RESEARCH ARTICLE | *Higher Neural Functions and Behavior*

Conjoint representation of texture ensemble and location in the parahippocampal place area

Jeongho Park[1] and Soojin Park[1,2]

[1]*Department of Cognitive Science, Johns Hopkins University, Baltimore, Maryland; and* [2]*Department of Psychology, Yonsei University, Seoul, South Korea*

Park J, Park S. Conjoint representation of texture ensemble and location in the parahippocampal place area. *J Neurophysiol* 117: 1595–1607, 2017. First published January 25, 2017; doi:10.1152/ jn.00338.2016.—Texture provides crucial information about the category or identity of a scene. Nonetheless, not much is known about how the texture information in a scene is represented in the brain. Previous studies have shown that the parahippocampal place area (PPA), a scene-selective part of visual cortex, responds to simple patches of texture ensemble. However, in natural scenes textures exist in spatial context within a scene. Here we tested two hypotheses that make different predictions on how textures within a scene context are represented in the PPA. The Texture-Only hypothesis suggests that the PPA represents texture ensemble (i.e., the kind of texture) as is, irrespective of its location in the scene. On the other hand, the Texture and Location hypothesis suggests that the PPA represents texture and its location within a scene (e.g., ceiling or wall) conjointly. We tested these two hypotheses across two experiments, using different but complementary methods. In *experiment 1*, by using multivoxel pattern analysis (MVPA) and representational similarity analysis, we found that the representational similarity of the PPA activation patterns was significantly explained by the Texture-Only hypothesis but not by the Texture and Location hypothesis. In *experiment 2*, using a repetition suppression paradigm, we found no repetition suppression for scenes that had the same texture ensemble but differed in location (supporting the Texture and Location hypothesis). On the basis of these results, we propose a framework that reconciles contrasting results from MVPA and repetition suppression and draw conclusions about how texture is represented in the PPA.

NEW & NOTEWORTHY This study investigates how the parahippocampal place area (PPA) represents texture information within a scene context. We claim that texture is represented in the PPA at multiple levels: the texture ensemble information at the across-voxel level and the conjoint information of texture and its location at the within-voxel level. The study proposes a working hypothesis that reconciles contrasting results from multivoxel pattern analysis and repetition suppression, suggesting that the methods are complementary to each other but not necessarily interchangeable.

IN A CONSTANTLY CHANGING WORLD, the ability to process visual scenes is important for recognition of specific locations and navigation. Previous literature reveals that people can recog-

nize scenes even with just a brief glance (Potter 1976; Schyns and Oliva 1994; Thorpe et al. 1996; VanRullen and Thorpe 2001). Such an effortless ability is supported by a network of cortical regions specialized in processing visual scenes. Among these cortical regions, the best-studied region is perhaps the parahippocampal place area (PPA; Epstein and Kanwisher 1998), which is located in the posterior part of the parahippocampal gyrus in the medial temporal lobe. The PPA responds more strongly to pictures of scene categories, houses, landmarks, and spatial structure (e.g., lines that define spatial layout of 3-dimensional scene) than to pictures of objects or faces (Aguirre et al. 1998; Epstein et al. 1999; Epstein and Kanwisher 1998; Epstein and Morgan 2012; Kravitz et al. 2011; Morgan et al. 2011; Park et al. 2011; Park and Chun 2009; Walther et al. 2009, 2011). Neurological evidence also supports the specialized role of the parahippocampal cortex (PHC) in scene recognition. Patients with lesions in the PHC show difficulty in recognizing pictures of familiar places or in wayfinding (Aguirre and D'Esposito 1999; Barrash et al. 2000; Takahashi and Kawamura 2002). Since its discovery, the PPA has been the main focus of studies investigating visual scene perception, recognition, and memory.

Although the PPA has been extensively studied as an area that primarily represents the geometric structure of scenes, recent studies suggest that the representational properties of the PPA are much richer. For example, Bar and Aminoff (2003) proposed that the PHC is activated by highly contextual objects (i.e., objects that are frequently associated with a certain context, such as road signs or a refrigerator) (Bar et al. 2008). Moreover, the posterior PHC, including the PPA, shows a preference for real-world-sized large objects (e.g., a couch or refrigerator) compared with small objects (e.g., a cup or ring) even when the retinal sizes of the objects are controlled (Konkle and Oliva 2012). The PPA even responds to small objects that have crucial value in navigation. In navigation in a virtual reality maze, the PPA shows higher activation to objects that are placed at navigational decision points (e.g., a corner where participants could turn either left or right) than to objects that are placed at navigationally irrelevant points (Janzen and Jansen 2010; Janzen and van Turennout 2004). In accordance with findings from Konkle and Oliva and Janzen's group, a study by Troiani et al. (2014) also shows that the PPA activity is closely related to the visual size and landmark suitability of objects. These findings together suggest that the representa-

tional properties of the PPA are complex and go beyond the spatial aspects of a scene.

In this report, we focus on one of the PPA's newly discovered representational properties: texture information. Cant and Goodale (2007, 2011) have shown that the PPA responds to patches of textures or to textures on object surfaces. For example, the PPA showed higher activation when participants paid attention to the surface of objects (e.g., texture or material properties) than when they paid attention to the shape of objects, whereas the lateral occipital area demonstrated the reverse pattern of activation. Further study using a repetition suppression method showed that the PPA is sensitive to change in the ensemble (statistical summary) of texture information. Repetition suppression is the reduction of a signal when a stimulus repeats twice. Importantly, the amount of repetition suppression allows testing of representational properties of a particular brain region; for example, it tests whether a particular brain region treats two stimuli that share a particular property similarly or differently. Cant and Xu (2012, 2015) found that the PPA showed a repetition suppression effect for images that were not identical but shared the same texture ensemble.

How is texture represented in the PPA? Cant and Xu (2012) suggest that the response patterns of the PPA to texture ensemble are distinct from the response patterns of the lateral occipital area, hinting at the PPA's unique involvement in processing textures. Interestingly, these studies used simple texture patches that did not have any spatial or navigational context associated with them. However, in natural scene perception, texture patches often appear at specific locations within a scene. The combination of texture and other attributes, such as the texture's location in a scene, helps to define the category or identity of a scene (Epstein and Julian 2013; Oliva and Torralba 2006). For example, a scene that has grass texture at the bottom of the image is more likely to be "a field," whereas a scene that has the same grass texture at the top of the image is more likely to be "a mountain." Similarly, indoor scenes with identical spatial layout (i.e., spatial layout refers to permanent geometric shape of a space that is not changed by viewpoints) may be recognized as two different places based on whether a certain texture is located on the wall or the floor. Thus it seems plausible that texture information in a scene is represented conjointly with spatial location information.

Indeed, recent evidence from macaque monkey functional MRI (fMRI) and single-cell recording (Kornblith et al. 2013) suggests that neurons in scene-specific regions of the monkey cortex are modulated not only by texture alone but also by the combination of texture and other factors, such as depth, viewpoint, and object identity. Kornblith and colleagues (2013) localized two anatomically discrete areas in the macaque monkey brain with fMRI and cell recording techniques: the lateral place patch, located in the occipitotemporal sulcus anterior to the V4 area, and the medial place patch in the medial parahippocampal gyrus. When monkeys viewed artificial room images, spatial features such as viewpoint or depth alone did not suffice to modulate neurons in the lateral place patch or the medial place patch. Instead, these neurons were most strongly modulated by texture alone, followed by texture with viewpoint or depth (Kornblith et al. 2013). These findings suggest that there are neurons in scene-selective cortex of monkeys that code texture information conjointly with other spatial attributes in a scene.

There were two aims of this study. First, we investigated how the PPA represents texture information within a scene. Two competing hypotheses about texture representation in the PPA were directly tested. The first hypothesis is that the PPA represents the texture ensemble (i.e., the kind of texture) as is, irrespective of the spatial locations in a scene (Texture-Only hypothesis). The second hypothesis is that the PPA represents the texture and its location information conjointly (Texture and Location hypothesis). To test these hypotheses, we generated synthetic room images that consisted of back wall, ceiling, floor, and left and right walls. Critically, we manipulated the location of specific textures within a scene by swapping the texture of the left and right walls with that of the ceiling and the floor. This manipulation results in different combinations of texture and location while maintaining the same overall texture ensemble of a scene. If the PPA represents texture ensemble independent of texture locations (Texture-Only hypothesis), then the PPA will not be sensitive to this manipulation; in other words, even if these two images look like different places, because the texture ensemble is the same PPA will not distinguish between the images. On the other hand, if the PPA represents each texture conjointly with its location (Texture and Location hypothesis), we would expect different representations for different combinations of texture and location.

Second, we aimed to shed light on an issue that has been a puzzle in the field: contradicting results from multivoxel pattern analysis (MVPA) and repetition suppression. An increasing number of studies have reported an apparent discrepancy between the results from these two methods (Drucker and Aguirre 2009; Epstein and Morgan 2012; Ward et al. 2013). Although a rough idea that these methods capture different aspects of neural representation was suggested, here we attempt to provide a more specific and mechanistic explanation for such difference. To investigate this question, we conducted two experiments using MVPA and the repetition suppression paradigm. In *experiment 1*, we used representational similarity analysis to ask whether multivoxel patterns (MVPs) in the PPA align with the Texture-Only or Texture and Location hypothesis. In *experiment 2*, we used the repetition suppression paradigm to ask whether the two scenes that share a texture ensemble or both texture and location show repetition suppression in the PPA. Both methods test representational properties of the PPA, but they capture different aspects of neural similarity. In DISCUSSION, we propose a working hypothesis that can reconcile the observed discrepancy between MVPA and repetition suppression and suggest that the combined use of MVPA and repetition suppression methods allows us to investigate different levels of neural organization in an area.

MATERIALS AND METHODS

Experiment 1

Subjects. Eleven participants (4 women, 7 men; 2 left-handed; ages 19–28 yr) were recruited from the Johns Hopkins University community for financial compensation. One participant in *experiment 1* was excluded from the analysis because his/her PPA was not localized. All had normal or corrected-to-normal vision. Written informed consent was obtained, and the study protocol was approved by the

Institutional Review Board of the Johns Hopkins University School of Medicine.

Stimuli. To systematically manipulate the type and location of the textures within a scene, we generated artificial room images with the software SketchUp and Adobe Photoshop CS6. Images were presented in 600 × 600-pixel resolution (4.5° × 4.5° visual angle) in the scanner with an Epson PowerLite 7350 projector (type: XGA, brightness: 1,600 ANSI lumens). Each room image was composed of back wall, ceiling, floor, left wall, and right wall. The back wall included a door and light fixtures and was held constant across different images (Fig. 1*A*). Two kinds of texture composed a single texture "ensemble": for example, in one of the scenes, *texture 1a* was used on both the ceiling and floor and *texture 1b* was used on the left and right walls to comprise *texture ensemble 1*. The ceiling and floor consistently had the same texture, as did the left and right walls. A texture ensemble (e.g., *texture 1a* and *texture 1b*) was used for each image, instead of a single texture (e.g., *texture 1a* only), to enable the Texture and Location manipulation. For example, in half of the stimuli that have *texture ensemble 1*, the location of textures that comprise a texture ensemble was swapped, such that *texture 1a* was on the left and right walls and *texture 1b* was on the ceiling and floor. Changing the location of textures while keeping the same texture ensemble was critical to test the Texture and Location hypothesis, because this hypothesis predicts that the PPA will be sensitive not only to the texture but also to the combination of the texture and its location.

In *experiment 1*, there were eight different stimuli, each representing one condition (Fig. 1*A*). These eight conditions were created from the combination of three factors: 2 texture ensembles (Ens) × 2 texture locations (Loc) × 2 viewpoints (View). First, the texture ensemble was changed by changing the kinds of texture used in an image (e.g., hypothetically, *ensemble 1* is composed of wood and brick; *ensemble 2* is composed of stone and bamboo). Second, different combinations of the texture and location were generated, while keeping the texture ensemble constant. Within the images that share the same texture ensemble (e.g., *ensemble 1*), the location of texture was changed. For example, the wood texture (on the ceiling and floor) and the brick texture (on the left and right walls) swapped locations so that the brick texture was on the ceiling and floor and the wood texture was on the left and right walls. Critically, this manipulation allowed us to test whether the PPA represents the texture ensemble in a scene (Cant and Xu 2012) regardless of their location in a scene. Third, viewpoints of the images changed within the images that share the same texture ensemble and texture location. Half of the stimuli were front views, while the other half were side views. Note that this viewpoint change, unlike the spatial structure change in Lowe et al. (2016), did not necessarily alter the spatial layout of the room. For that reason, we did not expect a strong modulation by the viewpoint, and it was not a factor of interest. Nonetheless, this factor

was included in the analysis to test whether the PPA shows an effect of a specific viewpoint over the texture ensemble or its location.

Stimuli ratings. Before scanning, an independent behavioral experiment was run to test whether participants recognized the change in texture location. Behavioral ratings were obtained from Amazon Mechanical Turk ($n = 26$) for all possible pairs within the complete stimuli set. Participants saw pairs of images and were asked to judge the likelihood that the two images were from the same room. Responses were selected from four possible options: 1 (definitely different rooms), 2 (somewhat different rooms), 3 (somewhat the same room), or 4 (definitely the same room).

The average rating difference between the same and different texture ensemble conditions was significant [$t(31) = 7.63$, $P < 0.01$]. This indicates that images that shared the same texture ensemble were considered to be more similar to each other than to the images that had different texture ensembles. On the other hand, within the same texture ensemble, there was a significant difference between two scenes that had the same texture but differed in location [$t(15) = 80.47$, $P < 0.01$]. In other words, even when images shared the same texture ensemble, if the location of specific texture was different within a scene people perceived them as different places. This indicates that participants were able to recognize changes in the texture location within our stimuli set and that two scenes with the same texture ensemble are recognized as two different places when the texture locations within a scene differ.

Experimental design. Experiment 1 consisted of six runs [each 5.3 min, 160 repetition time (TR)]. In each run, there were eight experimental conditions with five trials per conditions, resulting in 40 trials per run. Each image was displayed for 1.5 s, followed by 0.5 s blank with a jittered interstimulus interval (2–14 s; average of 6 s). Participants were asked to judge whether the current image was the same as or a different room from the previous image (Fig. 1*B*) by pressing a corresponding button (1: the same room, 2: different rooms).

In both experiments, the trial order was optimized by using Optseq2 (http://surfer.nmr.mgh.harvard.edu/optseq) to maximize the efficiency of estimating beta weights for a fast event-related design. The stimuli presentation and the experiment program were produced and controlled by MATLAB and Psychophysics Toolbox (Brainard 1997; Pelli 1997).

Multivoxel pattern analysis and representational similarity analysis. In *experiment 1*, MVPs of activity across time were extracted from the PPA. A general linear model (GLM) was run with six motion parameters added as predictors of no interest, and each condition was modeled with 10 finite impulse response (FIR) functions. An impulse function was set every TR (2 s); thus the model covered 20 s after the onset of stimulus. To obtain the statistically valid peak response for each condition, we ran a group-based *t*-test, following a conventional method that was used in many previous studies (Cant and Xu 2012;
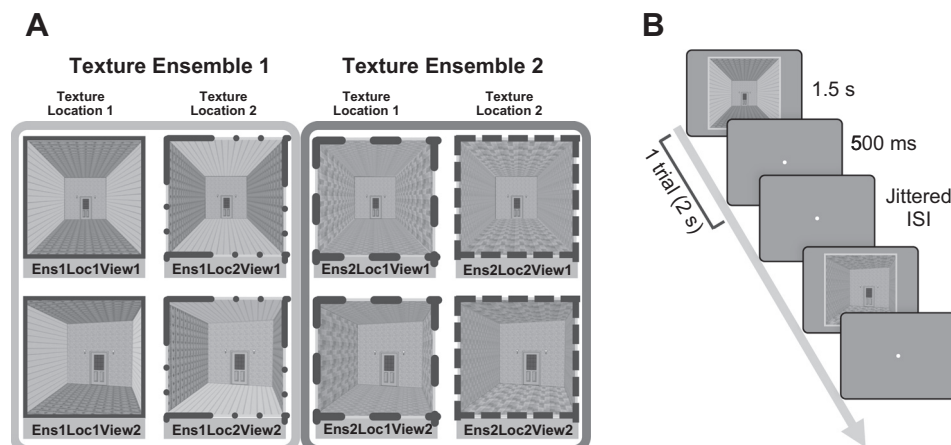
A

Texture Ensemble 1 | Texture Ensemble 2

Ens1Loc1View1, Ens1Loc2View1, Ens2Loc1View1, Ens2Loc2View1
Ens1Loc1View2, Ens1Loc2View2, Ens2Loc1View2, Ens2Loc2View2

B

1 trial (2 s) — 1.5 s — 500 ms — Jittered ISI

Fig. 1. *Experiment 1* conditions and procedure. *A*: gray-scaled versions of stimuli used in the present study. The combination of 3 factors (Texture Ensemble, Texture Location, and Viewpoint) resulted in 8 different conditions. Four images on *left* (e.g., Ens1Loc1View1) share *texture ensemble 1*, and 4 images on *right* (e.g., Ens2Loc1View1) share *texture ensemble 2*. Images in the same column (e.g., Ens1Loc1View1 and Ens1Loc1View2) share the same "Texture and Location" combination. *B*: in each trial, an image was displayed for 1.5 s, followed by 0.5 s blank and a jittered interstimulus interval.

Dilks et al. 2011; Kourtzi and Kanwisher 2001; O'Craven et al. 1999; Xu and Chun 2006).

After running the GLM with 10 FIR functions, we obtained 10 betas for each voxel for each condition. We averaged the betas across voxels for each condition, resulting in an 8 (conditions) × 10 (time points) matrix. Then, we averaged the betas across conditions, resulting in a 1 × 10 (time points) matrix. Among those 10 time points, we identified the numerically highest one (e.g., 5th time point). However, the difference between the numerical peak and its neighboring time points (e.g., 4th and 6th time points if the 5th time point is the numerical peak) could be marginal. To account for such cases, we did a pairwise t-test between the numerical peak and each of its neighboring time points (t test, $P < 0.05$, 1-tailed; Epstein et al. 2003; Marois et al. 2004; Park et al. 2007). If there was a significant difference, then only the numerical peak was chosen. If there was not a significant difference, then the compared neighboring time point was also chosen. In case where there was more than one chosen time point, the average value of the chosen time points was used for further analysis.

Then we constructed a pattern similarity matrix, using MVPs (beta) of the statistical peak time point. A pairwise correlation (Pearson r) was computed for MVP of all possible condition pairs. This generated an 8 × 8 similarity matrix of the correlations between each pair of conditions. The similarity matrix was computed separately for each individual participant. These similarity matrices were later averaged together for a groupwise similarity matrix.

To test our main hypotheses, we ran a multiple linear regression with the fMRI similarity matrix. The regression analysis was done with both the groupwise similarity matrix and each participant's similarity matrix to confirm the results from the group level. Predictive representational model matrices were constructed based on each hypothesis and used as predictors in the regression analysis. The Texture-Only predictor is based on the hypothesis that the PPA is sensitive to the texture ensemble only, independent of its location and viewpoint within a scene. According to this hypothesis, the conditions that share the same texture ensemble (e.g., Ens1Loc1View1 and Ens1Loc2View1) will be represented similarly in the PPA (Fig. 2*A*)

regardless of different texture locations. The Texture and Location predictor is based on the hypothesis that the PPA conjointly represents the texture and its location information within a scene. In this case, if two scenes have different combinations of texture and location, they will be represented differently, even when they share the same texture ensemble (Fig. 2*B*). The Viewpoint predictor is based on the hypothesis that the PPA is sensitive to a specific viewpoint of a scene independent of texture ensemble and texture location in a scene (Fig. 2*C*). We included the three predictors in the regression analysis to predict the PPA similarity patterns. Since half of off-diagonal cells (i.e., the lower left and the upper right triangle in the matrix) had repeated values, we excluded those cells, as well as the diagonal cells, when computing the correlation (Kriegeskorte et al. 2008).

Since our predictive matrices are not orthogonal to each other, we computed the variance inflation factor (VIF) to address the multicollinearity issue. The VIF is one of the most widely used diagnostics for multicollinearity, and this estimates how much the variance of coefficient is inflated because of the dependence among predictors. When there are two predictors in the regression, the VIF between Texture-Only and Texture and Location is 1.3. When there are three predictors in the regression, the VIFs between one factor and the other two factors are 1.3 for Texture-Only, 1.4 for Texture and Location, and 1.1 for Viewpoint. VIF's lower bound is 1, and the upper bound is not clearly defined, but a VIF smaller than 2 is generally considered as little or no collinearity.

It is noteworthy to mention that our predictive matrices are different from confusion matrices, the sum of which needs to be matched for each row between the matrices. In case of confusion matrices, true values are known and the actual performance is compared with them. Thus all predictive confusion matrices must be matched to the true values. However, the predictive models in the present and many other studies using representational similarity analysis (Kriegeskorte et al. 2008; Kriegeskorte and Kievit 2013; Nili et al. 2014; Watson et al. 2016) are conceptual models reflecting hypotheses about the representations of a specific brain region. In other words, in our predictive models there are no known true values to be matched.
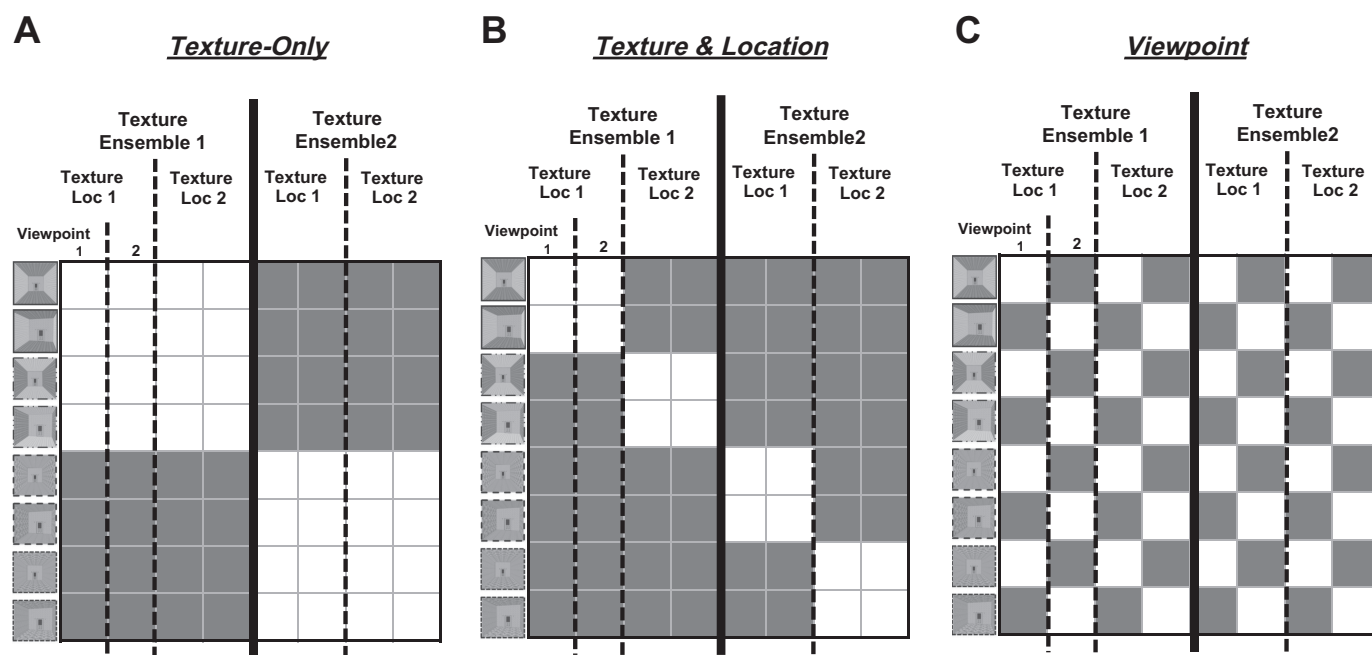

Fig. 2. Predictive representational similarity matrices. Three predictive representational matrices are constructed. The white cells indicate high similarity between 2 conditions, and the gray cells mean low similarity between the pair. *A*: the first model is based on the Texture-Only hypothesis, so conditions that share the same texture ensemble are colored white. *B*: the second model is based on the Texture and Location hypothesis, so conditions that share both same texture ensemble and same location are colored white. *C*: the third model is based on the Viewpoint factor, so conditions that share the same viewpoint are colored white.

A multidimensional scaling (MDS) plot was constructed to visualize the similarity structure of the activation patterns in the PPA, using the RSA (Representational Similarity Analysis) Toolbox (Kriegeskorte et al. 2008; Nili et al. 2014). The MDS plot is a useful way to visualize representational similarity, as it visually shows the degree of similarity as proximity between conditions. Because the location of each dot (condition) on the plot is determined in a way that preserves the distance between each pair of conditions, conditions that have high similarity in patterns of activity in the PPA will be plotted close to each other, forming a cluster.

In addition to multiple linear regression, we used a rank correlation method (Nili et al. 2014) to test *experiment 1* data. We computed a rank correlation (Kendall τ_A) between our fMRI data and each of the three representational models (Henriksson et al. 2015; Nili et al. 2014; Skerry and Saxe 2015; Wardle et al. 2016). In addition, we performed two reliability tests as specified below.

First, we tested for potential biases due to an inequality of signal energy among the model matrices. For example, it might be possible that a model with more 1s (i.e., white cells in Fig. 2) than other models could always lead to a higher correlation than other two models. To test this, we performed simulation analyses for two major cases. First, we simulated a case in which the data represent an ideal pattern for each model types by adding noise to each model matrices. Second, we simulated a case in which the actual data is purely noise. We generated a similarity matrix according to each hypothetical case and then compared this simulated data to each model matrix using a rank correlation (Kendall τ). Such simulation was repeated 10,000 times, and the mean correlation coefficients for each model matrix were computed. We found that a difference in model matrix energy does not prevent an appropriate model from producing the highest correlation value among three models. For example, even when the model matrix energy is the lowest, it was able to produce the highest correlation coefficient for the simulated data based on its own ideal pattern. For the simulation case in which the actual data are purely noise, we observed differences in correlation coefficients among the models. To account for this potential bias between models, the correlation coefficient values from this simulation were subtracted from each model's actual correlation coefficients.

Second, we computed the noise ceiling, which is the expected correlation that can be obtained from the "true" model. This true model was estimated by the average similarity matrix of all subjects, and the lower and upper bound of the noise ceiling were calculated using the subject variance in the data (Nili et al. 2014): the lower bound was computed by averaging each individual's correlation to the average of the other subjects' similarity matrices, whereas the upper bound was computed by averaging each individual's correlation to the average of all subjects' similarity matrices.

Experiment 2

Subjects. Eighteen participants (10 women, 8 men; 1 ambidextrous; ages 18–30 yr) were recruited from the Johns Hopkins University community for financial compensation. All had normal or corrected-to-normal vision. Written informed consent was obtained, and the study protocol was approved by the Institutional Review Board of the Johns Hopkins University School of Medicine.

Stimuli. In *experiment 2*, the images were almost identical to those used in *experiment 1*, except that *experiment 2* had 256 texture ensembles and one viewpoint. Similar to *experiment 1*, two kinds of texture were used for every image to form a texture ensemble: one texture on the ceiling and floor and another texture on the left and right walls. To avoid the repetition of the same image, we made 256 different room images with nonrepeating textures. We then swapped the texture on the ceiling and floor with the texture on the left and right walls of these 256 images to obtain another 256 images that have the same texture ensemble but different texture locations. Since the same kind of stimulus manipulation was used in *experiment 1*, no stimuli ratings were collected.

Experimental design. Experiment 2 consisted of four runs (each 5.8 min, 174 TR). In each run, there were three experimental conditions of 16 trials, resulting in 48 trials per run. The number of each image appearing in each condition was counterbalanced for every six participants. For a set of six participants, the stimuli (256 images) were first randomly divided into four groups (a group for each run), which were further divided into four subgroups. Among these subgroups, one subgroup was used for SameTexture-SameLocation, another subgroup for SameTexture-DifferentLocation, and the other two subgroups for the DifferentTexture-DifferentLocation condition (Fig. 3A). The subgroups in each run were assigned according to the predetermined combination, and the presentation order of images within each subgroup was randomly decided. After running six participants, the whole stimuli set of 256 images were shuffled again and used for another set of six participants.

Each trial was 4 s (2 TR) long: 500-ms fixation, followed by two images that were sequentially presented for 1 s each (each stimulus was presented for 600 ms, followed by 400-ms blank period), with a 1.5-s blank period at the end (Fig. 3B). Each trial was followed by a jittered intertrial interval with an average of 3 s (2–8 s). Participants observed two images and responded as to whether the two images
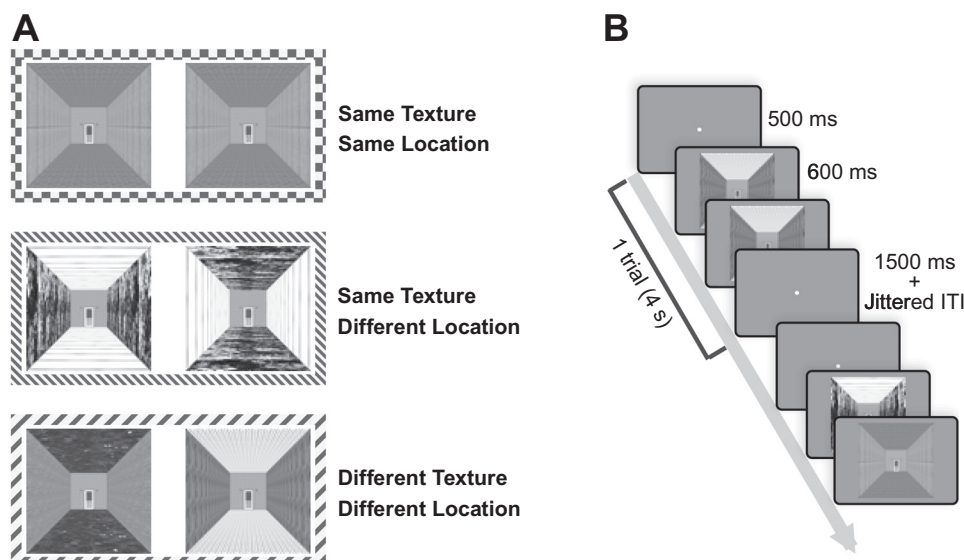

Fig. 3. *Experiment 2* conditions and procedure. *A*: there were 3 experimental conditions in which an image could repeat: Same Texture-Same Location condition, Same Texture-Different Location condition, and Different Texture-Different Location condition. *B*: for each trial, 2 images were presented in sequence. Each stimulus was presented for 600 ms, followed by a 400-ms blank screen. At the end of each trial, there was a 1.5-s blank period and a jittered intertrial interval (ITI).

were from the same room or different rooms by pressing the button (1: the same room, 2: different rooms).

Repetition suppression. In *experiment 2*, we ran a GLM within each region of interest (ROI) (ROI-based GLM), with six motion parameters added as predictors of no interest. Each condition was modeled with 10 FIR functions. We obtained the statistical peak time point by using the same method as in *experiment 1*. The peak beta values of each condition were compared by the pairwise *t*-test. The critical difference between *experiments 1* and *2* was that *experiment 2* used univariate responses that were averaged across multiple voxels within an ROI.

MRI acquisition and preprocessing. Imaging data were acquired with a 3-T Philips fMRI scanner with a 32-channel phased-array head coil at the F. M. Kirby Research Center for Functional Neuroimaging at Johns Hopkins University. Structural T1-weighted images were acquired by magnetization-prepared rapid-acquisition gradient echo (MPRAGE) with $1 \times 1 \times 1$-mm voxels. Functional images were acquired with a gradient echo-planar T2* sequence [$2.5 \times 2.5 \times 2.5$-mm voxels; TR 2 s; TE 30 ms; flip angle = 70°; 36 axial 2.5-mm sliced (0.5-mm gap); acquired parallel to the anterior commissure-posterior commissure (ACPC) line]. Functional data were analyzed with Brain Voyager QX software (Brain Innovation, Maastricht, The Netherlands). Preprocessing included slice scan-time correction, linear trend removal, and three-dimensional motion correction. No spatial or temporal smoothing was performed, and the data were analyzed in individual ACPC space.

Parahippocampal place area. The PPA was defined for each participant with a functional localizer independent of the main experimental runs. In a localizer run, participants saw blocks of faces, scenes, objects, and scrambled object images. A localizer run consisted of four blocks per each of these four image conditions, each block with 20 images (7.1 min, 213 TR). In each trial, a stimulus was presented for 800 ms. Participants performed one-back repetition detection tasks, in which they were asked to press a button whenever they detected an immediate repetition of an image.

The PPA of each participant was functionally defined by contrasting brain activity of scene blocks and face blocks (Epstein and Kanwisher 1998). A cluster of contiguous voxels in the posterior parahippocampal gyrus and collateral sulcus region that passed the threshold ($P < 0.0001$) was defined as the PPA.

RESULTS

Experiment 1

We compared the representational similarity matrices generated from the actual fMRI data (Fig. 4A) to the three hypothetical models. First, we ran a multiple linear regression to compare the similarity matrix from the fMRI data to the model matrices. Three predictors, Texture-Only hypothesis, Texture and Location hypothesis, and Viewpoint, were put into the regression model. We tested which predictor fits better with the group-level similarity matrix from fMRI data, which was the averaged similarity matrix of all participants. The beta coefficients from the regression were 0.014 ($P = 0.001$) for the Texture-Only predictor, -0.0061 ($P = 0.31$) for the Texture and Location predictor, and 0.0003 ($P = 0.93$) for the Viewpoint predictor. Only the Texture-Only predictor had a significant beta coefficient, which means that only the Texture-Only predictor, and not the Texture and Location or Viewpoint predictor, was related to changes in the representational similarity matrix obtained from the fMRI data. To confirm this result, we also ran the multiple linear regression for each individual, separately estimating each participant's fMRI data with the predictive models. The average beta coefficients of individual data were compared among the models. One-way ANOVA showed that there is a significant difference among the beta of the models [$F(2, 27) = 9.14$, $P < 0.001$]. Then, we compared the difference between each pair of predictors, using a paired *t*-test (Fig. 4B). The beta of Texture-Only was significantly higher than that of Texture and Location [$t(9) = 3.08$, $P = 0.01$] and Viewpoint

Fig. 4. *Experiment 1* results. *A*: representational similarity matrix was constructed from the actual fMRI data. The conditions that share the same texture ensemble showed higher correlation to each other (lighter gray cells), similar to the hypothetical Texture-Only model. *B*: result of the multiple linear regression. Only the Texture-Only predictor could significantly explain the variance in the fMRI data. *P < 0.05; n.s., not significant. *C*: multidimensional scaling (MDS) plot. Each dot represents a condition. Dark gray dots indicate the conditions that have *texture ensemble 1*, whereas light gray dots represent the conditions that have *texture ensemble 2*. The frames around each stimulus image are depicted based on the specific combination of texture and location (as in Fig. 1*A*). This plot shows that the dots with the same shade of gray (same texture ensemble) are positioned toward the same side, which is consistent with the Texture-Only hypothesis. However, no such pattern was observed according to the Texture and Location combination or Viewpoint.
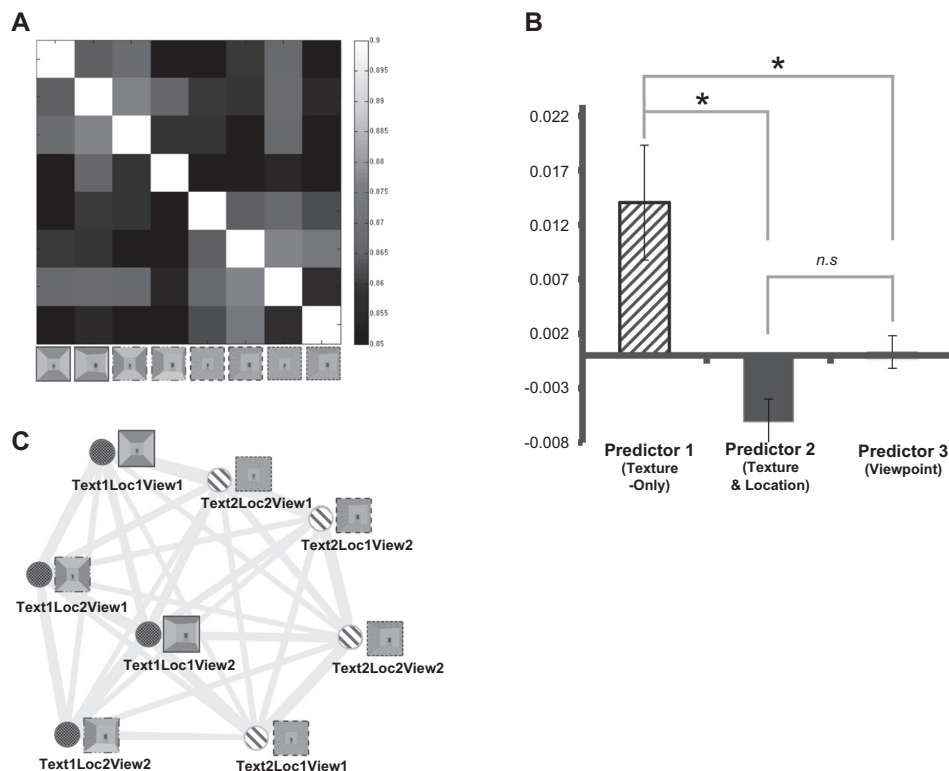
$[t(9) = 2.79, P = 0.02]$. However, there was no significant difference between betas of Texture and Location and Viewpoint $[t(9) = -2.18, P > 0.05]$. The results suggest that the Texture-Only predictor explains the PPA's response significantly better than the other two predictors.

We also used rank correlation method (Nili et al. 2014) and replicated results from the multiple linear regression, suggesting that the Texture-Only model is a significantly better model than the other two (Texture-Only model $\tau = 0.1038$; Texture and Location model $\tau = 0.0021$; Viewpoint model $\tau = 0.0158$). This was true after accounting for a potential model bias (see MATERIALS AND METHODS for more details). Furthermore, the correlation coefficient for the Texture-Only model was just below the lower bound of the noise ceiling, suggesting that this model can highly account for the non-noise variance in the data.

Figure 4*C* illustrates a two-dimensional MDS plot based on the averaged similarity matrix. Each dot in the plot represents a condition; the conditions that elicit similar response patterns are placed close to each other, whereas the conditions that elicit different response patterns are placed far from each other. By visualizing representations of the conditions in such a way, this two-dimensional MDS plot allows us to easily observe the representational geometries of our stimuli in the PPA. Consistent results were also found in a three-dimensional MDS plot, in which the third dimension was allowed to emerge. Thus we chose to report the two-dimensional plot for simplicity's sake. Figure 4*C* shows that the conditions that share *texture ensemble 1* and those that share *texture ensemble 2* are positioned toward different ends of the horizontal axis, consistent with the Texture-Only hypothesis but not the other two hypotheses. In summary, *experiment 1* results support the Texture-Only hypothesis and suggest that the PPA represents the kind of textures in a scene, regardless of the location of a texture in a scene.

Experiment 2

The results of *experiment 1* suggest that the PPA represents texture ensemble information but not the location information of a texture within a scene, at least not in the MVPs. Solely on the basis of the *experiment 1* data, however, we could not draw a definite conclusion about whether or not the combined information of texture and location is represented in the PPA.

There were both theoretical and technical reasons that led to further testing of the texture representation in the PPA. First, participants consistently rated stimuli that shared the same texture and same location as similar and stimuli that had the same texture and different locations as different. However, MVPA results from *experiment 1* did not reflect such behavioral similarity judgment, which required further testing.

Second, Kornblith et al. (2013) showed the existence of neurons that are tuned to the combination of texture and spatial information in the macaque monkey brain. However, that was a single-cell recording study, which suggests that the conjoint information of texture and space might be represented at the neuronal level. If so, MVPs used in *experiment 1* may not be an ideal source from which to observe such conjoint representation. MVPA treats a single-voxel response as a minimum unit, which contains a huge number of neurons (~630,000 neurons in a $3 \times 3 \times 3$-mm functional voxel). Hence, it might

be problematic to assume homogeneity of neural responses within a voxel. For example, there are many different patterns of neural response that would produce the same average response at the voxel level. However, since the MVPA measures patterns of the average voxel activity, the MVPA method will be insensitive to such different cases. That is, if we are interested in a representation with a scope that is smaller than a voxel, then we need a more fine-grained measurement than MVPA. Repetition suppression can be a useful alternative method in this case. The basic assumption of repetition suppression is that a neuron activated by the same stimulus for the second time shows a reduced amount of activation (Grill-Spector and Malach 2001). Since this assumption is applied at neuronal level, instead of at voxel level, the different representations within a voxel may be detected by using the repetition suppression method (Drucker and Aguirre 2009; Grill-Spector and Malach 2001). We come back to this point in DISCUSSION.

Third, several recent papers have reported discrepant results from MVPA and repetition suppression (Drucker and Aguirre 2009; Epstein and Morgan 2012; Ward et al. 2013), suggesting that MVPA and repetition suppression may be sensitive to different levels of representation, such as the representation across multiple voxels and the representation at a more local, or within-voxel, level. Moreover, a recent study (Hatfield et al. 2016) showed that combining results from MVPA and repetition suppression analysis provides a richer interpretation of the topography of a neural representation. Motivated by these three reasons, we tested our hypotheses again in *experiment 2* by using repetition suppression.

In *experiment 2*, we used average univariate beta estimates of the peak time point for each condition and compared them to the amount of activation for each condition. To test the significance of repetition suppression results, pairwise *t*-tests were conducted for each pair of conditions. The SameTexture-SameLocation condition served as the lower baseline, because the pairs of images presented in the SameTexture-SameLocation condition were completely identical. The critical question concerns the relative level of activation of the SameTexture-DifferentLocation condition, which has image pairs that share the same texture ensemble but have different location of textures within each image. If the image pair in SameTexture-DifferentLocation is represented as similar in the PPA, then the activation level of SameTexture-DifferentLocation should be similar to that of SameTexture-SameLocation. This is what the Texture-Only hypothesis predicts, because the image pair in SameTexture-DifferentLocation shares the exact same texture ensemble. On the other hand, if the images in SameTexture-DifferentLocation pairs are represented as different from each other, the activation level of SameTexture-DifferentLocation should be significantly greater than that of SameTexture-SameLocation; we would expect the activation level of SameTexture-DifferentLocation to be similar to that of DifferentTexture-DifferentLocation, which plays a role as the upper baseline. This is what the Texture and Location hypothesis predicts, because the combination of texture and location changes from that of the SameTexture-SameLocation condition, even though the texture ensemble remains the same.

A paired-samples *t*-test showed that the SameTexture-SameLocation condition had a significantly lower response than the DifferentTexture-DifferentLocation condition $[t(10) = -2.21,$

$P < 0.05$]. This confirms the significant repetition suppression effect for the SameTexture-SameLocation condition. Most critically, the SameTexture-DifferentLocation condition had a significantly higher response than the SameTexture-SameLocation condition [$t(10) = -3.06$, $P < 0.05$], suggesting that the PPA is sensitive to the change of texture location in a scene even when the texture ensemble is kept the same. The difference of repetition suppression effects between SameTexture-DifferentLocation and DifferentTexture-DifferentLocation is not significant [$t(10) = 0.98$, $P = 0.34$; Fig. 5], suggesting that the effect of changing texture location was as great as changing the texture itself. The result is in line with the prediction from the Texture and Location hypothesis. This contradicts with results from *experiment 1* on the surface, which supported the Texture-Only hypothesis. We discuss this contradiction in more depth in DISCUSSION.

DISCUSSION

In this study, we investigated how the texture information in a scene is represented in the PPA. Two competing hypotheses were tested. The Texture-Only hypothesis states that the PPA represents texture as is; that is, the PPA is sensitive to the change in texture ensemble (kinds of textures) irrespective of the texture's spatial locations within a scene. The Texture and Location hypothesis states that the PPA encodes texture conjointly with location in a scene. In this case, the PPA will be sensitive to changes in the position of texture even when texture itself has not changed.

The critical difference between the two hypotheses is that they make different predictions on whether or not the PPA distinguishes images that share the same texture ensemble but have different locations of texture. To test which hypothesis explains the neural representation in the PPA better, we created stimuli in which the texture ensemble was the same but the location of the texture is different. Interestingly, we found different results in *experiment 1* and *experiment 2*. In *experiment 1*, the MVPs of the PPA activation did not show any

significant difference among scenes with different combinations of texture and location. The representational similarity matrix, made from MVPs, indicated that the PPA is sensitive to the difference between texture ensembles but not to the difference between combinations of texture and location. In contrast, in *experiment 2*, we found a repetition suppression effect to images that share the same texture and location combination but not to images that only share the same texture ensemble. This result indicates that the PPA is sensitive to the spatial information (e.g., specific location within a scene), in addition to the texture ensemble.

MVPA and Repetition Suppression

The observed difference between *experiment 1* and *experiment 2* results might be explained by the inherent differences between two methods we used: MVPA and the repetition suppression paradigm.

Several recent studies have reported discrepant results across these methods, similar to the findings in this study. Such studies evoke further questions and provide an opportunity to see a broader picture about what each method measures. Epstein and Morgan (2012) examined the PPA's representation of scene categories and landmarks. With MVPA, they found a successful classification of both category and landmarks in the PPA. In contrast, they found the repetition suppression effect only for landmark images and not for category images (except for a slight effect in a medial part of the left PPA). In other words, the MVPA method could successfully distinguish the stimuli at both the category and exemplar (in this case, landmark) levels, whereas repetition suppression was sensitive to only exemplar-level stimuli such as landmarks (Epstein and Morgan 2012). Another study showed a clear difference between MVPA and repetition suppression, and each method was linked to explicit and implicit memory, respectively (Ward et al. 2013). In this study, participants were shown repeated scene images and asked to categorize them. The amount of behavioral repetition priming was predicted by the repetition sup-
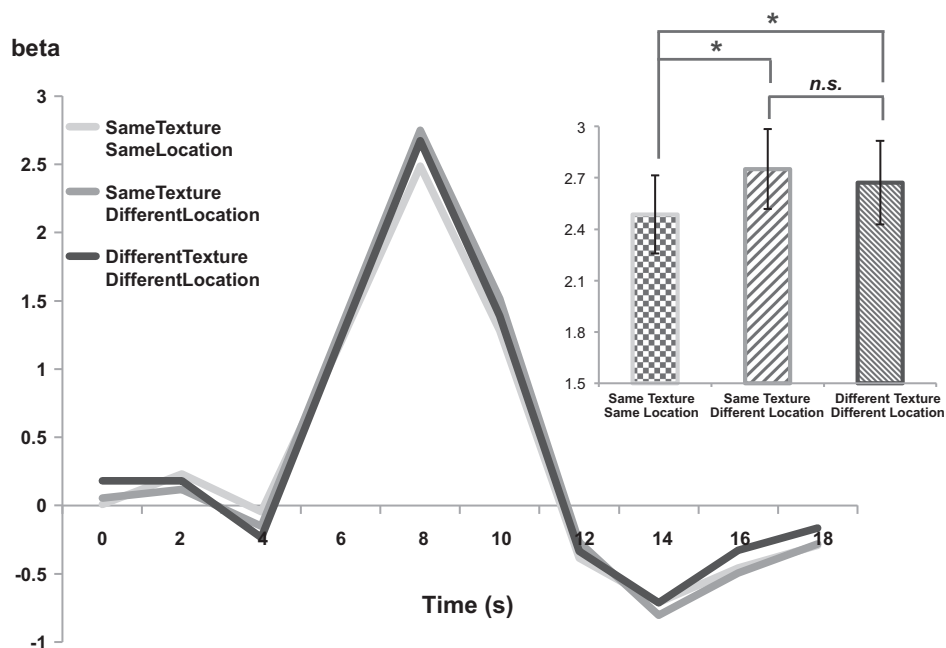

Fig. 5. *Experiment 2* results. Average hemodynamic responses for *experiment 2* conditions are shown for PPA. Average peak beta estimates are shown as a bar graph depicting the statistical peak of the time course. *Peak beta estimates showed a significant difference between the SameTexture-SameLocation condition and the other 2 conditions. n.s., The SameTexture-DifferentLocation condition and DifferentTexture-DifferentLocation condition did not show any statistically significant difference.

pression effect in the occipitotemporal region but not by the MVP similarity. Conversely, the subsequent recognition performance was predicted by the pattern similarity in the occipitotemporal region but not by the repetition suppression effect. These reports add weight to the idea that MVPA and repetition suppression might be based on different neural mechanisms.

The question then arises: What mechanism underlies this difference between MVPA and the repetition suppression paradigm? To further discuss possible reasons, we first need to clarify what each method measures. The MVPA (Haxby et al. 2001; Haynes and Rees 2005; Kamitani and Tong 2005) measures a distributed pattern of neural activity across voxels. If MVPs are similar between two stimuli, this indicates that similar voxels are activated. When the same voxel is activated, however, it is not obvious how neurons within each voxel are responding. As briefly mentioned in RESULTS, the minimum unit we are looking at in fMRI studies is a voxel, and there are a large number of neurons within each voxel. This fact suggests a possibility that neurons with different tuning curves or response patterns may coexist within a voxel (Drucker and Aguirre 2009; Epstein and Morgan 2012). Since what we can observe using fMRI is the average activation of all neurons within a voxel, the fact that the same voxel is activated could mean either that the same set of neurons are activated or that different sets of neurons are activated but they are located within the same voxel.

Repetition suppression, on the other hand, is sensitive to within-voxel patterns (Aguirre 2007; Drucker and Aguirre 2009; Grill-Spector and Malach 2001). Although multiple models have been proposed to explain the neural mechanism underlying the repetition suppression effect, there is no unified answer yet. However, all models share a common assumption that the repetition suppression effect is observed when the same set of neurons respond to two stimuli. Thus when the repetition suppression is observed over two stimuli, this suggests that the same set of neurons is activated by the stimuli. Thus by looking at the repetition suppression effect one can speculate how much overlap there is in population of neurons within a voxel that respond to each stimulus. Taken together, MVPA and repetition suppression both examine neural similarity but at different levels: MVPA measures similarity across voxels, and repetition suppression measures similarity at the within-voxel level.

Neural Organization of Texture Representation in PPA

With the above framework in mind, we now interpret our results in terms of different neural levels. First, MVPA results suggest that different texture ensembles evoke different patterns of activation across voxels, so we will assume that different sets of voxels are involved in representing different texture ensembles (e.g., *voxel 1* vs. *voxel 2*; Fig. 6). At the same time, MVPA results also suggest that the same texture ensemble in a scene is represented similarly at the across-voxel level. That is, images that share the same texture ensemble, regardless of the texture's location in a scene, are likely to cause a similar level of activation for a given voxel. Such a result can be driven by at least two cases. In the first case, two images stimulate the same set of neurons. In this case, the repetition suppression effect is expected across the presentation of those two images. In the second case, two images stimulate different sets of neurons, but these different sets of neurons are anatomically close to each other, such that they fall within the same voxel. In this case, the repetition suppression effect is hardly expected, because only a negligible number of neurons will be activated for both stimuli.

In Fig. 6, the pattern that would be observed with MVPA is shown as small parallelograms labeled *voxel 1* and *voxel 2*. For MVPA, we consider only the voxel-level pattern for each "initial" stimulus presented in the figure. Note that the shaded patterns of *voxel 1* and *voxel 2* are the same for Fig. 6, *A* and *B*, but different for Fig. 6, *A* and *C*. In other words, in this hypothetical case we would not see different patterns across voxels for the stimuli that have the same texture in different locations in a scene, as long as they share the same texture ensemble.

Crucially, the similar pattern at the across-voxel level does not guarantee the same neural patterns at the within-voxel level, which is required for repetition suppression to occur. If we observe repetition suppression between the stimuli that share the same texture ensemble but with different locations of each texture, it would suggest that there is a substantial number of neurons that respond to both stimuli. On the other hand, however, if the repetition suppression does not occur for those stimuli, it would suggest a possibility that different sets of neurons within the same voxel are activated for those stimuli. For example, we can assume a situation in which there are four neurons (e.g., *N1*–*N4*, Fig. 6) in a voxel (e.g., *voxel 1*) with different tuning curves, such that *N1* and *N2* respond to the stimulus Ens1Loc1View1 and *N3* and *N4* respond to the stimulus Ens1Loc2View1 (Fig. 6). In this case, presentation of the stimuli Ens1Loc1View1 and Ens1Loc2View1 in sequence would not cause repetition suppression, because they evoke different populations of neurons within this particular voxel. However, even if they evoked different neurons, their output at the voxel level would not be distinguishable.

Therefore, we propose that texture information in the PPA is represented at multiple levels: texture ensemble information is represented at the voxel level, and conjoint texture ensemble and texture location information is represented at the within-voxel level. Although we cannot make a claim about within-voxel patterns on the basis of the MVPA results, this does hint at the topographical organization of the PPA when representing texture information. Note that the "topography" here refers to how a neuronal population that represents certain information (e.g., texture ensemble) is clustered closely together (e.g., within a voxel) or not (e.g., across voxels). In our results, when two images share the same kinds of texture, even with the different texture locations in a scene, they were likely to activate similar MVPs. A similar MVP means that the same voxels are likely to be activated across the stimuli, suggesting that neurons activated by the two stimuli are likely to exist within the same voxel. Since each voxel is determined by the spatial location on the cortices, a high probability of being in the same voxel indicates that the neurons coding the stimuli are anatomically close to each other. Following this logic, the MVPA results hint about how neuronal populations exist topographically close together. Given our data, we propose that the PPA neurons representing the same texture ensemble are clustered together in such a way that they are more likely to fall in the same voxel, compared with the neurons representing
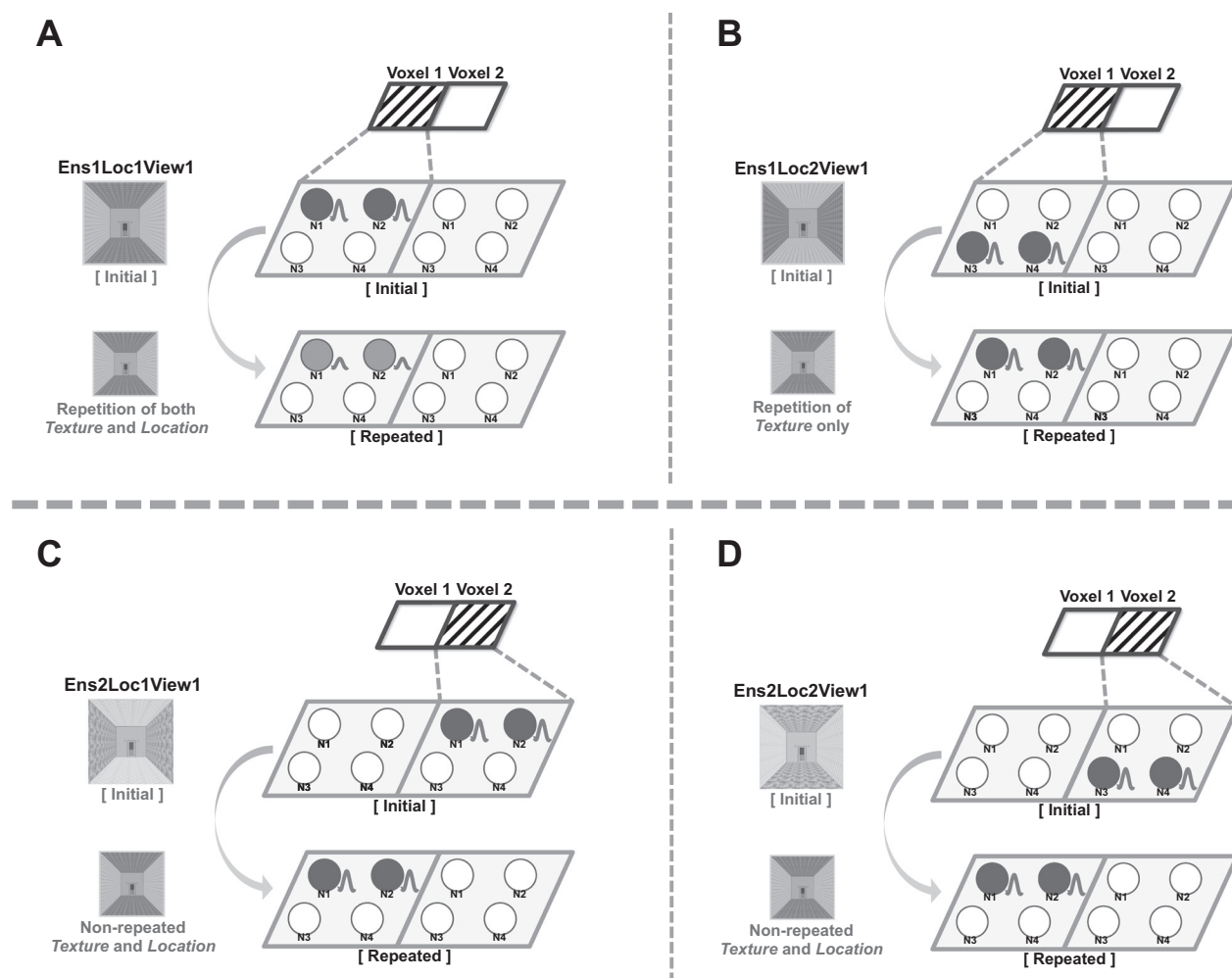
Fig. 6. Hypothetical models of texture representation in the PPA. Two small parallelograms at *top* of each diagram describe the pattern that is observable with MVPA. Note that only the stimuli that share the same texture ensemble show the same pattern at the voxel level (e.g., *A* and *B*). Four circles (*N1–N4*) within each parallelogram shape (voxel) represent neurons located in each voxel. The circles colored in gray indicate that they are activated. A particular set of neurons (e.g., *N1* and *N2* in *voxel 1*) are activated by presentation of a particular stimulus (e.g., Ens1LocView1). Each diagram (*A–D*) illustrates the presumed activation patterns of the PPA neurons for separate possible cases in the repetition suppression paradigm.

different texture ensembles. This kind of topographic representation is prevalent across the cortices, perhaps because it is an efficient way of organizing neurons such that the distance between them is minimized (Kaas 1997; Silver and Kastner 2009).

At the within-voxel level, texture ensemble information is represented conjointly with texture location. Although it is difficult to make a direct comparison to the monkey cell recording results, our results are in line with previous studies. Kornblith et al. (2013) showed that a significant number of neurons in scene-selective regions were modulated by the combination of texture and spatial factors. Similarly, *experiment 2* (repetition suppression paradigm) results suggest the existence of neurons that are sensitive to different combinations of texture and spatial location. As a note, it is entirely possible that there are also neurons responding to texture only, rather than conjointly with location, but that is perfectly compatible with the present results. Since those neurons tuned to a specific texture will respond to both conditions that share the same texture ensemble but in different locations (e.g., Ens1Loc1View1 and Ens1Loc2View1), they will not contribute to the difference in the repetition suppression effect.

Our results show the benefit and the significance of using both MVPA and the repetition suppression method in investigating a research question. These two methods are both widely used in current fMRI studies. Nevertheless, most studies concern the results from only one method, perhaps with an implicit assumption that MVPA and repetition suppression are alternative methods. However, the present study demonstrates that an answer to the research question could be different depending on the method, because each method investigates a different level of representation. Thus we suggest that MVPA and repetition suppression methods are complementary to each other and not necessarily interchangeable per se. For example, an interpretation of MVPA results could be complemented by results from the repetition suppression method. Without considering the different repetition suppression effect for the stimuli that share the same texture ensemble, the PPA neurons' sensitivity to the conjoint information of texture and location would not have been noticed. Conversely, by looking at the repetition suppression results only, we would not have learned about the topographic organization of texture ensemble in the PPA.

Consequently, the most comprehensive understanding of the neural representation would come from observing both levels,

which can be accomplished by combining MVPA and repetition suppression. Although we do not necessarily assert that both methods should always be used in a single study, an experimental design such as the continuous carryover design (Aguirre 2007) would be a useful way to observe both within-voxel and across-voxel patterns in a single experiment.

Conjoint Representation of Texture and Location

This study contributes to the existing understanding about the PPA's scene representation. The results of *experiment 2* (repetition suppression) suggest a complex relationship between the texture and the spatial information of a scene in the PPA. Texture patches used in previous experiments (Cant and Xu 2012; Jacobs et al. 2014) did not have cues for spatial structure (e.g., lines separating the floor from the left or right walls), which are crucial factors to define an image as a scene. Cant and Xu (2012) found a repetition suppression effect in the PPA for the patches with the same texture ensemble and argued that the PPA represents texture ensemble information. However, the present study extends such results by adding spatial structure to an image. When the texture was embedded in the spatial structure of a scene, the repetition suppression effect in the PPA was released and two images with the same texture ensemble were treated as different when location of texture changed.

The present study is in line with a recent study that investigated how scene-selective areas represent multiple visual features (e.g., texture and spatial structure), depending on the task demands. In Lowe et al. (2016), participants paid attention to either global texture or spatial layout of the stimuli, which consisted of four scene categories defined by the scene content (natural vs. manufactured) and the spatial boundary (open vs. closed). They suggest that texture and spatial layout are both crucial features in the scene processing, but the contribution or importance of each varies depending on the scene; overall, Lowe et al. (2016) support a greater sensitivity of the PPA to a scene structure than to texture. On the surface, this result may seem to contradict the present results. However, one critical difference between Lowe et al. (2016) and the present study reconciles the finding: the spatial layout changes of scenes used in Lowe et al. (2016) are fundamentally different from the location of texture changes in the present study. In Lowe et al. (2016), the level of spatial boundary was either "open" or "closed." In other words, changing the spatial layout indicated a substantial amount of structural change, such as the difference between an open desert image and a closed cave image. On the other hand, in the present study even when two stimuli had different locations of texture the structure or layout of the stimuli remained the same. In other words, if the texture and color information is removed from images that differ in location of texture, it will not be possible to differentiate the two of them. This suggests that the location of texture alone is not informative enough for scene recognition in the present study. To be a diagnostic feature for scene recognition, the location information needs to be conjointly encoded with texture information. On the contrary, if there is a structural change between the two stimuli, as in the Lowe et al. (2016) study, it will still be possible to differentiate the two images when the texture or color information is removed. This suggests that the spatial layout alone is informative for scene recognition in Lowe et al.

(2016). This is a critical difference, which makes it hard to make a direct comparison between the results of the present study and Lowe et al. (2016).

What would be the benefit of coding texture and its spatial information together? Separate from the question of how the PPA represents the texture information (e.g., Texture-Only vs. Texture and Location), it is also interesting to ask why the PPA represents the texture information in such a way. We speculate that texture information plays a crucial role in identifying and differentiating different scene images, especially when the spatial layout is not informative (e.g., rooms with identical geometric features). In such cases, the location of a specific texture can provide valuable information for determining a scene's identity (Epstein and Julian 2013). Taken together, the overall story is that the PPA is generally more sensitive to the structure information of a scene; however, if the structure information is not available, then texture information becomes the most salient diagnostic feature, resulting in higher sensitivity to texture in the PPA.

The place identity judgment (stimuli rating) result also supports this idea. It is shown that if two scenes have different combinations of texture and location, people perceive them as different places. This result suggests that people use the combined information of texture and location as a cue to judge the place identity of a scene. Perhaps this provides the ecological reasons for the PPA neurons to encode the texture and its location information together. This raises the possibility that PPA neurons are tuned to scene identity; however, testing such a possibility would require further study with conditions dissociating place identity from texture and location information.

Conclusions

The present study observed the representation of texture information within a scene, which is an aspect of scene perception that has not been explored much to date. We investigated how texture information is represented in the PPA using MVPA and repetition suppression. Throughout *experiments 1* and *2*, we provide evidence that MVPA and repetition suppression examine brain activation patterns at different levels: across-voxel and within-voxel, respectively. We propose that texture information within a scene context is represented in the PPA at multiple levels: the texture ensemble information is represented in the PPA at the across-voxel level, and the conjoint information of the texture and its location is represented at the within-voxel level. These results are in line with the rapidly evolving view that the PPA encodes not only spatial information but also nonspatial information (Cant and Xu 2012, 2015; Epstein and Julian 2013; Harel et al. 2013; Kornblith et al. 2013).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

J.P. and S.P. conceived and designed research; J.P. performed experiments; J.P. analyzed data; J.P. and S.P. interpreted results of experiments; J.P.

prepared figures; J.P. drafted manuscript; J.P. and S.P. edited and revised manuscript; J.P. and S.P. approved final version of manuscript.

REFERENCES

Aguirre GK. Continuous carry-over designs for fMRI. *Neuroimage* 35: 1480–1494, 2007. doi:10.1016/j.neuroimage.2007.02.005.

Aguirre GK, D'Esposito M. Topographical disorientation: a synthesis and taxonomy. *Brain* 122: 1613–1628, 1999. doi:10.1093/brain/122.9.1613.

Aguirre GK, Zarahn E, D'Esposito M. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* 21: 373–383, 1998. doi:10.1016/S0896-6273(00)80546-2.

Bar M, Aminoff E. Cortical analysis of visual context. *Neuron* 38: 347–358, 2003. doi:10.1016/S0896-6273(03)00167-3.

Bar M, Aminoff E, Ishai A. Famous faces activate contextual associations in the parahippocampal cortex. *Cereb Cortex* 18: 1233–1238, 2008. doi:10.1093/cercor/bhm170.

Barrash J, Tranel D, Anderson SW. Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. *Dev Neuropsychol* 18: 355–381, 2000. doi:10.1207/S1532694205Barrash.

Brainard DH. The psychophysics toolbox. *Spat Vis* 10: 433–436, 1997. doi:10.1163/156856897X00357.

Cant JS, Goodale MA. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb Cortex* 17: 713–731, 2007. doi:10.1093/cercor/bhk022.

Cant JS, Goodale MA. Scratching beneath the surface: new insights into the functional properties of the lateral occipital area and parahippocampal place area. *J Neurosci* 31: 8248–8258, 2011. doi:10.1523/JNEUROSCI.6113-10.2011.

Cant JS, Xu Y. Object ensemble processing in human anterior-medial ventral visual cortex. *J Neurosci* 32: 7685–7700, 2012. doi:10.1523/JNEUROSCI.3325-11.2012.

Cant JS, Xu Y. The impact of density and ratio on object-ensemble representation in human anterior-medial ventral visual cortex. *Cereb Cortex* 25: 4226–4239, 2015. doi:10.1093/cercor/bhu145.

Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N. Mirror-image sensitivity and invariance in object and scene processing pathways. *J Neurosci* 31: 11305–11312, 2011. doi:10.1523/JNEUROSCI.1935-11.2011.

Drucker DM, Aguirre GK. Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cereb Cortex* 19: 2269–2280, 2009. doi:10.1093/cercor/bhn244.

Epstein R, Graham KS, Downing PE. Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37: 865–876, 2003. doi:10.1016/S0896-6273(03)00117-X.

Epstein R, Harris A, Stanley D, Kanwisher N. The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23: 115–125, 1999. doi:10.1016/S0896-6273(00)80758-8.

Epstein RA, Julian JB. Scene areas in humans and macaques. *Neuron* 79: 615–617, 2013. doi:10.1016/j.neuron.2013.08.001.

Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 392: 598–601, 1998. doi:10.1038/33402.

Epstein RA, Morgan LK. Neural responses to visual scenes reveals inconsistencies between fMRI adaptation and multivoxel pattern analysis. *Neuropsychologia* 50: 530–543, 2012. doi:10.1016/j.neuropsychologia.2011.09.042.

Grill-Spector K, Malach R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107: 293–321, 2001. doi:10.1016/S0001-6918(01)00019-1.

Harel A, Kravitz DJ, Baker CI. Deconstructing visual scenes in cortex: gradients of object and spatial layout information. *Cereb Cortex* 23: 947–957, 2013. doi:10.1093/cercor/bhs091.

Hatfield M, McCloskey M, Park S. Neural representation of object orientation: a dissociation between MVPA and Repetition Suppression. *Neuroimage* 139: 136–148, 2016. doi:10.1016/j.neuroimage.2016.05.052.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425–2430, 2001. doi:10.1126/science.1063736.

Haynes JD, Rees G. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8: 686–691, 2005. doi:10.1038/nn1445.

Henriksson L, Khaligh-Razavi SM, Kay K, Kriegeskorte N. Visual representations are dominated by intrinsic fluctuations correlated between areas. *Neuroimage* 114: 275–286, 2015. doi:10.1016/j.neuroimage.2015.04.026.

Jacobs RH, Baumgartner E, Gegenfurtner KR. The representation of material categories in the brain. *Front Psychol* 5: 146, 2014. doi:10.3389/fpsyg.2014.00146.

Janzen G, Jansen C. A neural wayfinding mechanism adjusts for ambiguous landmark information. *Neuroimage* 52: 364–370, 2010. doi:10.1016/j.neuroimage.2010.03.083.

Janzen G, van Turennout M. Selective neural representation of objects relevant for navigation. *Nat Neurosci* 7: 673–677, 2004. doi:10.1038/nn1257.

Kaas JH. Topographic maps are fundamental to sensory processing. *Brain Res Bull* 44: 107–112, 1997. doi:10.1016/S0361-9230(97)00094-4.

Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8: 679–685, 2005. doi:10.1038/nn1444.

Konkle T, Oliva A. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74: 1114–1124, 2012. doi:10.1016/j.neuron.2012.04.036.

Kornblith S, Cheng X, Ohayon S, Tsao DY. A network for scene processing in the macaque temporal lobe. *Neuron* 79: 766–781, 2013. doi:10.1016/j.neuron.2013.06.015.

Kourtzi Z, Kanwisher N. Representation of perceived object shape by the human lateral occipital complex. *Science* 293: 1506–1509, 2001. doi:10.1126/science.1061133.

Kravitz DJ, Peng CS, Baker CI. Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *J Neurosci* 31: 7322–7333, 2011. doi:10.1523/JNEUROSCI.4588-10.2011.

Kriegeskorte N, Kievit RA. Representational geometry integrating cognition, computation, and the brain. *Trends Cogn Sci* 17: 401–412, 2013. doi:10.1016/j.tics.2013.06.007.

Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci* 2: 4, 2008. doi:10.3389/neuro.06.004.2008.

Lowe MX, Galliva JP, Ferber S, Cant JS. Feature diagnosticity and task context shape activity in human scene-selective cortex. *Neuroimage* 125: 681–692, 2016. doi:10.1016/j.neuroimage.2015.10.089.

Marois R, Yi DJ, Chun MM. The neural fate of consciously perceived and missed events in the attentional blink. *Neuron* 41: 465–472, 2004. doi:10.1016/S0896-6273(04)00012-1.

Morgan LK, Macevoy SP, Aguirre GK, Epstein RA. Distances between real-world locations are represented in the human hippocampus. *J Neurosci* 31: 1238–1245, 2011. doi:10.1523/JNEUROSCI.4667-10.2011.

Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N. A toolbox for representational similarity analysis. *PLOS Comput Biol* 10: e1003553, 2014. doi:10.1371/journal.pcbi.1003553.

O'Craven KM, Downing PE, Kanwisher N. fMRI evidence for objects as the units of attentional selection. *Nature* 401: 584–587, 1999. doi:10.1038/44134.

Oliva A, Torralba A. Building the gist of a scene: the role of global image features in recognition. *Prog Brain Res* 155: 23–36, 2006. doi:10.1016/S0079-6123(06)55002-2.

Park S, Brady TF, Greene MR, Oliva A. 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. doi:10.1523/JNEUROSCI.3885-10.2011.

Park S, Chun MM. Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage* 47: 1747–1756, 2009. doi:10.1016/j.neuroimage.2009.04.058.

Park S, Intraub H, Yi DJ, Widders D, Chun MM. Beyond the edges of a view: boundary extension in human scene-selective visual cortex. *Neuron* 54: 335–342, 2007. doi:10.1016/j.neuron.2007.04.006.

Pelli DG. Pixel independence: measuring spatial interactions on a CRT display. *Spat Vis* 10: 443–446, 1997. doi:10.1163/156856897X00375.

Potter MC. Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn* 2: 509–522, 1976. doi:10.1037/0278-7393.2.5.509.

Schyns PG, Oliva A. From blobs to boundary edges: evidence for time-and spatial-scale-dependent scene recognition. *Psychol Sci* 5: 195–200, 1994. doi:10.1111/j.1467-9280.1994.tb00500.x.

Silver MA, Kastner S. Topographic maps in human frontal and parietal cortex. *Trends Cogn Sci* 13: 488–495, 2009. doi:10.1016/j.tics.2009.08.005.

Skerry AE, Saxe R. Neural representations of emotion are organized around abstract event features. *Curr Biol* 25: 1945–1954, 2015. doi:10.1016/j.cub.2015.06.009.

Takahashi N, Kawamura M. Pure topographical disorientation—the anatomical basis of landmark agnosia. *Cortex* 38: 717–725, 2002. doi:10.1016/S0010-9452(08)70039-X.

Thorpe S, Fize D, Marlot C. Speed of processing in the human visual system. *Nature* 381: 520–522, 1996. doi:10.1038/381520a0.

Troiani V, Stigliani A, Smith ME, Epstein RA. Multiple object properties drive scene-selective regions. *Cereb Cortex* 24: 883–897, 2014. doi:10.1093/cercor/bhs364.

VanRullen R, Thorpe SJ. The time course of visual processing: from early perception to decision-making. *J Cogn Neurosci* 13: 454–461, 2001. doi:10.1162/08989290152001880.

Walther DB, Caddigan E, Fei-Fei L, Beck DM. Natural scene categories revealed in distributed patterns of activity in the human brain. *J Neurosci* 29: 10573–10581, 2009. doi:10.1523/JNEUROSCI.0559-09.2009.

Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L. Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proc Natl Acad Sci USA* 108: 9661–9666, 2011. doi:10.1073/pnas.1015666108.

Ward EJ, Chun MM, Kuhl BA. Repetition suppression and multi-voxel pattern similarity differentially track implicit and explicit visual memory. *J Neurosci* 33: 14749–14757, 2013. doi:10.1523/JNEUROSCI.4889-12.2013.

Wardle SG, Kriegeskorte N, Grootswagers T, Khaligh-Razavi SM, Carlson TA. Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. *Neuroimage* 132: 59–70, 2016. doi:10.1016/j.neuroimage.2016.02.019.

Watson DM, Young WY, Andrews TJ. Spatial properties of objects predict patterns of neural response in the ventral visual pathway. *Neuroimage* 126: 173–183, 2016. doi:10.1016/j.neuroimage.2015.11.043.

Xu Y, Chun MM. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440: 91–95, 2006. doi:10.1038/nature04262.