

How Reliable Are Visual Context Effects in the Parahippocampal Place Area?

Russell A. Epstein and Emily J. Ward

Department of Psychology and Center for Cognitive Neuroscience, University of Pennsylvania, Philadelphia, PA 19104-6241, USA

The parahippocampal place area (PPA) is a region of human cortex that responds more strongly to visual scenes (e.g., landscapes or cityscapes) than to other visual stimuli. It has been proposed that the primary function of the PPA is encoding of contextual information about object co-occurrence. Supporting this context hypothesis are reports that the PPA responds more strongly to strong-context than to weak-context objects and more strongly to famous faces (for which contextual associations are available) than to nonfamous faces. We reexamined the reliability of these 2 effects by scanning subjects with functional magnetic resonance imaging while they viewed strong- and weak-context objects, scrambled versions of these objects, and famous and nonfamous faces. "Contextual" effects for objects were observed to be reliable in the PPA at slow presentation rates but not at faster presentation rates intended to discourage scene imagery. We were unable to replicate the earlier finding of preferential PPA response to famous versus nonfamous faces. These results are difficult to reconcile with the hypothesis that the PPA encodes contextual associations but are consistent with a competing hypothesis that the PPA encodes scenic layout.

Keywords: context effects, fMRI, parahippocampal place area, spatial layout, visual recognition, visual scenes

Introduction

Parahippocampal cortex (PHC) is believed to play a key role in both memory and visuospatial cognition; consequently, understanding the precise information processing function of this region is an important challenge for cognitive neuroscience. Human neuroimaging studies have demonstrated the existence of a functional locus in the posterior portion of PHC known as the parahippocampal place area (PPA) that responds more strongly to visual scenes (i.e., images of landscapes, cityscapes, or buildings) than to other complex visual stimuli (Epstein and Kanwisher 1998). Although the response of the PPA to nonscene objects is not negligible (Haxby et al. 2001; Downing et al. 2006; Diana et al. 2008), the fact that it responds maximally to scenes has been taken to be particularly important for understanding its function. The current study tests competing explanations for the scene-preferential response of the PPA.

In the original report on this topic, Epstein and Kanwisher (1998) proposed that the PPA encodes a representation of the spatial layout of the currently visible environment (Epstein and Kanwisher 1998; see also Epstein 2005, 2008). In this view, the PPA responds preferentially to scenes because they convey information about local spatial layout, whereas nonscene stimuli such as blenders, vehicles, and faces typically do not. "Spatial layout" refers to the geometric structure of the scene

as defined primarily by fixed background elements such as walls or other immovable topographical features. Support for this hypothesis comes from a finding that the PPA responds strongly to indoor scenes irrespective of whether they are empty (i.e., just bare walls) or filled with potentially movable objects (i.e., furniture) but only responds weakly to objects removed from the scenes and displayed as a multi-item array on a blank background. In other words, the PPA responds strongly to the fixed background elements of the scene but is largely indifferent to the presence or absence of smaller objects.

Note that in this view, the PPA does not encode "scenes" per se—rather, it encodes spatial layout information which can be extracted from scenes but not from nonscene objects. Other brain regions such as perirhinal cortex and the hippocampus may encode other aspects of the scene, such as the identities of the foreground objects (perirhinal cortex) and the locations of the objects within the scene (hippocampus) (King et al. 2002; Goh et al. 2004; Summerfield et al. 2006; Hartley et al. 2007). Nor does the proposal necessarily imply that the PPA encodes the background elements in detail. Rather, it might be primarily concerned with encoding a local spatial coordinate frame (Shelton and Pippitt 2007), in which case the background elements might only be important insofar as they are fixed entities in the environment to which the spatial coordinate frame can be attached (Epstein 2008).

A prominent alternative to the spatial layout hypothesis has been offered by Bar and colleagues. These authors propose that PHC (including the PPA) encodes visual context, which they define as information about which objects "typically co-occur in the environment around us" (Bar, Aminoff, and Ishai 2008). Under this hypothesis, the parahippocampal response to scenes reflects the activation of a "context frame" representation that includes information about which objects typically appear in that context and where they are likely to be located relative to each other. These authors further propose that there is a division of labor within PHC such that anterior PHC primarily encodes information about the identities of the typical objects, whereas posterior PHC (i.e., the PPA) primarily encodes information about their typical locations.

There are commonalities between the spatial layout hypothesis and the context hypothesis. Notably, both propose that the posterior portion of PHC (i.e., the PPA) encodes spatial information. However, there are several important differences. First, the spatial layout hypothesis suggests that the PPA is largely unconcerned with the discrete objects within the scene, whereas the context hypothesis suggests that the primary function of the PPA is to represent relationships between these objects. For example, on viewing a kitchen scene in which a toaster rests on a countertop, the spatial layout hypothesis

predicts that the PPA represents the countertop but not the toaster, whereas the context hypothesis predicts that the PPA is *primarily* concerned with representing the fact that the toaster is on the counter (or, alternatively, the fact that toasters typically appear on counters). Second, the spatial layout hypothesis suggests that the PPA represents some quantity that is physically present in the scene—the geometry defined by the background elements. In contrast, the context hypothesis suggests that the PPA represents something quite abstract—the typical spatial relationships between objects (Bar and Aminoff 2003; Bar, Aminoff, and Ishai 2008). Finally, the 2 accounts differ on their view of the ultimate function of the PPA. The spatial layout hypothesis emphasizes the importance of geometric structure as a cue for identification of a scene as a particular location in the world (O’Keefe and Burgess 1996; Burgess, Becker, et al. 2001; Epstein 2005; Byrne et al. 2007; Epstein and Higgins 2007; Epstein et al. 2007), whereas the context hypothesis emphasizes the usefulness of context frames for quickly and accurately identifying objects within the scene based on information about their typical co-occurrence (Bar 2004).

The evidence presented thus far for the context hypothesis is 3-fold. First, PHC in general and the PPA in particular are reported to respond more strongly when subjects view objects with strong contextual associations than when they view objects with weak contextual associations (Bar 2004). For example, PHC responds more strongly to images of microscopes or tractors than to images of light bulbs or telephones because microscopes and tractors are strongly associated with particular context frames (laboratory and farm), whereas light bulbs and telephones are not. This effect is reported to occur both when the objects form part of a larger scene including background elements (Bar, Aminoff, and Schacter 2008) and also when they are presented alone on a blank background (Bar and Aminoff 2003). Second, PHC is reported to respond more strongly to famous faces than to nonfamous faces, which is postulated to reflect retrieval of contextual associations that are only available for the famous faces (Bar, Aminoff, and Ishai 2008). Finally, a third study examined neural activity when subjects retrieved information about novel “contexts” consisting of 3 meaningless visual objects whose association was learned over time (Aminoff et al. 2007). Activity in several medial temporal lobe regions including PHC was greater when subjects viewed one object from a context and recalled the visual features of the other associated objects than when they viewed an object that had no contextual associations and made a perceptual judgment about it.

Although the evidence outlined above is at first glance convincing, there are at least 3 methodological aspects of the experiments that might have impacted the results. The first is the fact that stimuli in all experiments were shown at relatively slow presentation rates of 0.33–0.5 Hz, which leaves open the possibility that PHC activity could reflect mental imagery of scenes rather than rapid activation of contextual representations as postulated by the theory (Bar 2004). Under this interpretation, the actual contextual associations that lead to retrieval of an appropriate scene might be encoded elsewhere. The second methodological issue is the fact that no low-level visual control was examined; thus, it is possible that strong- and weak-context objects might differ on low-level visual properties, which are known to affect the PHC (Levy et al. 2001; Rajimehr et al. 2008). Finally, in the study examining response to famous and nonfamous faces, the relevant stimulus classes

were never presented within the same scan runs, making it difficult to directly compare these conditions (Ishai et al. 2005; Bar, Aminoff, and Ishai 2008). For these reasons, the reliability of visual context effects in the PHC must currently be considered tentative.

To clarify these issues, we reexamined the strong- versus weak-context object and the famous versus nonfamous face effects in the PHC. Specifically, we scanned subjects while they viewed images of strong-context objects and weak-context objects, as well as famous and nonfamous faces, under both fast presentation rates (1.25 Hz; Experiment 1) and slow presentation rates (0.33 Hz; Experiment 2). We hypothesized that use of a faster presentation rate would reduce the incidence of scene imagery; thus, replication of the previously observed effects under these presentation rates would strengthen the evidence for the context hypothesis, whereas failure to replicate would challenge this hypothesis. To control for low-level visual differences, scrambled versions of the strong- and weak-context objects were also shown. Finally, all stimulus classes were presented within the same scan runs to facilitate between-condition comparisons. In addition to strong/weak objects, famous/nonfamous faces, and scrambled versions of the strong/weak objects, subjects also viewed images of famous and nonfamous places in order to establish a comparative scene response. To anticipate, our results indicate that once the methodological points above are considered, the evidence for the context hypothesis becomes much less clear.

Materials and Methods

Subjects

Twenty right-handed subjects (5 males, 15 females, median age 21 years, range 19–27) with normal or corrected-to-normal vision participated in Experiment 1, and 14 subjects meeting the same criteria (8 males, 6 females, median age 22 years, range 20–31) participated in Experiment 2. All subjects were recruited from the University of Pennsylvania community and gave written informed consent according to procedures approved by the local institutional review board.

Magnetic Resonance Imaging Acquisition

Scanning was performed at the Hospital of the University of Pennsylvania on a 3-T Siemens Trio equipped with an 8-channel multiple array Nova Medical head coil. T_2^* -weighted images sensitive to blood oxygenation level-dependent contrasts were acquired using a gradient-echo echo-planar pulse sequence (time repetition [TR] = 3000 ms, time echo [TE] = 30 ms, voxel size = $3 \times 3 \times 3$ mm, matrix size = $64 \times 64 \times 45$). Structural T_1 -weighted images for anatomical localization were acquired using a 3D magnetization-prepared rapid gradient-echo pulse sequence (TR = 1620 ms, TE = 3 ms, time to inversion = 950 ms, voxel size = $0.9766 \times 0.9766 \times 1$ mm, matrix size = $192 \times 256 \times 160$). Visual stimuli were rear projected onto a Mylar screen at the head of the scanner with an Epson 8100 3-LCD projector equipped with a Buhl long-throw lens and viewed through a mirror mounted to the head coil. Responses were recorded using a 4-button fiber-optic response pad system.

Stimuli

Stimuli were 400×400 pixel color images from 8 categories: strong-context objects, weak-context objects, famous faces, nonfamous faces, famous places, nonfamous places, and scrambled versions of both strong- and weak-context objects (see Fig. 1). Eighty-eight images were used for each category for a total of 704 images in the complete stimulus set. Strong- and weak-context object images were those graciously provided for public use by M. Bar and colleagues on their laboratory website (<http://barlab.mgh.harvard.edu/ContextLocalizer.htm>) except that to fit the current design one strong- and one weak-context object from the

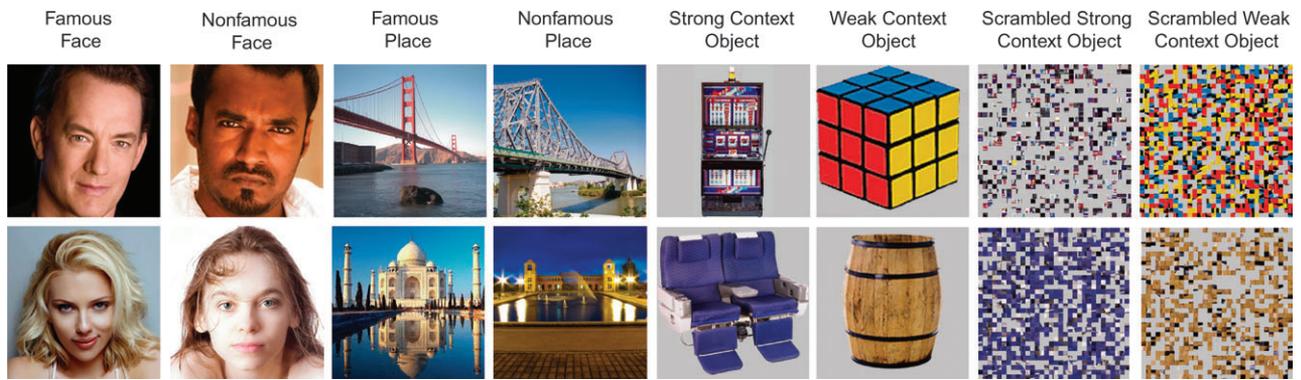


Figure 1. Examples of stimuli. Eighty-eight stimuli were collected for each of the 8 stimulus categories, which were presented at different presentation rates in Experiments 1 and 2. Famous faces were contemporary Hollywood celebrities, whereas famous places were well-known landmarks from around the world. Objects in the strong-context condition in this stimulus set were associated strongly with a specific context (e.g., a slot machine is associated with a casino), whereas objects in the weak-context condition did not have similar associations.

original stimulus set of 89 was discarded. Because some of these object images were smaller than 400×400 pixels, Adobe Photoshop was used to expand them to be equivalent in size to the other images used in the current experiment. Strong-context objects were associated strongly with a specific context, whereas weak-context objects were not; each object belonged to a different context. Photographs of contemporary Hollywood celebrities (e.g., George Clooney and Jennifer Aniston) were included in the famous faces condition, whereas portraits of approximately equal attractiveness and age were selected from a stock image website (www.sxc.hu) for the nonfamous faces condition. Commonly recognizable places (e.g., Eiffel Tower and the Great Wall of China) were used in the famous places condition, and generic landscapes and cityscapes were used in the nonfamous places condition. Scrambled versions of the strong- and weak-context object images were created by scrambling the images within a 30×30 grid.

Procedure

In both experiments, subjects made 1-back repetition detection judgments on visual stimuli that were sequentially presented in a standard block design. The primary difference between Experiments 1 and 2 was the stimulus presentation rate (1.33 Hz in Experiment 1; 0.33 Hz in Experiment 2).

Experiment 1 consisted of 4 scan runs, each 6 min 15 s long and composed of 16 18-s blocks during which visual stimuli were shown interleaved with four 18-s periods of fixation and a 15-s fixation period at the end of the scan. In each stimulus block, 22 unique images from a single stimulus category (e.g., 22 strong-context objects) were shown along with two 1-back repetitions; each image appeared for 400 ms and was followed by 350-ms blank interstimulus interval. The order of the condition blocks was randomized for each run, subject to the constraint that all 8 stimulus conditions must be shown before a block of the same type appeared again. Thus, each stimulus condition appeared twice within each run. Each of the 704 images in the stimulus set was presented once in runs 1–2 and once in runs 3–4 (except for images selected for 1-back repetition, which were presented additional times).

Experiment 2 consisted of 4 scan runs, each 10 m 45 s long and composed of 32 18-s blocks during which visual stimuli were shown interleaved with eight 6-s periods of fixation and a 21-s fixation period at the end of the scan. In each stimulus block, 5 unique images from a single stimulus category were shown along with one repeated image; each image appeared for 2800 ms and was followed by a 200-ms blank interstimulus interval. The order of the stimulus blocks was randomized. Each of the 704 images in the stimulus set was presented once during the experiment (except for images selected for 1-back repetition, which were presented twice).

Data Analysis

Functional magnetic resonance imaging (fMRI) data were corrected for differences in slice timing by resampling slices in time to match the first

slice of each volume, realigned with respect to the first image acquired during a scanning session, spatially normalized to the Montreal Neurological Institute template, and then spatially smoothed with a 6-mm full-width half-maximum Gaussian filter. Data were analyzed using the general linear model as implemented in VoxBo (www.voxbo.org) including an empirically derived $1/f$ noise model, filters that removed strong and low temporal frequencies, regressors to account for global signal variations, and nuisance regressors to account for between-scan differences. Eight regressors (one for each stimulus condition) were used to model the effects of interest; each consisted of a boxcar function convolved with a standard hemodynamic response function.

Functional regions of interest (ROIs) were defined using data from the experiment rather than from separate functional localizer scans. The PPA ROI was defined as the set of contiguous voxels that responded more strongly to scenes (famous and nonfamous places) than to common objects (strong- and weak-context objects) in the posterior parahippocampal/collateral sulcus region. Significance thresholds were set on a subject-by-subject basis so that ROIs were consistent with those identified in previous studies (Epstein and Kanwisher 1998; Epstein et al. 1999, 2007); thresholds ranged from $t > 3.0$ to $t > 3.5$. Note that the defining contrast for the PPA (places > objects) was independent of the contrasts of interest (strong- vs. weak-context objects or famous vs. nonfamous faces); thus, use of an “internal” localizer did not bias the magnitude of these critical contrasts. In addition, to further examine response in the region showing the strongest context effect, a second ROI was defined based on greater response to strong-context versus weak-context objects in a random effects group analysis. The time course of magnetic resonance response was then extracted from each ROI (averaging over all voxels) and reentered into the general linear model, which was used to calculate parameter estimates (beta values) for the 8 conditions of interest, which were then used as dependent variables in a second-level random effects analysis of variance.

Results

Experiment 1

fMRI response in the left and right PPA for all 8 conditions is plotted in the left and middle panel of Figure 2a. As expected, both hemispheres responded much more strongly to places than to objects, reflecting the fact that the PPA is defined as the region that exhibits this effect. Our focus here, however, is on the differential PPA response to strong- versus weak-context objects and famous versus nonfamous faces, which are the effects predicted by the context hypothesis and are orthogonal to the contrasts used to define the region.

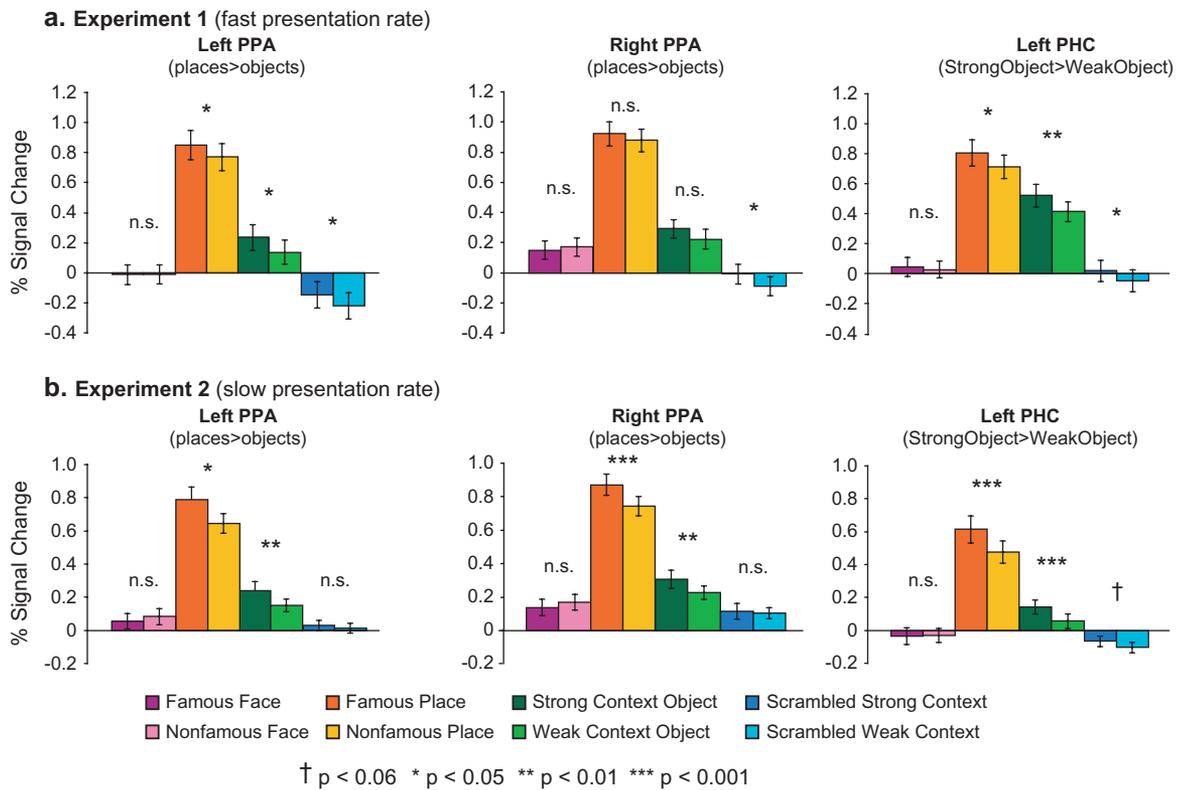


Figure 2. fMRI response within the left PPA, right PPA, and left PHC for (a) Experiment 1 and (b) Experiment 2. The PPA was functionally defined in each subject as the set of voxels that responded more strongly to places (famous and nonfamous) than to objects (strong context and weak context). The PHC was functionally defined as the set of voxels that responded more strongly to strong-context objects than to weak-context objects in a group analysis. Significance markers reflect the results of 4 pairwise comparisons (famous vs. nonfamous faces, famous vs. nonfamous places, strong-context vs. weak-context objects, and scrambled strong context vs. scrambled weak context); error bars are ± 1 standard error of the mean for each condition. Greater response was observed for strong-context objects than for weak-context objects in the left PPA and PHC in Experiment 1 and in all 3 ROIs in Experiment 2. However, a similar “strong” > “weak” advantage was observed for scrambled stimuli in Experiment 1, which is counter to what would be predicted by the context hypothesis. Neither the PPA nor the PHC responds differentially to famous versus nonfamous faces (contrary to previous reports), although they did respond more strongly to famous than to nonfamous places.

We observed a small advantage for strong- versus weak-context objects, which was significant in the left PPA ($t_{19} = 2.8$, $P < 0.05$) but not in the right PPA ($t_{19} = 1.8$, $P = 0.09$, not significant [NS]), partially replicating previous results. However, this differential response disappeared when the response to the scrambled objects was subtracted from each condition to control for low-level visual differences between the stimulus sets ([strong vs. weak] \times [intact vs. scrambled] interaction: F values < 1 , NS, both hemispheres). We observed no evidence that the PPA responded differentially to famous versus nonfamous faces (both t values < 1 , NS), contrary to previous reports. On the other hand, we did observe an unexpected effect of place familiarity: the PPA responded more strongly to famous than to nonfamous places in the left hemisphere ($t_{19} = 2.3$, $P < 0.05$) although not in the right ($t_{19} = 1.2$, NS).

One might argue that our apparent failure to fully replicate previously reported context effects was impacted by the choice of ROI. Because the PPA is defined as the region that responds more strongly to scenes than to objects, voxels that respond strongly to objects are selected against, making it possible that the voxels that represent the contextual associations to objects (and famous faces) were unnecessarily excluded. To guard against this scenario, we defined a left parahippocampal ROI based on a random effects group analysis of the strong-context versus weak-context contrast (for intact objects only). A group analysis was used because this effect was

not strong enough to define ROIs on a subject-by-subject basis. Results from this ROI are plotted in the rightmost panel of Figure 2a.

As expected, the differential response to places versus objects was not as dramatic in this context region as in the PPA. However, with regards to the contrasts relevant to the context hypothesis, the results were the same as in the PPA. Specifically, there was an advantage for strong- versus weak-context intact objects ($t_{19} = 3.4$, $P < 0.01$), but this advantage disappeared when the response to strong versus weak scrambled objects was subtracted from each condition ([strong vs. weak] \times [intact vs. scrambled] interaction: $F_{1,19} = 1.0$, $P = 0.32$, NS). Note that this means that the context effect was of equivalent magnitude for intact and for scrambled objects, even though the ROI was specifically chosen based on differential response to strong versus weak intact objects. We once again find no evidence that PHC responds more strongly to famous than to nonfamous faces ($t < 1$, NS), although it did respond more strongly to famous than to nonfamous places ($t_{19} = 2.8$, $P < 0.05$).

Experiment 2

In Experiment 2, subjects viewed the same stimuli as in Experiment 1 but at a slower presentation rate in order to more precisely replicate the conditions of previous studies of visual context. Of particular interest was whether the strong- versus weak-context and famous versus nonfamous effects in PHC,

which have been previously associated with contextual processing, would be more reliable at these slower presentation rates.

fMRI response in the PPA is plotted in the left and middle panels of Figure 2*b*. We observed greater response to strong-context objects than to weak-context objects in both hemispheres (left $t_{11} = 3.1$, right $t_{11} = 3.7$, both P values < 0.01). In contrast to Experiment 1, here the strong versus weak difference was significantly larger for intact than for scrambled objects ([strong vs. weak] \times [intact vs. scrambled] interaction: left $F_{1,11} = 7.8$, right $F_{1,11} = 7.7$, both P values < 0.05). Importantly, we once again found no evidence for a differential response to famous versus nonfamous faces (t values < 1 , NS) but did observe greater response to famous than to nonfamous places (left $t_{11} = 4.0$, $P < 0.05$; right $t_{11} = 4.5$, $P < 0.001$). These results were not substantially different when PHC was defined based on the strong-context $>$ weak-context contrast instead of places $>$ objects (Fig. 2*b*, rightmost panel).

Relative Strength of the Context and Place Effects

If the intrinsic function of PHC is to represent contextual associations, then one might expect the context effect (strong- vs. weak-context objects) and place effect (places vs. objects) to be of roughly equivalent magnitude. In contrast, if context effects are driven by second-order considerations such as stimulus differences or scene imagery, then one would expect the context effect to be smaller than the place effect. To examine this question, we directly compared the strength of these effects within the parahippocampal/lingual region for each experiment.

For this analysis, use of a PPA ROI would be inappropriate because the PPA is defined as the region that exhibits the strongest place versus object effect. We used the following technique to define a new ROI that would not be biased toward one effect or the other. First, we identified the 100 voxels in the left parahippocampal/lingual region that exhibited the strongest differential response for each of the 2 contrasts of interest in the random effects group analysis for each experiment. We then defined a joint ROI based on the union of these 2 component ROIs. This ROI consisted of 178 voxels for Experiment 1 and 157 voxels in Experiment 2, indicating partial but incomplete overlap between context-responsive and place-responsive regions. Crucially, this ROI was not differentially weighted toward a context-sensitive or place-sensitive response, except insofar as one of these 2 effects might be inherently stronger than the other. We focused on the left hemisphere in particular because our earlier analyses indicated that the context effect was more reliable in this hemisphere. Thus, by focusing on this hemisphere, we ensured that the context effect was observed at its point of maximum advantage.

When compared directly with each other (Fig. 3), the context effect was significantly weaker than the place effect in both Experiment 1 ($t_{19} = 2.7$, $P < 0.05$) and Experiment 2 ($t_{13} = 6.7$, $P < 0.0001$). Indeed, inspection of the figure reveals that there was no overlap in the strength of the 2 effects in Experiment 2. The relative strength of the 2 effects was even more apparent when their statistical significance was examined on a subject-by-subject basis. The place effect was found to be reliable ($P < 0.05$, 1 tailed) in 19/20 subjects in Experiment 1 and 14/14 subjects in Experiment 2 (average t values 4.1 and 8.4, respectively). In contrast, the context effect was only reliable in a small subset of subjects (3/20 in Experiment 1, 3/14 in Experiment 2, average t values 0.57 and 0.87, re-

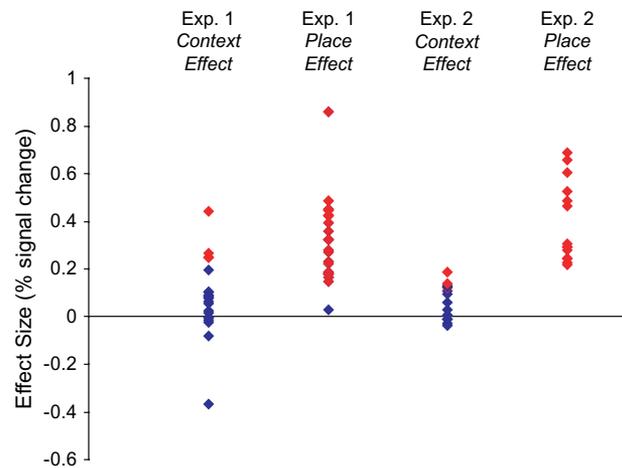


Figure 3. The magnitude and significance of the place and context effects within PHC, plotted on a subject-by-subject basis. “Place effect” is defined as greater response to places (famous and nonfamous) than to objects (strong and weak context); “context effect” is defined as greater response to strong-context objects than to weak-context objects. Red diamonds indicate that the effect was significant at the individual subject level ($P < 0.05$); blue diamonds indicate a nonsignificant subjectwise effect. Whereas the place effect was significant in almost every subject, the context effect was only reliable in a minority of subjects.

spectively). These results were not significantly different when a smaller ROI consisting of the top 50 voxels for each contrast was used.

Although the results of these analyses suggest that context effects in the parahippocampal-lingual region are significantly weaker than place effects, an alternative possibility is that there are separate parahippocampal-lingual subregions for processing contextual associations and scene layout. Under this scenario, use of a joint ROI could have led to a situation in which the response in the context-processing region was swamped by a larger yet anatomically distinct response in the place-processing region. To guard against this possibility, we performed an additional analysis in which we compared the strength of the place and context effects within the 100 left parahippocampal-lingual voxels that exhibited the strongest context effect. Even in this region, which was selected on the basis of its differential response to strong-versus weak-context objects, the place effect was significantly larger than the context effect (Experiment 1, $t_{19} = 2.8$, $P < 0.05$; Experiment 2, $t_{13} = 6.8$, $P < 0.001$). Furthermore, when statistical tests were performed in this region, on a subject-by-subject basis, the place effect was found to be significant in 14/20 subjects in Experiment 1 and 14/14 subjects in Experiment 2 (average t values 3.3 and 8.6), whereas the context effect was only significant in 4/20 subjects in Experiment 1 and 4/14 subjects in Experiment 2 (average t values 0.82 and 1.04). These results were not substantially different when response was examined within a smaller ROI consisting of the top 50 context-sensitive voxels.

In sum, the results above indicate that context effects are secondary to place effects, even within the territory that exhibits the strongest context effect. Furthermore, context effects were only significant in a minority of subjects, whereas place effects were significant in almost all subjects examined. These results give credence to the idea that response differences between strong- and weak-context objects might be attributed to secondary factors such as the tendency of the stimuli to elicit scene imagery, a factor that would be expected to vary considerably across subjects.

Anatomical Loci of the Context and Place Effects

The previous analysis indicated that the overlap between the set of voxels showing the largest context effect and the set of voxels showing the largest place effect was incomplete. In order to better understand the cortical loci of these effects, we discarded the ROI approach and performed random effects voxelwise group analyses on the data from Experiment 2. (We focused on this experiment because both the context and place effects were significant in this case.) Results are shown in Figure 4. The place and context effects appeared to be centered on the same point: along the collateral sulcus, encompassing both posterior PHC and the anterior lingual gyrus (Aguirre et al. 1998). This finding is consistent with the picture of a single mechanism housed by this region that underlies both effects—for example, a single region that supports both scene perception and scene imagery.

Outside of PHC, this whole-brain analysis revealed a larger network of regions that responded more strongly to places

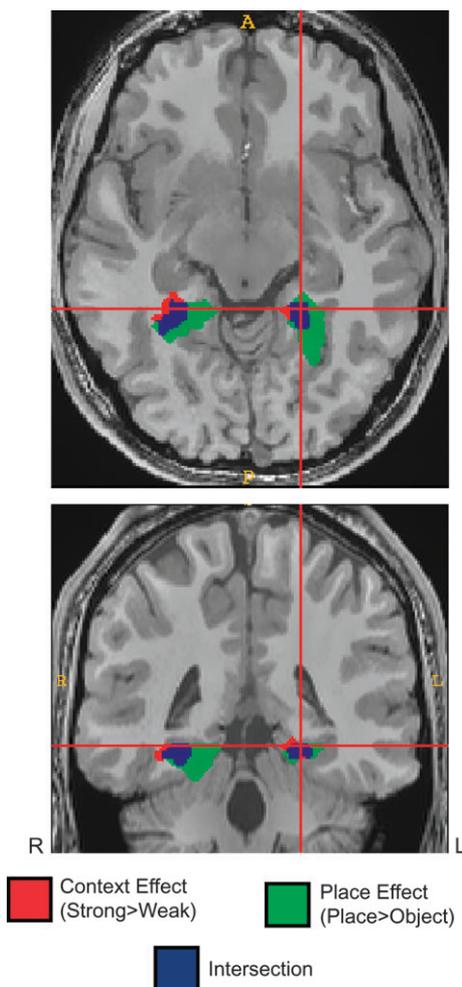


Figure 4. Anatomical loci of the place and context effects in the PHC region in Experiment 2. As in Figure 3, the place effect is defined as greater response to places than to objects, whereas the context effect is defined as greater response to strong-context objects than to weak-context objects. A more stringent threshold was applied to the place effect ($P < 0.0001$) than to the context effect ($P < 0.01$) in order to facilitate comparison of the loci of the 2 effects when plotted at comparable extent. The 2 effects exhibit largely overlapping locations in bilateral PHC/lingual gyrus, although the territory affected by place effect extends more posteriorly. (Note that activations outside of the PHC region are not shown; see text.)

than to objects ($P < 0.05$ corrected; permutation test), including the retrosplenial complex (extending from the parietal-occipital sulcus anteriorly into the anterior calcarine sulcus), the transverse occipital sulcus, a large swath of posterior visual cortex, the lateral geniculate nucleus, and orbitofrontal cortex. No region responded more strongly to strong-context objects than to weak-context objects at this threshold. Relaxing the threshold ($P < 0.001$ uncorrected; 6 contiguous voxels) revealed strong-context > weak-context activation in PHC (extending posteriorly along the collateral sulcus), the anterior calcarine sulcus, and the right middle occipital gyrus. Famous > nonfamous face activation was observed in the precuneus ($P < 0.05$, corrected), with additional subthreshold activity ($P < 0.001$ uncorrected, 6 contiguous voxels) in the medial retrosplenial region, 2 foci in the right temporal lobe (occipitotemporal sulcus and superior temporal sulcus), and the left superior temporal sulcus near the temporal pole. Famous > nonfamous place activation was observed in the right hippocampus, medial retrosplenial region, left occipitotemporal sulcus, transverse occipital sulcus, left middle temporal lobe, left inferior frontal gyrus (Brodmann Area 44), and left supplementary motor area (all $P < 0.05$ corrected). Interestingly, the medial retrosplenial region exhibiting the famous > nonfamous place effect was largely identical to the medial retrosplenial region exhibiting the famous > nonfamous face effect but distinct from the more lateral retrosplenial region that responded differentially to places versus objects.

Discussion

The current study reexamined the evidence for a contextual processing account of the PHC in general and the PPA in particular. Previous studies have reported greater PHC/PPA response to strong-context objects than to weak-context objects (Bar and Aminoff 2003; Diana et al. 2008; but see Yue et al. 2007) and greater response to famous faces than to nonfamous faces (Bar, Aminoff, and Ishai 2008), both of which were taken as evidence for parahippocampal instantiation of contextual associations. We only partially replicated these results. In particular, the advantage for strong- versus weak-context objects was only reliable at slow presentation rates (Experiment 2). At faster presentation rates, this context effect did not survive subtraction of the response to scrambled objects (Experiment 1). Furthermore, we completely failed to replicate the previously observed famous versus nonfamous face effect (Experiments 1 and 2). Taken as a whole, these results weaken the case for a context-processing account of PHC/PPA function.

In this discussion, we first consider the implications of the current results for the context hypothesis and then more broadly evaluate the functional role of the PHC/PPA. Note that we will use the terms PHC and PPA interchangeably, reflecting the fact that context effects, when observed, can be localized to a PHC region equivalent to the PPA (e.g., Fig. 4). However, this conflation of terms should not be taken to mean that the functionally defined PPA is equivalent to the anatomically defined PHC, as the anterior portion of PHC may be outside the PPA and may have a distinct function (Bar and Aminoff 2003).

Testing the Context Hypothesis

Using the stimuli of Bar and colleagues, we were able to replicate the finding that the PHC/PPA responded more vigorously to strong-context objects than to weak-context objects at slow presentation rates (0.33 Hz) comparable to those used in previous

experiments. However, these effects were not reliable at faster presentation rates of 1.33 Hz. Although we did observe a moderate advantage for strong- versus weak-context objects in the left PPA at these faster rates, this advantage was no greater for intact objects than for uninterpretable scrambled versions of the same objects, which are unlikely to evoke contextual associations. These results suggest that response differences between the strong- and weak-context objects at faster presentation rates are likely to be driven by low-level physical differences between the stimulus sets such as color differences, luminance differences, texture differences, or differences in the extent to which the stimulus covers the screen, all of which would be preserved by spatial scrambling. Indeed, previous studies have indicated that response in the parahippocampal-lingual-fusiform region is sensitive to these purely visual aspects of the stimulus (Levy et al. 2001; Cant and Goodale 2007; Rajimehr et al. 2008), although as in the current experiment these effects are secondary to the larger categorical effect of places versus objects.

At slower presentation rates, on the other hand, the strong-versus weak-context difference might be driven mostly by the tendency of the strong-context objects to evoke scene imagery, which is known to strongly activate the PPA (O'Craven et al. 1999). This hypothesis is consistent with the finding that the strong- versus weak-context effect is only significant in a subset of subjects, which suggests that PPA response to strong-context objects is not automatic but might depend on individual differences in how subjects respond mentally to each item—for example, whether they think of related scenes or not. Although we do not have direct evidence that subjects formed mental images of scenes in response to the strong-context objects during Experiment 2, the presentation rates were slow enough that they would have had time to do so. Thus, a mental imagery explanation for the context effect seems at least as plausible as a contextual processing account and has the advantage that it is more consistent with the body of evidence that implicates the PPA in scene processing.

It is important to note that a scene imagery account of PPA activation is substantially different from a contextual association account. Under the scene imagery account, the role of the PPA is to encode the visual structure of scenes; information about which scenes are associated with which objects might be encoded in the PPA, but might just as well be encoded elsewhere in the brain (e.g., in the anterior temporal lobes). Under the contextual association account, on the other hand, the intrinsic role of the PPA is to encode associations between contextually related objects; whether a subject experiences a scene image in response to a strong-context object is a secondary consideration.

Our second major finding was that the PPA responds equally strongly to famous and nonfamous faces. This finding, observed in both experiments, contradicts an earlier report indicating that PHC/PPA responds more strongly to famous than to nonfamous faces, which was taken to indicate activation of contextual associations in response to the famous face stimuli. Although it is not entirely clear why we did not replicate the earlier results, it is worth noting that a close reading of the original paper (Ishai et al. 2005) for which the report of Bar, Aminoff, and Ishai (2008) is a reanalysis indicates that famous and nonfamous faces were never shown within the same scan runs. This makes the earlier design suboptimal for comparing the response levels between famous and nonfamous faces because differences between conditions might be confounded

by scanwise differences in the fMRI response. The current experiment uses a more standard fMRI design in which blocks of each condition type are shown within every scan run, allowing for within-scan comparisons of response level. The fact that we do not observe any difference in PPA response to famous versus nonfamous faces in either experiment while using this more sensitive design and analysis method suggests that the previously reported effect is of questionable reliability. Indeed, previous studies of face recognition have generally failed to observe famous versus nonfamous face differences in PHC (Leveroni et al. 2000; Gorno-Tempini and Price 2001; Trinkler et al. 2009).

Finally, as noted in the Introduction, a third piece of evidence for the context hypothesis comes from a study by Aminoff et al. (2007) that measured fMRI activity while subjects recalled novel contexts consisting of 3 meaningless objects whose association was learned during an extensive prescan training regime. During the scan session of this earlier study, subjects viewed single objects from each triplet and were asked to report whether the associated (nonvisible) objects were multicolored or not. PHC/PPA responded more strongly during this memory retrieval task than during a perceptual task in which subjects reported whether or not a currently visible object (which had never been associated with other objects) was multicolored. These results were taken as evidence for PHC/PPA involvement in contextual processing because PHC/PPA activity was observed when retrieving information about objects that had been associated with the stimulus object.

Although the present data do not speak directly to these results, it is worth noting that context effects in the experiment of Aminoff et al. (2007) were found not only in PHC/PPA but also in perirhinal cortex, the hippocampus, and many other brain regions (retrosplenial cortex, fusiform gyrus, intraparietal sulcus, parietal-occipital junction, caudate nucleus, lateral-occipital complex, inferior frontal cortex, and medial prefrontal cortex). Thus, these results do not uniquely implicate PHC/PPA in contextual processing. In fact, a simpler interpretation of the data of Aminoff et al. (2007) is that tasks that require memory retrieval activate a wide cortical network (including medial temporal lobe regions such as PHC, perirhinal cortex, and the hippocampus) more strongly than tasks that do not require memory retrieval. This might be especially true when the memory retrieval task is significantly harder (i.e., longer response times and lower accuracies) than the perceptual task as was the case in Aminoff's study.

What Does the PPA Do?

The PPA was initially identified on the basis of its differential response to scenes versus nonscene objects. Unlike the differential response to strong- versus weak-context objects, this effect is extremely reliable, being observable in almost all subjects scanned. In order to understand the function of the PPA, it is essential to understand its scene-preferential response. The spatial layout hypothesis proposes that the PPA responds strongly to scenes because they convey information about the spatial layout of the local visual environment. In contrast, the context hypothesis proposes that the PPA responds strongly to scenes because they depict a group of objects that are strongly contextually associated with each other. Thus, a critical difference between these hypotheses, at least as formulated thus far, is in whether or not the PPA

encodes information about nonscene objects: whereas the spatial layout hypothesis suggests that the PPA plays little or no role in encoding these objects, the context hypothesis suggests that encoding relationships between discrete objects is its primary function.

Epstein and Kanwisher (1998) initially rejected the idea that the primary function of the PPA was to encode information about the relationship between discrete objects based on the finding that the PPA responds strongly to scenes that were denuded of objects (e.g., empty rooms and barren landscapes) but responds quite weakly to arrays of objects on a blank background. Indeed, the PPA response to a multiple object array was no greater than its response to a single object. The idea that the PPA encodes scenic background elements rather than foreground objects is further supported by the finding that the PPA responds more strongly to layouts made out of Lego blocks than to objects constructed from the same materials (Epstein et al. 1999) and also by neuropsychological findings, indicating that patients with parahippocampal-lingual damage cannot identify places based on their overall scenic structure but can identify individual objects within the scene (Habib and Sirigu 1987; Aguirre and D'Esposito 1999; Epstein et al. 2001; Mendez and Cherrier 2003). In contrast, a patient (DF) with almost complete obliteration of the object's form-processing pathway could identify scenes (and showed activation in the PPA while doing so) despite being unable to identify the objects within them. Neuroanatomical results are also consistent with the idea that there are separate processing streams for extracting object information and spatial layout information from scenes (Kim et al. 2006). In particular, a putative PPA role in spatial layout processing is consistent with its position as a major input to the neighboring hippocampal formation that processes locations relative to environmental surface geometry but not relative to single object-like landmarks (Doeller et al. 2008; see also Hartley et al. 2007).

In contrast, the evidence for the context hypothesis is less convincing. The initial finding of greater response to strong-versus weak-context objects only replicates at slow presentation rates, is only reliable in a minority of subjects, and can be alternatively explained in terms of scene imagery. The finding of greater response to famous versus nonfamous faces does not replicate. The finding that the PPA responds more strongly during recall of contextually associated objects than during perceptual judgments on noncontextually associated objects does not, on close analysis, definitively implicate the PPA in contextual processing. There have been no neuropsychological studies investigating the context hypothesis, and it is unclear how this hypothesis would explain findings such as intact PPA response in a patient with no ability to identify objects. Finally, there are several results in the neuroimaging literature that are hard to explain in terms of contextual processing, such as the weak PPA response to multi-object arrays or the differential response to Lego scenes versus Lego objects; to our knowledge, the proponents of the context hypothesis have not attempted to fit these earlier results into their theory.

Based on these observations, we conclude that extant data from the perception literature support the idea that the PPA encodes spatial/scenic information rather than contextual associations. (Although the extent to which "layout" is purely geometric as opposed to geometric + visual has not yet been established; for discussion, see Epstein 2008.) There is another line of research to consider, however. PHC is often activated in

fMRI studies of memory, even when the study involves non-scenic memoranda. In fact, the context hypothesis was first formulated as an explicit effort to describe a cognitive function that could explain parahippocampal involvement in both spatial navigation and episodic memory tasks. Before rejecting the context hypothesis, we must consider the evidence from these episodic memory experiments.

Parahippocampal activity during memory tasks often occurs in tandem with hippocampal activity and is associated with explicit recollection of the encoding episode. It contrasts with activity in perirhinal cortex, which is associated with a feeling of familiarity that can occur even in the absence of episodic recollection. Although it is not entirely clear whether the anatomical locus of this parahippocampal memory-related activity is the same as the PPA (for discussion, see Epstein 2008), for present purposes, we will assume that episodic memory and scene perception studies activate the same region. It is sometimes hypothesized that PHC/PPA contributes to episodic memory by encoding the contexts within which focally attended items are encountered (e.g., Diana et al. 2007). The critical question then becomes whether context in this case means the spatial surroundings of the item (i.e., the "scene") or whether it can mean something more general.

We believe that the memory literature does not compel adoption of the view that PHC/PPA encodes nonspatial/nonscenic episodic context. In a review of the literature, Diana et al. (2007) report that PHC activity during episodic recollection was observed in about half (14/26) of the studies that used either a remember/know or item/source paradigm. Although Diana and colleagues take this as support for PHC encoding of context, the fact that PHC activity is not found in almost half of the reviewed studies suggests that it may not be essential for episodic recollection. (In contrast, the hippocampus does appear to be essential because hippocampal activity was observed in 21/26 studies.) Furthermore, it is notable that in several of the experiments in which PHC activity was observed, scenes were either the study material (Sharot et al. 2004; Dolcos et al. 2005; Kensinger and Schacter 2006) or recollection was assessed by successful report of spatial/scenic information that accompanied studied words (Cansino et al. 2002; Davachi et al. 2003; Kahn et al. 2004; Johnson and Rugg 2007; see also Hayes et al. 2007). Even the few studies that have found PHC activity corresponding to successful recollection of nonscenic/nonspatial aspects of the learning episode (e.g., Ranganath et al. 2004) are susceptible to a spatial account because PHC activity might correspond to reinstatement of the spatial aspects of the learning episode even though these are not the aspects that subjects are required to report. Finally, it is worthwhile to note that a recent study (Awipi and Davachi 2008) reversed the standard stimulus assignments in an item-source paradigm by using scenes as the target items and assessing source memory by accurate recall of an associated "contextual" object. In this case, successful source memory recall (i.e., recollection) corresponded to a (albeit nonsignificant) decrease in PHC response rather than the increase predicted by the context hypothesis. This finding echoes the previous report by Burgess, Maguire et al. (2001) that PHC activity during episodic recollection reflects retrieval of spatial information (where did I get this object?) but not nonspatial information (who did I get this object from?).

In summary, we have reviewed the evidence presented thus far for the context hypothesis and reexamined two of the most

prominent effects. Our data suggest that these effects are less reliable than previously claimed. Contrary to recent claims that the evidence for contextual processing in the PPA/PHC is “unequivocal” (Bar, Aminoff, and Schacter 2008), we believe that the function of this region is very much open for debate and that the earlier idea that it represents the spatial layout of scenes is still quite viable.

Funding

National Institutes of Health (EY-016464 to RAE).

Notes

We thank Sean MacEvoy, Mary E. Smith, and 2 anonymous reviewers for helpful comments on the manuscript. *Conflict of Interest*: None declared.

Address correspondence to Russell Epstein, Department of Psychology, 3720 Walnut Street, Philadelphia, PA 19104-6241, USA. Email: epstein@psych.upenn.edu.

References

- Aguirre GK, D'Esposito M. 1999. Topographical disorientation: a synthesis and taxonomy. *Brain*. 122:1613-1628.
- Aguirre GK, Zarahn E, D'Esposito M. 1998. An area within human ventral cortex sensitive to “building” stimuli: evidence and implications. *Neuron*. 21:373-383.
- Aminoff E, Gronau N, Bar M. 2007. The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb Cortex*. 17:1493-1503.
- Awipi T, Davachi L. 2008. Content-specific source encoding in the human medial temporal lobe. *J Exp Psychol Learn Mem Cogn*. 34:769-779.
- Bar M. 2004. Visual objects in context. *Nat Rev Neurosci*. 5:617-629.
- Bar M, Aminoff E. 2003. Cortical analysis of visual context. *Neuron*. 38:347-358.
- Bar M, Aminoff E, Ishai A. 2008. Famous faces activate contextual associations in the parahippocampal cortex. *Cereb Cortex*. 18:1233-1238.
- Bar M, Aminoff E, Schacter DL. 2008. Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J Neurosci*. 28:8539-8544.
- Burgess N, Becker S, King JA, O'Keefe J. 2001. Memory for events and their spatial context: models and experiments. *Philos Trans R Soc Lond B Biol Sci*. 356:1493-1503.
- Burgess N, Maguire EA, Spiers HJ, O'Keefe J. 2001. A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage*. 14:439-453.
- Byrne P, Becker S, Burgess N. 2007. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol Rev*. 114:340-375.
- Cansino S, Maquet P, Dolan RJ, Rugg MD. 2002. Brain activity underlying encoding and retrieval of source memory. *Cereb Cortex*. 12:1048-1056.
- Cant JS, Goodale MA. 2007. Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cereb Cortex*. 17:713-731.
- Davachi L, Mitchell JP, Wagner AD. 2003. Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci USA*. 100:2157-2162.
- Diana RA, Yonelinas AP, Ranganath C. 2007. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn Sci*. 11:379-386.
- Diana RA, Yonelinas AP, Ranganath C. 2008. High-resolution multi-voxel pattern analysis of category selectivity in the medial temporal lobes. *Hippocampus*. 18:536-541.
- Doeller CF, King JA, Burgess N. 2008. Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proc Natl Acad Sci USA*. 105:5915-5920.
- Dolcos F, LaBar KS, Cabeza R. 2005. Remembering one year later: role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proc Natl Acad Sci USA*. 102:2626-2631.
- Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N. 2006. Domain specificity in visual cortex. *Cereb Cortex*. 16:1453-1461.
- Epstein R, DeYoe EA, Press DZ, Rosen AC, Kanwisher N. 2001. Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cogn Neuropsychol*. 18:481-508.
- Epstein R, Harris A, Stanley D, Kanwisher N. 1999. The parahippocampal place area: recognition, navigation, or encoding? *Neuron*. 23:115-125.
- Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. *Nature*. 392:598-601.
- Epstein RA. 2005. The cortical basis of visual scene processing. *Vis Cogn*. 12:954-978.
- Epstein RA. 2008. Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci*. 12:388-396.
- Epstein RA, Higgins JS. 2007. Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cereb Cortex*. 17:1680-1693.
- Epstein RA, Parker WE, Feiler AM. 2007. Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *J Neurosci*. 27:6141-6149.
- Goh JOS, Siong SC, Park D, Gutchess A, Hebrank A, Chee MWL. 2004. Cortical areas involved in object, background, and object-background processing revealed with functional magnetic resonance adaptation. *J Neurosci*. 24:10223-10228.
- Gorno-Tempini ML, Price CJ. 2001. Identification of famous faces and buildings: a functional neuroimaging study of semantically unique items. *Brain*. 124:2087-2097.
- Habib M, Sirigu A. 1987. Pure topographical disorientation—a definition and anatomical basis. *Cortex*. 23:73-85.
- Hartley T, Bird CM, Chan D, Cipolotti L, Husain M, Vargha-Khadem F, Burgess N. 2007. The hippocampus is required for short-term topographical memory in humans. *Hippocampus*. 17:34-48.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*. 293:2425-2430.
- Hayes SM, Nadel L, Ryan L. 2007. The effect of scene context on episodic object recognition: parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus*. 17:873-889.
- Ishai A, Schmidt CF, Boesiger P. 2005. Face perception is mediated by a distributed cortical network. *Brain Res Bull*. 67:87-93.
- Johnson JD, Rugg MD. 2007. Recollection and the reinstatement of encoding-related cortical activity. *Cereb Cortex*. 17:2507-2515.
- Kahn I, Davachi L, Wagner AD. 2004. Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *J Neurosci*. 24:4172-4180.
- Kensinger EA, Schacter DL. 2006. Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *J Neurosci*. 26:2564-2570.
- Kim M, Ducros M, Carlson T, Ronen I, He S, Ugurbil K, Kim DS. 2006. Anatomical correlates of the functional organization in the human occipitotemporal cortex. *Magn Reson Imaging*. 24:583-590.
- King JA, Burgess N, Hartley T, Vargha-Khadem F, O'Keefe J. 2002. The human hippocampus and viewpoint dependence in spatial memory. *Hippocampus*. 12:811-820.
- Leveroni CL, Seidenberg M, Mayer AR, Mead LA, Binder JR, Rao SM. 2000. Neural systems underlying the recognition of familiar and newly learned faces. *J Neurosci*. 20:878-886.
- Levy I, Hasson U, Avidan G, Hendler T, Malach R. 2001. Center-periphery organization of human object areas. *Nat Neurosci*. 4:533-539.
- Mendez MF, Cherrier MM. 2003. Agnosia for scenes in topographagnosia. *Neuropsychologia*. 41:1387-1395.
- O'Craven KM, Downing PE, Kanwisher N. 1999. fMRI evidence for objects as the units of attentional selection. *Nature*. 401:584-587.
- O'Keefe J, Burgess N. 1996. Geometric determinants of the place fields of hippocampal neurons. *Nature*. 381:425-428.
- Rajimehr R, Devaney K, Young J, Postelnicu G, Tootell RB. 2008. The ‘parahippocampal place area’ responds selectively to high spatial frequencies in humans and monkeys. *J Vision*. 8:85a.

- Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M. 2004. Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia*. 42:2-13.
- Sharot T, Delgado MR, Phelps EA. 2004. How emotion enhances the feeling of remembering. *Nat Neurosci*. 7:1376-1380.
- Shelton AL, Pippitt HA. 2007. Fixed versus dynamic orientations in environmental learning from ground-level and aerial perspectives. *Psychol Res*. 71:333-346.
- Summerfield JJ, Lepsien J, Gitelman DR, Mesulam MM, Nobre AC. 2006. Orienting attention based on long-term memory experience. *Neuron*. 49:905-916.
- Trinkler I, King JA, Doeller CF, Rugg MD, Burgess N. 2009. Neural bases of autobiographical support for episodic recollection of faces. *Hippocampus*. doi 10.1002/hipo.20556.
- Yue X, Vessel EA, Biederman I. 2007. The neural basis of scene preferences. *Neuroreport*. 18:525-529.