

Current Biology

Neural Representations Integrate the Current Field of View with the Remembered 360° Panorama in Scene-Selective Cortex

Highlights

- Visual experience of a 360° panorama forges memory associations between scene views
- Representations of discrete views of a 360° environment overlap in RSC and OPA
- The scene currently in view primes associated views of the 360° environment

Authors

Caroline E. Robertson,
Katherine L. Hermann, Anna Mynick,
Dwight J. Kravitz, Nancy Kanwisher

Correspondence

carolinerobertson@fas.harvard.edu

In Brief

How is panoramic visual memory formed in the brain? Robertson et al. report that specific regions of the brain—the RSC and OPA—integrate discrete views of a 360° environment. Subsequently, the scene in the current field of view implicitly triggers associated views of the panoramic environment, facilitating ongoing perception.

Neural Representations Integrate the Current Field of View with the Remembered 360° Panorama in Scene-Selective Cortex

Caroline E. Robertson,^{1,2,5,*} Katherine L. Hermann,² Anna Mynick,³ Dwight J. Kravitz,⁴ and Nancy Kanwisher²

¹Harvard Society of Fellows, Harvard University, Cambridge, MA 02138, USA

²McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02138, USA

³Neuroscience Program, Wellesley College, Wellesley, MA 02481, USA

⁴Department of Psychology, George Washington University, Washington, DC 20052, USA

⁵Lead Contact

*Correspondence: carolinerobertson@fas.harvard.edu

<http://dx.doi.org/10.1016/j.cub.2016.07.002>

SUMMARY

We experience our visual environment as a seamless, immersive panorama. Yet, each view is discrete and fleeting, separated by expansive eye movements and discontinuous views of our spatial surroundings. How are discrete views of a panoramic environment knit together into a broad, unified memory representation? Regions of the brain's "scene network" are well poised to integrate retinal input and memory [1]: they are visually driven [2, 3] but also densely interconnected with memory structures in the medial temporal lobe [4]. Further, these regions harbor memory signals relevant for navigation [5–8] and adapt across overlapping shifts in scene viewpoint [9, 10]. However, it is unknown whether regions of the scene network support visual memory for the panoramic environment outside of the current field of view and, further, how memory for the surrounding environment influences ongoing perception. Here, we demonstrate that specific regions of the scene network—the retrosplenial complex (RSC) and occipital place area (OPA)—unite discrete views of a 360° panoramic environment, both current and out of sight, in a common representational space. Further, individual scene views prime associated representations of the panoramic environment in behavior, facilitating subsequent perceptual judgments. We propose that this dynamic interplay between memory and perception plays an important role in weaving the fabric of continuous visual experience.

RESULTS AND DISCUSSION

How is memory of our broad panoramic environment built from discrete and fleeting views of our spatial surroundings? Here, we sought to understand the psychological and neural mechanisms by which different views of a 360° panoramic environment

are linked in memory. We further asked whether memory for the surrounding environment plays an important functional role in naturalistic scene perception, enabling the scene within the current field of view to prime views of the environment that are currently out of sight.

Participants studied novel 360° panoramic environments: photospheres of real-world locations, which were either dynamically revealed across a panoramic display (experiments 1 and 3) or actively explored using a virtual reality (VR) headset (experiments 2 and 4). These naturalistic, egocentric viewing experiences enabled participants to experience a seamless transition between the poles of each immersive panoramic scene (Figure 1; Movies S1 and S2). Written consent was obtained from all participants in accordance with a protocol approved by the Massachusetts Institute of Technology Institutional Review Board.

Controlled manipulation of these panoramic environments allowed us to ask three questions. First, how do discrete views of a panoramic environment become linked in memory? Second, once memory for a broad panoramic environment is formed, which regions of the brain represent the association between different views from within that environment? Finally, how is memory for the panoramic environment brought to bear on the scene within the current field of view during perception?

To address these questions, we first aimed to determine whether visual experience of a panoramic environment could link discrete views from within that environment in memory ($n = 21$, experiment 1). On each trial of the study phase, participants viewed two movie segments on a panoramic display (Figure 1A). These segments depicted two overlapping or non-overlapping quarters of a broad panoramic environment, dynamically revealed through a restricted sliding window on a computer screen (Figures 1B and 1C; Movie S1). We later tested whether scene views from opposite poles of this environment, 180° degrees apart, became associated in memory as a function of direct visual experience of their mutual panoramic context (overlap condition versus no-overlap condition; Figure 1C). Note that the studied panoramic environments were visually similar to each other (all came from a homogeneous urban neighborhood; Figure S1). Therefore, associations between test image pairs in this stimulus set could only be inferred from their remembered shared panoramic context (overlap versus no-overlap condition; Figure 1C).

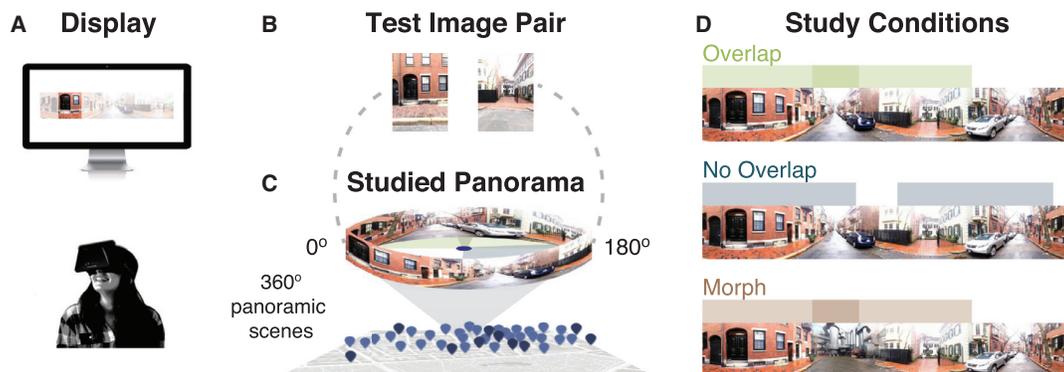


Figure 1. Experimental Paradigm and Stimuli

(A) Participants studied novel, 360° panoramic scenes, either dynamically revealed through a restricted, sliding window on a panoramic display (experiments 1 and 3, top) or actively explored using a virtual reality headset (experiments 2 and 4, bottom).

(B and C) During the memory tests, we presented 20 pairs of images, drawn from opposite poles of 360° panoramic scenes.

(D) For half of these pairs, the panoramic visual information linking them was known because participants had extensively studied overlapping quarters of the scene, for example, the 0–100° and 80°–180° quadrants of the 360° panorama (Overlap study condition, top panel, green). For the other half, the panoramic visual information linking them was unknown, either because participants had studied two non-overlapping quarters of the scene, for example, the 0–100° and 170°–270° quadrants of the 360° panorama (No-Overlap study condition, experiments 1 and 3, middle panel, blue), or because participants had studied spliced panoramas that smoothly morphed between two panoramic scenes (Morph study condition, experiment 2, bottom panel, brown).

See [Movies S1](#) and [S2](#) for screen videos of study trials, [Figure S1](#) for examples of all test images, and [Supplemental Experimental Procedures](#) for further experimental details.

On each trial of a subsequent memory test, participants were shown two images, drawn from opposite poles of the studied panoramas (Figure 2A), and asked to report whether the two images came from the same panoramic scene. Participants successfully formed memory associations between two images from the same scene as a function of panoramic visual experience: explicit memory (d') for the association between two images was significant for both the overlap and no-overlap conditions (overlap: $t(18) = 4.19$, Cohen's $d = 0.96$, $p = 0.001$; no-overlap: $t(18) = 2.43$, Cohen's $d = 0.56$, $p = 0.026$; Figure 2B). Importantly, though, memory performance was significantly stronger for the overlap compared with the no-overlap condition (overlap mean and SEM: 0.63 ± 0.13 , no-overlap: 0.25 ± 0.08 ; repeated-measures ANOVA $F(1,18) = 4.09$, $\eta_p^2 = 0.30$, $p = 0.001$; Figure 2A). These results indicate that direct visual experience of a broad panoramic environment serves to associate discrete views from within that environment in memory.

But which components of panoramic visual experience drive associative learning in natural scene perception? Thus far, our results argue that continuous visual experience through common visual content causes a stronger association between scene views than mere temporal co-occurrence on study trials, which was equivalent in the overlap and no-overlap conditions. This finding distinguishes panoramic visual learning from demonstrations of paired associative learning based on temporal co-occurrence, for example, between views of objects or faces [11] and their concomitant cortical changes [12, 13]. Next, we asked whether opposite poles of a panoramic expanse could be associated with each other via mutual association with common, but not panoramically consistent, visual content (i.e., via transitive learning). We also tested whether our finding from experiment 1—that continuous visual experience links scene views in memory—would extend to fully egocentric, active viewing conditions using a VR headset.

To test these hypotheses, we ran a control experiment ($n = 18$, experiment 2; [Supplemental Experimental Procedures](#)) in which participants actively explored novel panoramic environments using a VR headset (Figure 1A). In this experiment, half of the panoramas were continuously experienced as in experiment 1 (overlap condition; Figure 1D; [Movie S2](#)), and half were spliced in the middle with foreign panoramas (morph condition; Figure 1D; [Movie S2](#)). In theory, the two images on opposite poles of these panoramas could be associated via their mutual association with a common middle image, although this spliced image was from an obviously different place (e.g., a plaza or a factory setting) (Figure 1C; [Movie S2](#)).

However, continuous panoramic experience (overlap condition) led to significantly stronger associations between scene views than transitive experience (morph condition). Participants' sensitivity (d') to detect that two images came from the same panoramic environment was significant for both the overlap and morph conditions (overlap: $t(19) = 3.27$, Cohen's $d = 0.73$, $p = 0.004$; no-overlap: $t(19) = 3.74$, Cohen's $d = 0.84$, $p = 0.001$; Figure 2B). Critically, associative memory was significantly higher for the overlap compared with the morph condition (overlap mean d' and SEM: 1.61 ± 0.50 , morph: 0.63 ± 0.17 , repeated-measures ANOVA $F(1,19) = 2.218$, $\eta_p^2 = 0.21$, $p = 0.01$; Figure 2B). All in all, these results demonstrate that continuous, panoramic visual experience links individual scene views from within that panorama in memory under both passive (experiment 1) and active (experiment 2) viewing conditions, going beyond either associative or transitive learning. These findings lend support to models of memory integration that posit that graded exposure to continuous, morphed intermediate states between two distinct stimuli supports the merging of these representations in memory [14–16]. Future work should explore the benefits of active over passive viewing conditions on scene memory formation.

A Same or different place?



B Associative Memory Experiments 1 and 2

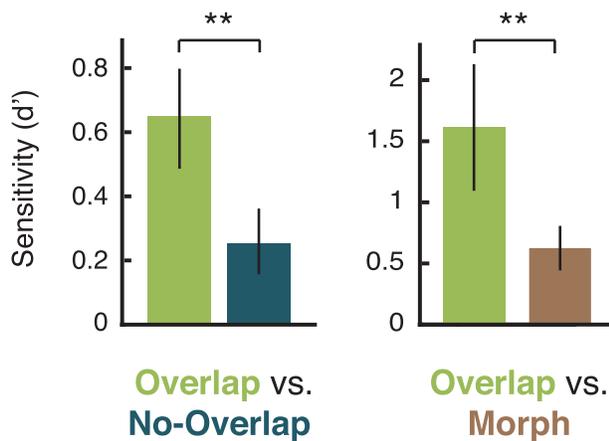


Figure 2. Panoramic Visual Experience Forges Associations between Scene Views in Behavior

(A) In the memory tests of both experiments 1 and 2, participants were asked whether two scene views, drawn from opposite poles of a 360° panoramic environment, depicted the same place or two different places.

(B) In both experiments, sensitivity (d') to detect that two images came from the same panoramic environment was higher for the Overlap condition (light green) compared with control conditions (Overlap vs. No-Overlap: $F(1,18) = 4.09$, $\eta_p^2 = 0.30$, $p = 0.001$; Overlap vs. Morph: $F(1,19) = 2.218$, $\eta_p^2 = 0.21$, $p = 0.01$). These results demonstrate that discrete views of a panoramic environment become associated in memory as a result of direct visual experience of the panoramic content that unites them (Overlap condition). Further, these effects go beyond simple temporal co-occurrence between scene views on study trials (No-Overlap condition) or transitive learning between views with intermediate visual content (Morph condition).

In all plots, error bars represent 1 SEM. ** $p \leq 0.01$, *** $p \leq 0.001$ difference between the two conditions. See also [Movies S1](#) and [S2](#) for screen videos of study trials and [Figure S2](#) for replication of experiment 1.

Having demonstrated the conditions under which two discrete views from opposite poles of a 360° panoramic environment are integrated into a broad structural representation of a scene in memory, distinct from either associative or transitive representations, we aimed to determine the neural substrate of this visual memory for views of a panoramic environment. We specifically predicted that this effect would manifest in certain regions of the scene network of the brain—the parahippocampal place

area (PPA), the retrosplenial complex (RSC), and the occipital place area (OPA)—which are visually selective for scenes [2, 3] but also sensitive to memory for navigationally relevant information, such as landmark familiarity (PPA) [5, 6] and known position on an allocentric map (RSC) [7, 8] and hence serve as candidate loci for integrated representations between retinal input and memory.

After studying the panoramas, participants were scanned in an event-related fMRI paradigm ($n = 12$, experiment 3; [Supplemental Experimental Procedures](#); see also [Figure S2](#) for behavioral replication of experiment 1). On each trial, participants viewed a discrete scene view from the studied panoramas and indicated whether it had appeared on the left or the right side of the studied panorama. Crucially, this question did not require explicit recall of the associated scene view (the other pole of the 360° panoramic environment).

Using multivariate pattern analysis, we tested whether representational similarity was stronger for pairs of scene views whose shared panoramic context was known (overlap condition) compared to unknown (no-overlap condition), demonstrating representational overlap between discrete scene views that were associated in memory ([Supplemental Experimental Procedures](#)). We first computed panorama decoding indices for each region of interest (ROI): the average correlation in the pattern of fMRI response between two different views from the same panorama minus the correlation between two views from different panoramas. Panorama decoding indices significantly greater than zero indicate greater similarity between the response patterns evoked by two scene views that shared panoramic context than two scenes that did not.

A two-way ANOVA on these panorama decoding indices, with ROI (PPA, RSC, and OPA) and memory condition (overlap versus no-overlap) as repeated-measures factors, revealed a significant main effect of overlap versus no-overlap ($F(1,11) = 6.8687$, $\eta_p^2 = 0.38$, $p = 0.026$), which was qualified by an interaction with ROI ($F(2,22) = 3.66$, $\eta_p^2 = 0.41$, $p = 0.038$). We therefore investigated each region of the scene-responsive network separately in a series of post hoc comparisons.

Our results demonstrate a robust effect of memory for the panoramic environment on the visual responses of two regions of the scene network: RSC and OPA ([Figure 3](#)). Specifically, these regions demonstrated a stronger similarity for pairs of images from the overlap compared with the no-overlap condition: OPA ($F(1,11) = 2.07$, $\eta_p^2 = 0.28$, $p = 0.003$) and RSC ($F(1,11) = 2.65$, $\eta_p^2 = 0.39$, $p = 0.001$). These results indicate that the RSC and OPA represent scene views in the context of visual memory for the broader panoramic environment outside of the current field of view.

These effects were specific to RSC and OPA: responses in PPA were not modulated by memory for the broader panoramic scene ([Figure 3](#)). In PPA, responses between pairs of scene views were equally similar for pairs whose panoramic context was known (overlap condition) as opposed to unknown (no-overlap condition) ($F(1,11) = 0.048$, $\eta_p^2 = 0.0$, $p = 0.928$). The absence of these effects in PPA is consistent with previous literature: the PPA is particularly implicated in visual recognition of landmarks [5, 17]—objects or buildings of navigational relevance that typically appear in the distance and rarely span multiple fields of view [1]. Further, control analyses revealed that face-selective

fMRI, Experiment 3

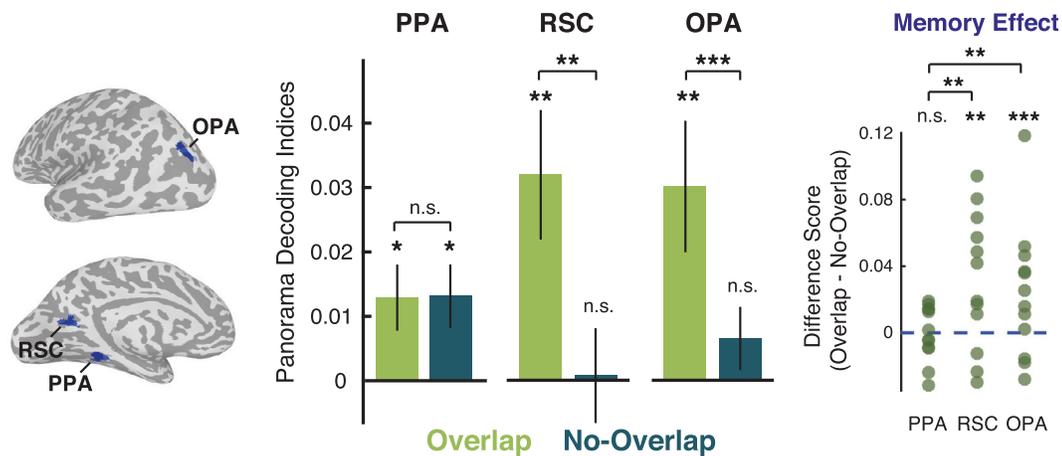


Figure 3. Memory for the Panoramic Environment Forges Representational Similarity between Scene Views in RSC and OPA

Two specific regions of the scene network evidenced stronger similarity for pairs of images from the Overlap as compared with the No-Overlap condition: RSC ($F(1,11) = 2.65$, $\eta_p^2 = 0.39$, $p = 0.001$) and OPA ($F(1,11) = 2.07$, $\eta_p^2 = 0.28$, $p = 0.003$). Conversely, responses in PPA were not modulated by memory for the broader panoramic scene ($F(1,11) = -0.048$, $p = 0.928$). Individual difference scores for each individual subject are plotted in the far right panel. These results demonstrate that discrete views of a panoramic environment increase in representational similarity in two specific regions of the scene network, the RSC and OPA, as a function of visual memory for the spatial information that unites them. In all plots, error bars represent 1 SEM. ** $p \leq 0.01$, *** $p \leq 0.001$ difference between the two conditions. See also [Movie S1](#) for screen videos of study trials and [Figure S2](#) for behavioral data.

(FFA) and early visual (V1) regions of the brain were not modulated by memory condition (FFA: $F(1,11) = 0.067$, $\eta_p^2 = 0$, $p = 0.899$; V1: $F(1,11) = -1.085$, $\eta_p^2 = 0.06$, $p = 0.20$).

Finally, it should be noted that all scene ROIs displayed stereotypical signatures of scene-selective visual regions [18–20]: sensitivity to the identity of each scene view across repetitions (all $t > 3.80$; $p = 0.001$) as well as the spatial layout within each scene view (open versus closed, PPA: $t = 2.20$, $p = 0.049$; OPA: $t = 5.22$, $p = 0.007$; RSC: $t = 3.49$, $p = 0.023$). Importantly, neither of these signatures (image discrimination or spatial layout discrimination) interacted with our main condition of interest, memory condition (both $p > 0.340$).

These results demonstrate neural representations of the scene within the current field of view are imbued with our memory for the broader panoramic environment. We hypothesized that this association might serve an important functional role in scene perception, causing non-overlapping views of a learned panoramic environment to automatically prime each other in perception. We tested this hypothesis in a final behavioral experiment ($n = 20$, experiment 4; [Supplemental Experimental Procedures](#)). On each trial of experiment 4, participants were shown an image from one of the studied panoramas and asked to remember the position of the image (whether it had been shown on the left or right side of the panoramic scene at study). Critically, the image was briefly primed (300 ms) by a task-irrelevant stimulus: either another image from the same panorama (valid prime) or a black square (neutral prime) ([Figures 4A and 4B](#)).

Overall, memory accuracy for where a scene view had appeared in the broader panorama (left or right side) was high for both the panoramic and morph conditions (overlap mean and SEM: 0.81 ± 0.03 , morph: 0.81 ± 0.02 , $F(1,18) = 0.03$, $\eta_p^2 = 0.0$, $p = 0.9$), indicating that location memory for discrete scene

views was similar between the conditions. Critically, however, the perceptual priming effect (efficiency valid prime – efficiency neutral prime) was significantly stronger in the overlap condition compared to the morph condition (mean and SEM, overlap: 0.046 ± 0.016 ; morph: -0.003 ± 0.02 ; $F(1,18) = 1.43$, $\eta_p^2 = 0.27$, $p = 0.02$; [Figure 4C](#)). Post hoc comparisons revealed a significant effect of perceptual priming in the overlap condition ($F(1,18) = 2.15$, $\eta_p^2 = 0.33$, $p = 0.008$), but not in the morph condition ($F(1,18) = 0.14$, $\eta_p^2 = 0.0$, $p = 0.82$). An additional control study demonstrated that the perceptual priming we observed in the overlap condition was specific to primes that share panoramic context with the target, as compared with invalid primes from another panorama (see [Supplemental Information](#)). These results demonstrate that memory for a broad, panoramic environment causes discrete views from within this environment to prime each other in subsequent perception, facilitating perceptual judgments.

Our findings dovetail with predictions from integrative encoding models of memory, which posit that memory representations of prior related events are reactivated during encoding of novel events, contextualizing ongoing experience [21–23]. Neural support for such models derives from associative inference paradigms, where discrete stimuli (e.g., a baseball and a hat) are paired via mutual association with a third stimulus (e.g., a car) [24, 25]. Here, we present a concrete example of integrative memory encoding in naturalistic visual experience: the scene within the current field of view triggers memory of the broader panoramic environment, priming perception of associated views.

Much of the content of a visual percept is based on inference and memory that goes far beyond the information available from retinal input [26]. Here, we demonstrate that the percept of the current field of view is imbued with memory of the broader

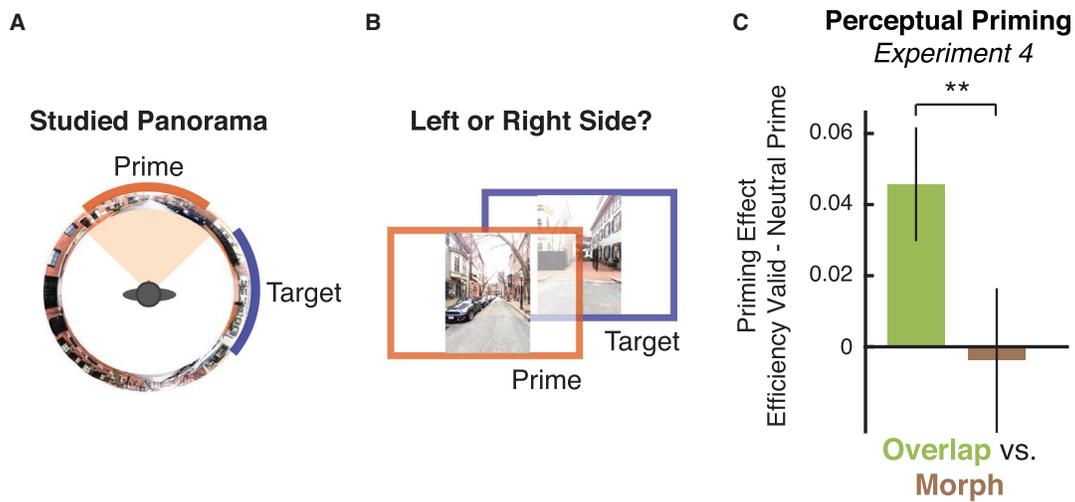


Figure 4. Associated Scene Views Prime Each Other in Perception

(A–C) Participants were faster and more accurate to remember the spatiotopic location of a scene view (Target, A) if it was briefly primed by another view from the broader panoramic environment (Prime, B). Performance was measured as efficiency scores (accuracy / reaction time). The perceptual priming effect (efficiency valid prime – efficiency neutral prime) was significantly stronger in the Overlap condition compared with the no-Overlap condition ($F(1, 18) = 1.43$, $\eta_p^2 = 0.27$, $p = 0.02$). These results demonstrate that the scene within the current field of view implicitly triggers memory for associated views of the panoramic environment. In all plots, error bars represent 1 SEM. ** $p \leq 0.01$, *** $p \leq 0.001$ difference between the two conditions. See also [Movie S2](#) for screen videos of study trials.

panoramic environment. In this way, ongoing scene representations are affixed to a broader representation of the surrounding environment, which may help to support our sense of a seamless panoramic visual expanse.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two figures, one table, and two movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2016.07.002>.

AUTHOR CONTRIBUTIONS

Conceptualization, C.E.R., N.K., and D.J.K.; Methodology, C.E.R., K.L.H., N.K., and D.J.K.; Investigation, C.E.R., A.M., and K.L.H.; Formal Analysis, C.E.R. and D.J.K.; Visualization, C.E.R. and A.M.; Writing – Original Draft, C.E.R.; Writing – Review & Editing, C.E.R., N.K., D.J.K., K.L.H., and A.M.

ACKNOWLEDGMENTS

We thank S. Chang for help with stimulus creation; C. Baker, M. Cohen, S. Gershman, and L. Isik for comments on the manuscript; and R. Epstein for illuminating discussion. This research was funded by grants awarded to N.K. by the NSF Science and Technology Center for Brains, Minds, and Machines (CCF-1231216) and to C.E.R. from the Harvard Milton Fund.

Received: June 3, 2016
Revised: June 24, 2016
Accepted: July 1, 2016
Published: September 8, 2016

REFERENCES

- Epstein, R.A., and Vass, L.K. (2013). Neural systems for landmark-based wayfinding in humans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20120533.
- Epstein, R., and Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature* 392, 598–601.
- Dilks, D.D., Julian, J.B., Paunov, A.M., and Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* 33, 1331–6a.
- Kravitz, D.J., Saleem, K.S., Baker, C.I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230.
- Marchette, S.A., Vass, L.K., Ryan, J., and Epstein, R.A. (2015). Outside looking in: landmark generalization in the human navigational system. *J. Neurosci.* 35, 14896–14908.
- Janzen, G., and van Turenout, M. (2004). Selective neural representation of objects relevant for navigation. *Nat. Neurosci.* 7, 673–677.
- Vass, L.K., and Epstein, R.A. (2013). Abstract representations of location and facing direction in the human brain. *J. Neurosci.* 33, 6133–6142.
- Marchette, S.A., Vass, L.K., Ryan, J., and Epstein, R.A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nat. Neurosci.* 17, 1598–1606.
- Park, S., and Chun, M.M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *Neuroimage* 47, 1747–1756.
- Golomb, J.D., Albrecht, A.R., Park, S., and Chun, M.M. (2011). Eye movements help link different views in scene-selective cortex. *Cereb. Cortex* 21, 2094–2102.
- Eichenbaum, H., and Bunsey, M. (1995). On the binding of associations in memory: clues from studies on the role of the hippocampal region in paired-associate learning. *Curr. Dir. Psychol. Sci.* 4, 19–23.
- Sakai, K., and Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature* 354, 152–155.
- Schlack, A., and Albright, T.D. (2007). Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. *Neuron* 53, 881–890.
- Wallis, G., and Bühlhoff, H.H. (2001). Effects of temporal association on recognition memory. *Proc. Natl. Acad. Sci. USA* 98, 4800–4804.
- Blumenfeld, B., Preminger, S., Sagi, D., and Tsodyks, M. (2006). Dynamics of memory representations in networks with novelty-facilitated synaptic plasticity. *Neuron* 52, 383–394.

16. Preminger, S., Blumenfeld, B., Sagi, D., and Tsodyks, M. (2009). Mapping dynamic memories of gradually changing objects. *Proc. Natl. Acad. Sci. USA* *106*, 5371–5376.
17. Aguirre, G.K., and D'Esposito, M. (1999). Topographical disorientation: a synthesis and taxonomy. *Brain* *122*, 1613–1628.
18. Epstein, R., Graham, K.S., and Downing, P.E. (2003). Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* *37*, 865–876.
19. Kravitz, D.J., Peng, C.S., and Baker, C.I. (2011). Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *J. Neurosci.* *31*, 7322–7333.
20. Park, S., Brady, T.F., Greene, M.R., and Oliva, A. (2011). Disentangling scene content from spatial boundary: complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *J. Neurosci.* *31*, 1333–1340.
21. Holland, P.C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learn. Motiv.* *12*, 1–18.
22. Hall, G. (1996). Learning about associatively activated stimulus representations: implications for acquired equivalence and perceptual learning. *Anim. Learn. Behav.* *24*, 233–255.
23. Schlichting, M.L., and Preston, A.R. (2015). Memory integration: neural mechanisms and implications for behavior. *Curr Opin Behav Sci* *7*, 1–8.
24. Zeithamova, D., Dominick, A.L., and Preston, A.R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* *75*, 168–179.
25. Horner, A.J., Bisby, J.A., Bush, D., Lin, W.-J., and Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nat. Commun.* *6*, 7462.
26. Bruner, J.S. (1973). *Beyond the Information Given: Studies in the Psychology of Knowing* (Oxford).

Current Biology, Volume 26

Supplemental Information

Neural Representations Integrate the Current

Field of View with the Remembered 360° Panorama

in Scene-Selective Cortex

Caroline E. Robertson, Katherine L. Hermann, Anna Mynick, Dwight J. Kravitz, and Nancy Kanwisher

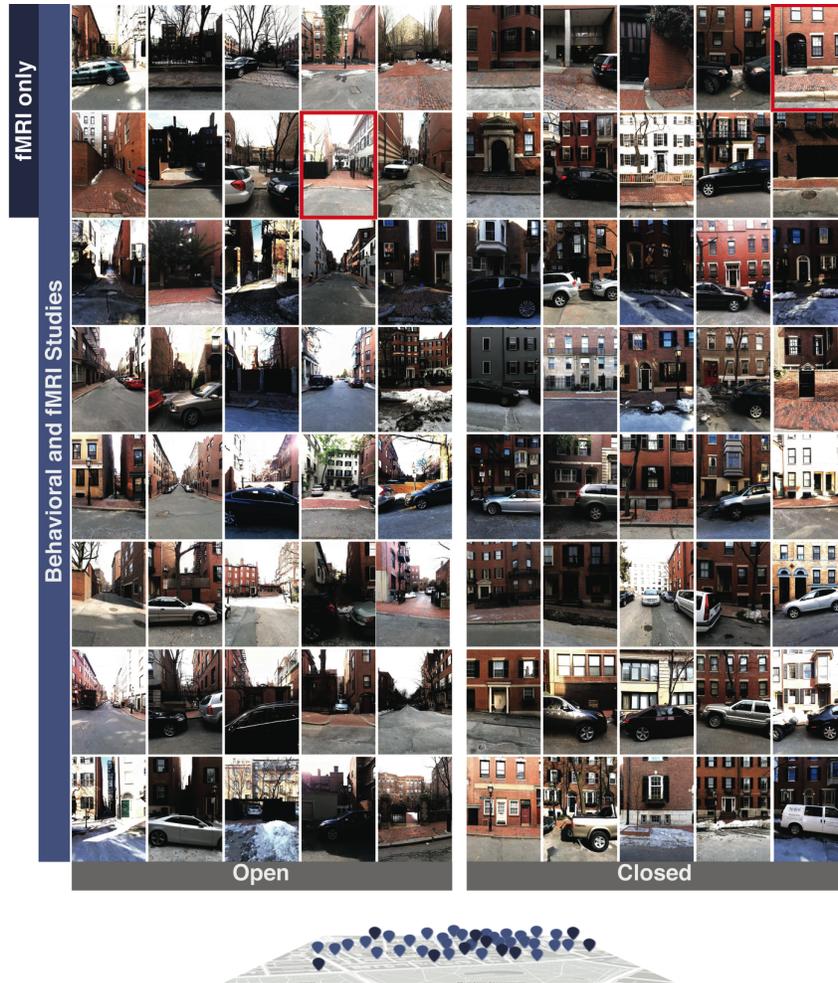


Figure S1. Related to Figure 1. Example scene views used in the memory tests. Each pair of test images (example pair circled in red) came from opposite ends of a studied panoramic environment, 180° degrees apart. Each panoramic scene was shot in the middle of a street, at one of 40 locations in the Beacon Hill neighborhood of Boston, Massachusetts (chosen based on its visual homogeneity: colonial-style row houses with narrow streets and redbrick sidewalks). At each of these locations, opposite sides of the street had different spatial layouts: one scene view was open (e.g. a courtyard) and the other was closed (e.g. the face of an apartment building). As illustrated in the figure, the studied panoramic environments were all visually similar to each other. Therefore, associations between test image pairs could only be inferred from their remembered panoramic context.

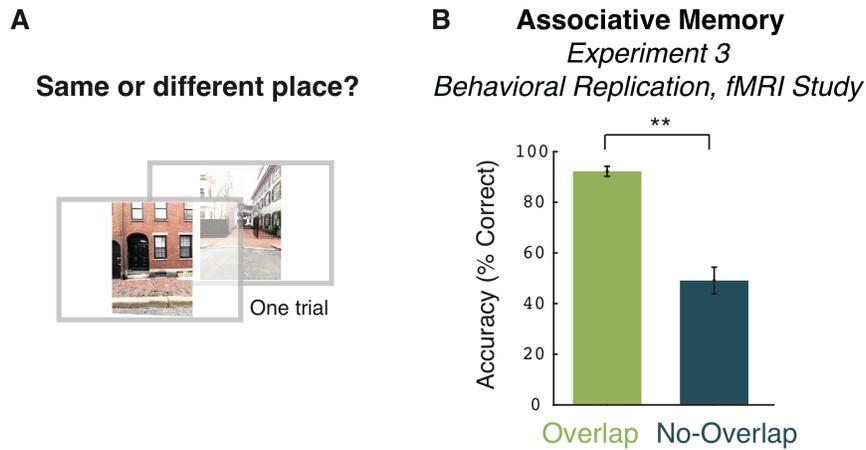


Figure S2. Related to Figure 3. Panoramic Visual Experience Forges Associations Between Scene Views in Behavior, Replication in fMRI Study Participants. (A) In the memory test, participants were asked whether two scene-views, drawn from opposite poles of a 360° panoramic environment, depicted the same place or two different places. (B) Sensitivity (d') to detect that two images came from the same panoramic environment was higher for the Overlap Condition (light green) compared with the No Overlap Condition (Overlap/No-Overlap, $P = 0.01$). In all plots, error bars represent 1 SEM. ** $P \leq 0.01$, *** $P \leq 0.001$ difference between the two conditions.

Supplemental Experimental Procedures

Participants

61 adults participated in our study: 21 (Experiment 1), 20 (Experiments 2 and 4), and 12 (Experiment 3). Participants were recruited based on two criteria: 1) never having spent time in the Beacon Hill neighborhood of Boston, Massachusetts (where our stimuli were filmed), and 2) having normal or corrected-to-normal vision. Participants were all adults, mean age and STD: 25.43 \pm 8.96 years (Experiment 1); 24.7 \pm 5.82 years (Experiments 2 and 4); 24.75 \pm 7.3 years, (Experiment 3). Written consent was obtained from all participants in accordance with a protocol approved by the Massachusetts Institute of Technology Institutional Review Board.

Stimuli

In the Study Phase of all experiments, participants studied novel 360° panoramic scenes of Beacon Hill. This neighborhood was chosen based on its visual homogeneity: colonial-style row houses with narrow streets and redbrick sidewalks. Each scene was shot in the middle of a street, at one of 40 locations in Beacon Hill (Figure 1B). At each of these locations, opposite sides of the street had different spatial layouts: one scene view was open (e.g. a courtyard) and the other was closed (e.g. the face of an apartment building). These two views, henceforth referred to as “opposite poles” of the 360° panorama, were used as the test image pairs in the memory tests. Each scene was shot as a “photosphere” using the image processing software Bubbli (www.bubb.li).

Stimulus Creation and Display: Experiments 1 and 3

Each photosphere was later processed into a movie that smoothly panned 360° around the horizontal equator of the sphere using MATLAB, revealing a restricted view of the scene on each frame (~60°). These movies were split into four segments, each containing 100° of the 360° visual scene. This process allowed us to precisely control the speed at which each degree of the visual scene was displayed on the panoramic screen. Participants viewed a widescreen ASUS LED monitor (diagonal: 60.96 cm; resolution: 1920 x 1080; refresh rate: 120 Hz) from a distance of approximately 60 cm. Stimuli were displayed using Psychtoolbox and responses were collected through button press.

Stimulus Creation and Display: Experiments 2 and 4

Each photosphere was later applied to a Virtual Reality (VR) environment built in Unity (www.unity3d.com), which integrated with a Virtual Reality Headset (Oculus Rift, Development Kit 2, www.oculus.com). The Oculus Rift display (Low persistence OLED screen, resolution: 960 x 1080; field-of-view: ~100 degrees; 75 Hz refresh) allowed participants to actively explore the panoramic environment through turns in head direction, providing a self-directed opportunity to explore the naturalistic panoramic environment from an egocentric perspective.

Table S1

	Behavior / fMRI	Study Display	Panorama Study Conditions	Test on Scene View Pairs
Experiment 1	Behavior	Computer	Overlap vs. No-Overlap	Associative memory
Experiment 2	Behavior	VR Headset	Overlap vs. Morph	Associative memory
Experiment 3	fMRI	Computer	Overlap vs. No-Overlap	Representational Overlap
Experiment 4	Behavior	VR Headset	Overlap vs. Morph	Perceptual Priming

Table S1, Overview of Experimental Design. Our study consisted of four experiments. *Experiments 1-2* determined the conditions under which a scene views are associated into a panoramic visual memory. *Experiment 3* aimed to determine the neural substrate of panoramic visual memory. *Experiment 4* aimed to determine whether panoramic memory enabled perceptual priming between scene views in an environment. Please see below for full details of the designs and analyses of *Experiments 1 – 4*.

Procedure: Experiment 1

Experiment 1 had two phases: Study Phase (50 mins) and Associative Memory Test (10 mins). On each trial of the Study Phase (6 trials / panorama), participants studied two movie segments that each captured 100° of the 360° of a panoramic scene in Beacon Hill. As depicted in Figure 1A and Movie S1, each movie was displayed through a moving aperture (8 visual degrees) that slid between the edge and center of the screen or vice versa (direction counterbalanced across trials) (Figure 1A; Movies S1). Two movie segments were presented sequentially on each trial, each revealing approximately one quarter (100°) of the panoramic scene, one on each side of the screen (left/right). As such, 180° panoramic scene was revealed on each trial in two fixed parts (each 100/180 degrees) (Movie S1). The directions from which these movie segments panned across the panoramic scene (e.g. 0° to 100° or 100° to 0°) were counterbalanced across trials. On each trial of the Study Phase, participants were instructed to judge whether or not the two movies were drawn from the same location in Beacon Hill. Participants were instructed to indicate their responses through button press and were given instruction and practice with the task before the Study Phase.

Three types of trials occurred during the Study Phase (Figure 1D), either: 1) the two movies were drawn from the same location and overlapped in 25% of their content (Overlap); 2) the two movies were drawn from the same location but did not share any content (No-Overlap); 3) the two movies were drawn from different locations (Different Location). The assignment of movies to conditions was counterbalanced across participants in each experiment.

Before the Study Phase began participants were given practice with the task. Specifically, participants were shown demo trials where the two movies came from unique places around the world (Easy Practice Run) or another neighborhood in Massachusetts (Difficult Practice Run), and asked to indicate whether two movies were drawn from the same location. They were given feedback, explicitly informed of the three potential trial types, and instructed to study the novel scenes carefully in preparation for two tests that would "test their memory of the panoramic scenes".

The Memory Test Phase tested whether discrete images from opposite edges of these panning movies become associated in memory. On each trial of a memory test, participants were shown two images (presented sequentially: 2s display time, 0.5s ISI), taken from opposite poles of the studied panoramas (Figure 1B), and asked to report whether the two images came from the same panoramic scene (*Associative Memory Test*).

To test whether associative memory was stronger for the Overlap compared with the No-Overlap Condition, we calculated d' for these conditions separately, with reference to false alarms and correct rejections derived from the Different Location Condition. For example: $[Hits (Overlap Trials) / \# Overlap Trials] - [False Alarms (Different Location Trials) / Correct Rejections (Different Location Trials)]$.

Procedure: Experiments 2 and 4

In a separate study, a new set of participants ($N = 20$) first took part in a Study Phase (50 mins) and two memory tests (henceforth *Experiment 2* and *Experiment 4*). On each trial of the Study Phase (3 trials / panorama), participants studied 360° panoramic scenes from Beacon Hill through a head-mounted virtual reality display (Oculus Rift DK2). In all scenes, an occluding wall was placed behind the participant so that only the 180° from the participant's right to left was visible (Movie S2). This procedure ensured that participants only viewed 180° of the panoramic scene, as in *Experiments 1 and 3*.

At the start of each trial, participants were visually immersed in a "cloud world", standing on top of a compass rose. To begin each trial, participants were required to face North, and then press a button on a remote control. Subsequently, a panoramic scene would appear around the participant. This procedure ensured that participants only ever viewed a polar scene-view on the left or right side of their body. This scene would remain on the screen for 20 seconds, after which the scene would time-out and proceed to the next trial (Movie S2).

Two types of trials occurred during the Study Phase (Figure 1D; Movie S2). Half of these panoramas were typical scenes (Overlap Study Condition). The other half were spliced panoramas that smoothly morphed between two panoramic scenes (Morph Study Condition). These spliced panoramas had an *ABA* structure: left side of street in Beacon Hill, middle of a different panoramic scene, right side of the street in Beacon Hill. In theory, the two images on opposite poles of the Morph Condition panoramas could be associated via their mutual association with a common middle image, although this spliced image was from an obviously different place (e.g. a plaza or a factory setting) (Movie S2). On each trial of the Study Phase, participants were explicitly informed of the two potential trial types, and instructed to study the novel scenes carefully in preparation for two tests that would "test their memory of the panoramic scenes".

Participants subsequently performed two Memory Tests (*Experiment 2* and *Experiment 4*). *Experiment 2* probed participants' explicit association memory for scene-views drawn from opposite edges of the panoramic scene. On each trial, participants judged whether two discrete views of a panoramic scene (presented sequentially) depicted the same location. *Experiment 4* tested whether two images from a learned panorama primed each other in perception. On each trial, participants were shown a discrete still frame from one of the panoramas, and were asked to report the spatiotopic position of the image (whether it was shown on the left/right side of the panoramic scene during the Study Phase). Critically, each discrete

view was preceded by a brief (300 ms) presentation of either: 1) a still from the same panorama from the study phase (Valid Prime) or 2) a black square (Neutral Prime). See *Perceptual Priming Control Study* below for comparison between a Valid and Invalid Prime.

To test whether associative memory was stronger for the Overlap compared with the Morph Condition (*Experiment 2*), we calculated d' for these conditions separately. To test whether perceptual priming was stronger for the Overlap compared with the Morph Condition, we calculated Efficiency Scores (mean reaction time / accuracy) for the Valid and Neutral trials of each Condition. We next calculated the Perceptual Priming Effect for each Condition (Efficiency Valid Prime – Efficiency Neutral Prime). In calculating mean RTs for each Condition, only correct trials were included.

Procedure: Experiment 3 (Behavioral)

12 new participants took part in the functional Magnetic Resonance Imaging (fMRI) study. Participants first took part in the Learning and Memory Test Phases *Experiment 1* (day 1), and then returned for an fMRI study (day 2). Before the scan, participants repeated half of the Study Phase to refresh their memories of the studied panoramas. All procedures were identical to those of *Experiment 1*, except that that the number of sites was reduced to 10 (5 Overlap and 5 No-Overlap).

Procedure: Experiment 3 (fMRI Study)

Magnetic Resonance Imaging (MRI) data were collected using a Siemens Trio 3T MRI scanner, equipped with a 32-channel head coil and located at the Massachusetts Institute of Technology McGovern Institute for Brain Research. We first acquired a high-resolution, whole-brain anatomical volume using a T1-weighted magnetization-prepared rapid acquisition gradient echo imaging sequence (MPRAGE) (124 slices; voxel size = 1 x 1 x 1 mm; repetition time = 2530 ms; echo time = 2.94 ms; field of view = 256 x 256 x 176 mm). Next, we acquired 8 runs of functional magnetic resonance imaging data for the main (event-related) experiment (25 slices, spanning the ventral most part of the temporal lobe through most of parietal; voxel size = 2 x 2 x 2 mm; repetition time = 2000 ms; echo time = 30; field of view = 25 x 96 x 96 mm; duration: 133 TRs), as well as two runs of a functional localizer scan with the same scan parameters (duration: 202 TRs).

On each trial of the event-related experiment, a single image was shown at the center of the screen at fixation (500 ms presentation; ISI: 4 – 6 s; width: 16 degrees visual angle; height: 11.5 degrees visual angle). Participants were asked to recall whether this image had been shown on the left or the right side of the screen during the study phase outside of the scanner and to respond using button press. Every image occurred three times in each run, with the presentation order randomized between and across runs, allowing 24 repetitions of each image throughout the experiment. A fixation cross was shown throughout the experiment, and participants were instructed to maintain fixation throughout.

During the block-design functional localizer scan, participants were shown images of faces, scenes, and objects to isolate scene and face-selective regions of visual cortex (image presentation duration: 1s; block duration: 16s; total number of blocks per condition: 16). To ensure engagement with the localizer stimuli, participants performed a simple 1-back task, pressing a button whenever an image was shown twice back-to-back.

MRI Analysis

Data were preprocessed using Analysis of Functional NeuroImages software (AFNI) (<http://afni.nimh.nih.gov/afni>) [S1]. Data from the first five repetition times of each event-related run were discarded to minimize the effect of transient magnetic saturation. All functional volumes were slice-time and motion corrected, and spatially smoothed using a 5 mm full-width at half-maximum Gaussian kernel. Motion parameters were modeled as regressors of no interest in the design matrices of all deconvolutions to model out potential head motion artifacts.

High-resolution anatomical scans were aligned to each participant's functional data, and all

analyses were done in each participant's native space. Cortical surfaces were generated from the high-resolution, skull-stripped anatomical volume using FreeSurfer [S2]. Anatomical regions of interest (ROIs) of the hippocampus and V1 were reconstructed for each participant using gyral and sulcal anatomy. Anatomical ROIs were aligned to the volumetric functional data using SUMA (<http://afni.nimh.nih.gov/afni/suma>).

Identification of FFA, PPA, OPA and RSC

We computed significance maps of the brain for each individual's localizer scan, thresholded at $P < 10^{-4}$ (uncorrected). PPA, OPA, and RSC were defined bilaterally as a contiguous cluster of suprathreshold voxels evidenced in the functional localizer contrast Scenes > Faces, and PPA was defined bilaterally using the contrast Faces > Objects. All ROIs were visually inspected for anatomical location [S3]. Each region of interest was further refined to ensure that all ROIs were mutually exclusive, using the following criteria. First, voxels within a cluster were required to lie within a 10 mm radius of the voxel demonstrating peak selectivity in the cluster corresponding to the ROI being defined. Second, any voxels common to multiple regions of interest were deleted. Results were comparable between the left and right hemispheres in all functionally-localized areas. We therefore collapsed across hemispheres for all reported analyses.

Identification of primary visual cortex

V1 was bilaterally defined on each participant's cortical surface, which was reconstructed from the high-resolution anatomical volume using FreeSurfer. For each participant, automated anatomical criteria were employed to reconstruct a probabilistic map of the location of V1 along the calcarine sulcus using the gyral and sulcal anatomy [S4]. Results were comparable between the left and right hemispheres in V1; we therefore collapsed across hemispheres for all reported analyses.

fMRI Analysis

We employed a standard generalized linear model (GLM) to estimate the haemodynamic response function to each stimulus in each voxel using AFNI. Our sparse event-related design allowed us to assess the response to each image in each voxel (20 stimuli, 24 repetitions/image). We employed two types of deconvolutions in our analysis: i) a standard GLM in which responses were convolved with a gamma function and averaged across repeated trials, resulting in 20 conditions; ii) a GLM in which responses were estimated for each trial, resulting in 480 conditions. Responses for each voxel were assessed by performing t-tests between each condition (each scene) and baseline.

For each ROI, we sought to establish the similarity in the spatial pattern of response (across voxels) for each pair of conditions. Thus, we performed an iterative variant [S5, S6] of split-half correlation analysis [S7], briefly described below. First, responses to each condition (t-values) were extracted from the voxels within each ROI. Second, the data was split in half by runs, and the mean signal was independently removed from each half of the data. Third, for each ROI, we extracted voxel response patterns from each ROI and cross-correlated these patterns for every possible pair of images (drawn from independent halves of the data). This yielded a similarity matrix for each ROI, in which each point represents an R-value: the response similarity between each pair of conditions.

We next computed Panorama Decoding Indices by subtracting the average between-panorama correlations from the average within-panorama correlations for every scene-view, resulting in a Context Decoding Index for each ROI, where any value significantly greater than zero reflects the presence of similarity between the response patterns evoked by two scene views that shared the same panoramic context.

We employed a region of interest analysis for three reasons: (i) our *a priori* hypotheses focused on the key regions of the scene processing circuit (PPA, RSC, and OPA); (ii) previous reports that the location of these regions varies widely between individuals; and (iii) our scan prescription was restricted for the sake of high spatial resolution (2 x 2 x 2 voxel size), but this precluded whole-brain coverage.

Statistical Analyses

In all analyses, two-tailed, uncorrected P values were calculated from bootstrapped confidence intervals, sampled 10,000 times with replacement. Participants whose task accuracy was determined to fall outside of 2 SDs of the group mean (*Experiment 1*: $N = 2$; *Experiment 4*: $N = 1$) were excluded from the results.

Perceptual Priming Control Study: Invalid vs. Valid Prime

In *Experiment 4*, we demonstrate that memory for the location of a scene view within a panorama (left/right side) is augmented if that view is briefly primed (300 ms) with another image from the same panorama, compared to a black square (Neutral Prime). These results suggest that discrete views from within a panoramic environment prime each other in subsequent perception.

In a control study, we tested whether the priming effect we report in *Experiment 4* could be explained by a non-specific benefit of priming scene representations of any landscape, potentially priming general processes related to scene perception and memory. To do this, we ran a control experiment ($N = 12$) to examine the effect of both a Valid Prime (a view from the same panoramic environment as the test image) and an Invalid Prime (a view from a different panoramic environment) on scene memory. This experiment used the same display methods as used in *Experiment 4* (scenes were explored using a Virtual Reality Headset), except that the Morph Condition was not included in this experiment. All 18 scenes studied in this experiment were drawn from a bank of diverse real-world stimuli.

Our results show that memory for the location of a scene view in a panorama (left/right side) is better when the scene-view is primed with another image from the same panorama (Valid Prime) compared with a view from a different panoramic environment (Invalid Prime). Overall, participants were faster (Valid mean and SEM: 1.56s +/- 0.68, Invalid: 1.87s +/- 0.85, $t(11) = -3.23$, $P = 0.007$) and more accurate (Valid mean and SEM: 0.93 +/- 0.08, Invalid: 0.87 +/- 0.13, $t(11) = 1.5$, $P = 0.15$) to report where a scene view had appeared in the broader panorama (left/right side) in the Valid as compared with the Invalid Prime Condition. This resulted in higher Efficiency Scores (Accuracy/RT) for the Valid compared with the Invalid Prime Condition (Valid mean and SEM: 0.73 +/- 0.37, Invalid: 0.54 +/- 0.20, $t(11) = 2.56$, $P = 0.02$).

Together with *Experiment 4*, this pattern of results demonstrates that memory for a broad, panoramic environment causes discrete views from within this environment to prime each other in subsequent perception, facilitating perceptual judgments.

Supplemental References

- S1. Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* *29*, 162–73.
- S2. Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage* *9*, 195–207.
- S3. Schwarzlose, R. F., Swisher, J. D., Dang, S., and Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 4447–52.
- S4. Hinds, O. P., Rajendran, N., Polimeni, J. R., Augustinack, J. C., Wiggins, G., Wald, L. L., Diana Rosas, H., Potthast, A., Schwartz, E. L., and Fischl, B. (2008). Accurate prediction of V1 location from cortical folds in a surface coordinate system. *Neuroimage* *39*, 1585–1599.
- S5. MacEvoy, S. P., and Epstein, R. A. (2007). Position selectivity in scene- and object-responsive occipitotemporal regions. *J. Neurophysiol.* *98*, 2089–98.
- S6. Kravitz, D. J., Kriegeskorte, N., and Baker, C. I. (2010). High-level visual object representations are constrained by position. *Cereb. Cortex* *20*, 2916–25.
- S7. Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* *293*, 2425–30.