivative of the response function was included to correct for slice timing and timing delays in the hemodynamic response function. The GLM for both the Localizer and the Experimental runs also included nuisance covariates for times that were corrupted by large movements (i.e., motion spikes), identified using FSL's motion outlier tool. Finally, beta weights associated with each covariate were extracted for further analyses.

After preprocessing, we functionally defined scene-selective ROIs bilaterally in each participant, using data from the Localizer runs. To ensure that each ROI in each hemisphere consisted of an equal number of voxels in all participants, we extracted the top 100 scene-selective voxels (defined as voxels that responded more to scenes than objects in each participant) from scene-selective parcels for PPA, RSC, and OPA that were defined using a Group-Constrained Subject Specific Method in a separate group of 42 adults (47). We then split the Experimental data into even and odd runs and subtracted the grand mean (i.e., the mean response to all tasks) from each voxel in each ROI in each participant. Next, we created an $8 \times 8$ matrix of all pairwise correlations between the pattern of activation of voxels for each building across the even and odd runs, and transformed the correlations into Fischer z-scores (Fig. 2*A*). For each ROI, we then ran a linear regression model that contained the correlation matrix as the dependent variable and a separate idealized matrix for each condition of interest (i.e., Category, Location, and Facing Direction) as regressors (Fig. 2*B*). Finally, repeated-measures ANOVAs were performed on the neural responses for each ROI. A 2 (ROI: PPA, RSC) × 2 (hemisphere: Left, Right) × 3 (condition: Category, Location, Facing Direction, Category) repeated-measures ANOVA did not reveal a significant ROI × Hemisphere × Task interaction [$F(2,26) = 0.20$; $P = 0.80$; $\eta_P^2 = 0.02$], nor did we find a significant interaction between Hemisphere and Condition or ROI and Hemisphere (both $P$ values $> 0.77$). Thus, data from each hemisphere were collapsed. We also ran these same analyses on $8 \times 8$ correlation matrices obtained by splitting the Experimental runs into all 35 possible spilt-half combinations of the data and then averaged across them. The results of this analysis were similar to the results we found when we split the Experimental data into even and odd runs (*SI Appendix*, Fig. S3). In addition, a Bayes Factor (BF) was calculated for each pairwise comparison between the 3 conditions of interest in each ROI when no significant difference was reported to confirm that these results indeed favored the null hypothesis (48).

We further probed the representations of the buildings in the virtual town in each ROI by attempting to reconstruct the layout of the town by applying multidimensional scaling to the correlation matrix from each ROI. Specifically, in each ROI, we first averaged the $8 \times 8$ correlation matrices across participants, converted the average matrix into a dissimilarity matrix by normalizing the values in the matrix between 0 and 1, and then subtracting each value from 1, and finally applied multidimensional scaling on the dissimilarity matrix. We also calculated the Euclidean distances between the buildings based on their actual locations in the town, and based on a layout in which the buildings were grouped by category. We then correlated the dissimilarity matrix from each ROI with the Euclidean distances between both the locations of the buildings in the town, and the buildings grouped by category. To calculate the significance of the correlations, we randomly shuffled the cells in each ROI dissimilarity matrix, correlated the random matrix to each building layout matrix 25,000 times to obtain a null distribution of correlations, and compared the correlation between the unshuffled dissimilarity matrix and each building layout matrix to this null distribution of correlations. Finally, we directly compared the PPA to the RSC by first calculating a difference score in each ROI (i.e., the correlation between the dissimilarity matrix and the location-based matrix minus the correlation between the dissimilarity matrix and the category-based matrix), and then directly compared the difference scores across ROIs.

**Data Availability.** The datasets generated during the current study are available at https://osf.io/hcbdt/.

**ACKNOWLEDGMENTS.** We thank the Facility for Education and Research in Neuroscience (FERN) Imaging Center in the Department of Psychology, Emory University, Atlanta, Georgia. We also thank Frederik Kamps for insightful conversation and Samuel K. Weiller for technical support. This work was supported by grants from the National Eye Institute (R01 EY29724 to D.D.D. and T32 EY007092 to A.S.P.) and a National Science Foundation Graduate Research Fellowship (DGE-1444932 to A.S.P.).

1. S. D. Auger, S. L. Mullally, E. A. Maguire, Retrosplenial cortex codes for permanent landmarks. *PLoS One* **7**, e43620 (2012).
2. R. A. Epstein, "Neural systems for visual scene recognition" in *Scene Vision: Making Sense of What We See*, M. Bar, K. Kveraga, Eds. (MIT Press, Cambridge, MA, 2014) pp. 105–134.
3. R. A. Epstein, E. Z. Patai, J. B. Julian, H. J. Spiers, The cognitive map in humans: Spatial navigation and beyond. *Nat. Neurosci.* **20**, 1504–1513 (2017).
4. R. Epstein, N. Kanwisher, A cortical representation of the local visual environment. *Nature* **392**, 598–601 (1998).
5. G. Janzen, C. Jansen, M. van Turennout, Memory consolidation of landmarks in good navigators. *Hippocampus* **18**, 40–47 (2008).
6. G. Janzen, M. van Turennout, Selective neural representation of objects relevant for navigation. *Nat. Neurosci.* **7**, 673–677 (2004).
7. G. Janzen, C. G. Weststeijn, Neural representation of object location and route direction: An event-related fMRI study. *Brain Res.* **1165**, 116–125 (2007).
8. J. Wegman, G. Janzen, Neural encoding of objects relevant for navigation and resting state correlations with navigational ability. *J. Cogn. Neurosci.* **23**, 3841–3854 (2011).
9. C. Bennett, M. Miller, G. Wolford, Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. *Neuroimage* **47**, S125 (2009).
10. N. Kriegeskorte, W. K. Simmons, P. S. Bellgowan, C. I. Baker, Circular analysis in systems neuroscience: The dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
11. E. Vul, C. Harris, P. Winkielman, H. Pashler, Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect. Psychol. Sci.* **4**, 274–290 (2009).
12. D. D. Dilks, J. B. Julian, J. Kubilius, E. S. Spelke, N. Kanwisher, Mirror-image sensitivity and invariance in object and scene processing pathways. *J. Neurosci.* **31**, 11305–11312 (2011).
13. F. S. Kamps, J. B. Julian, J. Kubilius, N. Kanwisher, D. D. Dilks, The occipital place area represents the local elements of scenes. *Neuroimage* **132**, 417–424 (2016).
14. F. S. Kamps, V. Lall, D. D. Dilks, The occipital place area represents first-person perspective motion information through scenes. *Cortex* **83**, 17–26 (2016).
15. A. S. Persichetti, D. D. Dilks, Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex* **77**, 155–163 (2016).
16. R. Epstein, K. S. Graham, P. E. Downing, Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* **37**, 865–876 (2003).
17. S. Park, M. M. Chun, Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage* **47**, 1747–1756 (2009).
18. A. S. Persichetti, D. D. Dilks, Dissociable neural systems for recognizing places and navigating through them. *J. Neurosci.* **38**, 10295–10304 (2018).
19. T. Wolbers, C. Büchel, Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *J. Neurosci.* **25**, 3333–3340 (2005).
20. K. R. Sherrill *et al*., Hippocampus and retrosplenial cortex combine path integration signals for successful navigation. *J. Neurosci.* **33**, 19304–19313 (2013).
21. L. K. Vass, R. A. Epstein, Abstract representations of location and facing direction in the human brain. *J. Neurosci.* **33**, 6133–6142 (2013).
22. A. S. Persichetti, D. D. Dilks, The parahippocampal place area is involved in scene categorization, not landmark recognition. *J. Vis.* **18**, 1239 (2018).
23. D. D. Dilks, J. B. Julian, A. M. Paunov, N. Kanwisher, The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* **33**, 1331–6a (2013).
24. M. F. Bonner, R. A. Epstein, Coding of navigational affordances in the human visual system. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4793–4798 (2017).
25. J. B. Julian, J. Ryan, R. H. Hamilton, R. A. Epstein, The occipital place area is causally involved in representing environmental boundaries during navigation. *Curr. Biol.* **26**, 1104–1109 (2016).
26. M. R. Dillon, A. S. Persichetti, E. S. Spelke, D. D. Dilks, Places in the brain: Bridging layout and object geometry in scene-selective cortex. *Cereb. Cortex* **28**, 2365–2374 (2018).
27. D. B. Walther, E. Caddigan, L. Fei-Fei, D. M. Beck, Natural scene categories revealed in distributed patterns of activity in the human brain. *J. Neurosci.* **29**, 10573–10581 (2009).
28. D. B. Walther, B. Chai, E. Caddigan, D. M. Beck, L. Fei-Fei, Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9661–9666 (2011).
29. D. J. Kravitz, C. S. Peng, C. I. Baker, Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *J. Neurosci.* **31**, 7322–7333 (2011).
30. S. Park, T. F. Brady, M. R. Greene, A. Oliva, Disentangling scene content from spatial boundary: Complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *J. Neurosci.* **31**, 1333–1340 (2011).
31. G. K. Aguirre, M. D'Esposito, Topographical disorientation: A synthesis and taxonomy. *Brain* **122**, 1613–1628 (1999).
32. R. Epstein, E. A. Deyoe, D. Z. Press, A. C. Rosen, N. Kanwisher, Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cogn. Neuropsychol.* **18**, 481–508 (2001).
33. S. Nasr *et al*., Scene-selective cortical regions in human and nonhuman primates. *J. Neurosci.* **31**, 13771–13785 (2011).
34. S. Nasr, K. J. Devaney, R. B. H. Tootell, Spatial encoding and underlying circuitry in scene-selective cortex. *Neuroimage* **83**, 892–900 (2013).
35. C. Baldassano, D. M. Beck, L. Fei-Fei, Differential connectivity within the parahippocampal place area. *Neuroimage* **75**, 228–237 (2013).
36. V. R. Schinazi, R. A. Epstein, Neural correlates of real-world route learning. *Neuroimage* **53**, 725–735 (2010).
37. N. Burgess, E. A. Maguire, H. J. Spiers, J. O'Keefe, A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage* **14**, 439–453 (2001).
38. S. D. Vann, J. P. Aggleton, E. A. Maguire, What does the retrosplenial cortex do? *Nat. Rev. Neurosci.* **10**, 792–802 (2009).
39. S. D. Auger, P. Zeidman, E. A. Maguire, A central role for the retrosplenial cortex in de novo environmental learning. *eLife* **4**, e09031 (2015).
40. R. A. Epstein, Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn. Sci.* **12**, 388–396 (2008).
41. S. A. Marchette, L. K. Vass, J. Ryan, R. A. Epstein, Anchoring the neural compass: Coding of local spatial reference frames in human medial parietal lobe. *Nat. Neurosci.* **17**, 1598–1606 (2014).
42. G. K. Aguirre, Continuous carry-over designs for fMRI. *Neuroimage* **35**, 1480–1494 (2007).
43. G. K. Aguirre, M. G. Mattar, L. Magis-Weinberg, de Bruijn cycles for neural decoding. *Neuroimage* **56**, 1293–1300 (2011).
44. S. M. Smith *et al*., Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23** (suppl. 1), S208–S219 (2004).
45. S. M. Smith, Fast robust automated brain extraction. *Hum. Brain Mapp.* **17**, 143–155 (2002).
46. M. Jenkinson, P. Bannister, M. Brady, S. Smith, Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).
47. J. B. Julian, E. Fedorenko, J. Webster, N. Kanwisher, An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage* **60**, 2357–2364 (2012).
48. A. F. Jarosz, J. Wiley, What are the odds? A practical guide to computing and reporting Bayes factors. *J. Probl. Solving* **7**, 2–9 (2014).