Distribution Agreement

Andrew Persichetti

Date

Perceived egocentric distance sensitivity and invariance across scene-selective cortex

By

Andrew S. Persichetti

Psychology

Daniel Dilks, Ph.D.
Advisor

Joseph Manns, Ph.D.
Committee Member

Phillip Wolff, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Perceived egocentric distance sensitivity and invariance across scene-selective cortex

By

Andrew S. Persichetti
B.A., Portland State University, 2010

Advisor: Daniel Dilks, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts in Psychology
2015

Abstract

Behavioral studies in many species and studies in robotics have demonstrated two sources of information critical for visually-guided navigation: sense (left-right) information and egocentric distance (proximal-distal) information. A recent fMRI study found sensitivity to sense information in two scene-selective cortical regions, the retrosplenial complex (RSC) and the occipital place area (OPA), consistent with hypotheses that these regions play a role in human navigation. Surprisingly, however, another scene-selective region, the parahippocampal place area (PPA), was not sensitive to sense information, challenging hypotheses that this region is directly involved in navigation. Here we conduct a stronger test of the involvement of these regions in navigation, by investigating how these regions encode egocentric distance information (e.g., a house seen from close up versus far away), another type of information crucial for navigation. Using fMRI adaptation in human adults, we found no sensitivity to egocentric distance information in PPA, while RSC and OPA showed sensitivity to such information. These findings further support that RSC and OPA may be directly involved in navigation, while PPA is not, and suggest that scenes are processed by distinct pathways guiding navigation and recognition.

Perceived egocentric distance sensitivity and invariance across scene-selective cortex

By

Andrew S. Persichetti
B.A., Portland State University, 2010

Advisor: Daniel Dilks, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts in Psychology
2015

Table of Contents

Figures

INTRODUCTION

The navigability of a scene is completely different when mirror reversed (e.g., walking through a cluttered room to exit a door either on the left or right), or when viewed from a proximal or distal perspective (e.g., walking to a house that is either 50 feet or 500 feet in front of you). Indeed, behavioral evidence across many different species [including insects (Wehner et al. 1996), fish (Sovrano et al. 2002), pigeons (Gray et al. 2004), rats (Cheng 1986), rhesus monkeys (Gouteux et al. 2001), and humans (Hermer and Spelke 1994; Fajen and Warren 2003)] has demonstrated that both sense (left-right) and egocentric distance (proximal-distal) information are used for navigating the local environment. Similarly, studies in robotics highlight the necessity of sense and egocentric distance information for successful visually-guided navigation (Schöner et al. 1995). The term navigation has been defined by the above studies and many other reports as a process of relating one's egocentric system to fixed points in the world as one traverses the local environment (Gallistel, 1990; Wang and Spelke 2002). Here we use this standard definition of navigation.

A recent fMRI study (Dilks et al. 2011) found sensitivity to at least one of the two critical types of information guiding navigation (i.e., sense information) in two human scene-selective cortical regions, the retrosplenial complex (RSC) (Maguire 2001), and the occipital place area (OPA) (Dilks et al. 2013), also referred to as the transverse occipital sulcus (Grill-Spector 2003), consistent with hypotheses that these regions play a direct role in human navigation (Maguire, 2001; Epstein 2008; Dilks et al., 2011;); . By contrast, another scene-selective region, the parahippocampal place area (PPA) (Epstein and Kanwisher 1998), was not sensitive to sense information, challenging hypotheses that this region is directly involved in navigation (Epstein and Kanwisher 1998; Ghaem et al. 1997; Janzen and van Turennout 2004; Cheng and Newcombe 2005; Rosenbaum et al. 2004; Rauchs et al. 2008; Spelke et al. 2010). Here we conduct a stronger test of the involvement (or lack thereof) of the three scene-selective regions in navigation, by investigating how these regions encode egocentric distance information. We predict that if PPA is not directly involved in navigation, as suggested by its insensitivity to sense information, then this region will also be insensitive to egocentric distance information. Further, if RSC or OPA play a direct role in navigation, as suggested by their sensitivity to sense information, these regions will also be sensitive to egocentric distance information.

To test our predictions, we used an event-related fMRI adaptation paradigm (Grill-Spector and Malach 2001) in human adults. Participants viewed trials consisting of two successively presented images of either scenes or objects. Each pair of images consisted of one of the following: (1) the same image presented twice; (2) two completely different images; or (3) an image viewed from either a proximal or distal perspective followed by the opposite version of the same stimulus. If scene representations in scene-selective cortex are not sensitive to egocentric distance information, then images of the same scene viewed from proximal and distal perspectives will be treated as the same image, and the neural activity in scene-selective cortex will show adaptation across egocentric distance changes. On the other hand, if scene representations are sensitive to egocentric distance information, then images of the same scene viewed from proximal and distal perspectives will be treated as different images, producing no adaptation across distance changes in scene-selective cortex. We examined the representation of egocentric distance information in the three known scene-selective regions (PPA, RSC, and OPA) in human cortex.

METHODS

Participants

Thirty healthy individuals (ages 18–54; 17 females) were recruited for the experiment. All participants gave informed consent. All had normal or corrected to normal vision. One participant was excluded for excessive motion, and another participant did not complete the scan due to claustrophobia. We excluded a further three participants from the main analyses due to an inability to measure the expected fMRI adaptation effect in *any* scene-selective cortical region when viewing identical and completely different images of scenes (i.e., Different > Same). Thus, we report the results from 25 participants. However, when we analyzed the data including the three participants who did not exhibit the normal fMRI adaptation effect, the overall results did not change.

Design

We localized scene-selective regions of interest (ROIs) and then used an independent set of data to investigate the response of these regions to pairs of scenes or objects that were identical, different, or varied in their perceived egocentric distance. For the localizer scans, we used a standard method described previously to identify ROIs (Epstein and Kanwisher 1998). Specifically, a blocked design was used in which participants viewed images of faces, objects, scenes, and scrambled objects. Each participant completed 3 runs. Each run was 336 s long and consisted of 4 blocks per stimulus category. The order of the stimulus category blocks in each run was palindromic (e.g., faces, objects, scenes, scrambled objects, scrambled objects, scenes, objects, faces) and was randomized across runs. Each block contained 20 images from the same category for a total of 16 s blocks. Each image was presented for 300 ms, followed by a 500 ms interstimulus interval (ISI). We also included five 16 s fixation blocks: one at the beginning, three in the middle interleaved between each palindrome, and one at the end of each run. Participants performed a one-back task, responding every time the same image was presented twice in a row.

For the experimental scans, participants completed 8 runs each with 96 experimental trials, and an average of 47 fixation trials, used as a baseline condition. Each run was 397 s long. On each fixation trial, a white fixation cross (subtending 1° of visual angle) was displayed on a gray background. On each non-fixation trial, an image of either a scene or an object was presented for 300 ms, followed by an ISI of 400 ms and then by another image of the same stimulus category presented for 300 ms – following the method of Kourtzi and Kanwisher (2001) and many subsequent papers. After presentation of the second image, there was a jittered interval of ~3 s (ranging from 1 to 6 s) before the next trial began. Each pair of images consisted of one of the following: (1) the same image presented twice (Same condition); (2) two completely different images (Different condition); or (3) an image viewed from either a proximal or distal perspective followed by the opposite perspective of that same image (Distance condition) (Fig. 1A). In the Distance condition there were equal numbers of trials in which a proximal image preceded a distal image, and vice versa. This aspect of the experimental design is important because it allowed us to test whether the effects we measured in our experiment were indeed due to changes in perceived egocentric distance, and not due to 'boundary extension' (Intraub and Richardson 1989), which is discussed in more detail in the Results section. Scene and object trials were intermixed within a run. Trial sequence was generated using the Free-Surfer optseq2 function, optimized for the most accurate estimations of hemodynamic response (Burock et al. 1998; Dale et al. 1999). The images used as stimuli were photographs of 20 different scenes (10 indoor, 10 outdoor) from both a proximal and distal perspective. Thus, there were 40 different images of scenes. The distance differences varied across stimuli, ranging from ~20 feet to ~100 feet. Similarly, we included 40 images of objects viewed from both proximal and distal perspectives against backgrounds of varying textures to test the specificity of distance information in the scene-selective regions. In total, each subject viewed 128 trials of each condition (Same, Different, Distance). All stimuli were grayscale and 7° x 9° in size. Subjects were instructed to remain fixated on a white fixation cross that was presented on the screen in between stimuli. Each image was presented at the central fixation and then moved either slightly left or right. Participants performed an orthogonal task (not related to whether it was proximal or distal, or whether an image was a scene or object), responding via button box whether images in a pair were moving in the same or opposite direction. The motion task was also particularly chosen to eliminate any early retinotopic confounds.

fMRI scanning

Scanning was done on a 3T Siemens Trio scanner at the Facility for Education and Research in Neuroscience (FERN) at Emory University (Atlanta, GA). Functional images were acquired using a 32-channel head matrix coil and a gradient echo single-shot echo planar imaging sequence. Sixteen slices were acquired for both the localizer scans (repetition time = 2 s), and the experimental scans (repetition time = 1 s). For all scans: echo time = 30 ms; voxel size = 3.1 x 3.1 x 4.0 mm with a 0.4 mm interslice gap; and slices were oriented approximately between perpendicular and parallel to the calcarine sulcus, covering the occipital and temporal lobes. Whole-brain, high-resolution T1 weighted anatomical images were also acquired for each participant for anatomical localization.

Data analysis

fMRI data analysis was conducted using the FSL software (Smith et al. 2004) and custom MATLAB code. Before statistical analysis, images were skull-stripped (Smith 2002), and registered to the subjects' T1 weighted anatomical image(Jenkinson et al. 2002). Additionally, localizer data, but not experimental data, were spatially smoothed (6 mm kernel), detrended, and fit using a double-gamma function. After preprocessing, scene-selective regions PPA, RSC, and OPA were bilaterally defined in each participant (using data from the independent localizer scans) as those regions that responded more strongly to scenes than objects ($p < 10^{-4}$, uncorrected), as described previously (Epstein and Kanwisher 1998) (Fig. 1B). PPA and OPA were identified bilaterally in all 25 participants, while RSC was identified in the right hemisphere in all 25 participants, and in the left hemisphere in 24 participants. As a control region, we also functionally defined a bilateral foveal cortex (FC) ROI—the region of cortex responding to foveal stimulation (Dougherty et al. 2003). Specifically, the FC ROI was bilaterally defined in each of the 25 participants (using data from the localizer scans) as the regions that responded more strongly to scrambled objects than to intact objects ($p < 10^{-6}$, uncorrected), as described previously (MacEvoy and Yang 2012; Linsley and MacEvoy 2014). For each ROI of each participant, the mean time courses (percentage signal change relative to a baseline fixation) for the experimental conditions were extracted across voxels. Next, for each ROI, the time courses across conditions and participants were averaged together to identify the peak response (i.e., 5 s after the trial onset for OPA and PPA, and 6 s for RSC). (Note that such interregional variability in the shape of the hemodynamic response has been well documented – e.g., Aguirre et al. 1998; Miezin et al. 2000; Handwerker et al. 2004.) In a second analysis, for each ROI and each condition, the time courses across participants were averaged together to identify an average response across time points, not simply the peak response. More specifically, we identified the first time point to exhibit the expected adaptation effect (i.e., Different > Same) to the last time point to exhibit adaptation. We determined which time points showed the expected adaptation effect by conducting a paired t-test between the Different and Same conditions at each time point. We found that the Different condition was significantly greater than the Same condition at all time points from 5 s to 9 s after trial onset (all p values < 0.02). Conversely, none of the time points before 5 s or after 9 s showed this adaptation effect (all p values > 0.30). Finally, for each participant, both the peak and average responses for each scene-selective region were extracted for each condition (Different, Distance, Same), and repeated-measures ANOVAs were performed on each.

A 2 (hemisphere : Left, Right) x 3 (condition: Different, Distance, Same) repeated-measures ANOVA for each ROI was conducted. We found no significant hemisphere x condition interaction in any of the scene-selective ROIs at the peak response (all p values > 0.50), or the average response (all p values > 0.14). Thus, both hemispheres were collapsed for further analyses.

RESULTS

As predicted, PPA was not sensitive to egocentric distance information in images of scenes. A 3 level (condition: Different, Distance, Same) repeated-measures ANOVA on the peak response (i.e., 5 s – identified by averaging the time courses for the ROI across conditions and participants) revealed a significant main effect of condition ($F_{(2,48)} = 15.42$, $p < 0.001$, $\eta_P^2 = 0.39$), with a significantly greater response to the Different condition compared to either the Same or Distance conditions (main effect contrasts, both p values < 0.001, d's > 0.90), and no significant difference between the Distance and Same conditions (main effects contrast, p = 0.46, d = 0.14) (Fig. 2A). These findings not only demonstrate the expected fMRI adaptation effect (Different > Same), but also the predicted perceived egocentric distance adaptation effect (i.e., no significant difference between Distance and Same), revealing that PPA treats images of scenes viewed from different egocentric distance perspectives as similar to identical image pairs, indicating insensitivity to egocentric distance information in scenes.

Our second planned analysis, on the average response from the first time point to exhibit adaptation to the last time to exhibit adapation (i.e., from 5 s to 9 s) (see Methods for details), not only at the peak response, also revealed insensitivity to perceived egocentric distance in PPA. A 3 level (condition: Different, Distance, Same) repeated-measures ANOVA on this average response also revealed a significant main effect of condition ($F_{(2,48)} = 12.15$, $p < 0.001$, $\eta_P^2 = 0.34$), with a significantly greater response to the Different condition compared to either the Same or Distance conditions (main effect contrasts, both p values < 0.001, both d's > 0.80), and no significant difference between the Distance and Same conditions (main effects contrast, $p = 0.81$, $d = 0.06$).

By contrast, we found that RSC and OPA were sensitive to egocentric distance information in images of scenes. For RSC, a 3 level (condition: Different, Distance, Same) repeated-measures ANOVA at the peak response (i.e., 6 s) revealed a significant main effect of condition ($F_{(2,48)} = 9.43$, $p < 0.001$, $\eta_P^2 = 0.28$), with a significantly greater response to the Different condition compared to the Same condition (main effect contrast, $p < 0.001$, $d = 0.96$) – the expected fMRI adaptation effect – and a significant difference between the Distance and Same conditions (main effect contrast, $p < 0.05$, $d = 0.51$), demonstrating sensitivity to egocentric distance information in RSC (Fig. 2B). There was a marginally significant difference between the Distance and Different conditions (main effect contrast, $p = 0.05$, $d = 0.50$). Likewise, a 3 level (condition: Different, Distance, Same) repeated-measures

ANOVA on the average response (i.e., from 5 s to 9 s) revealed a significant main effect of condition ($F_{(2,48)} = 6.97$, $p < 0.002$, $\eta_P^2 = 0.23$), with a significantly greater response to the Different condition compared to the Same condition (main effect contrast, $p < 0.001$, $d = 1.06$), and a marginally significant difference between the Distance and Same conditions (main effect contrast, $p = 0.07$, $d = 0.48$). There was no significant difference between the Distance and Different conditions (main effect contrast, $p = 0.11$, $d = 0.50$). These results demonstrate the expected fMRI adaptation effect (i.e., Different > Same) in RSC, but no adaptation across perceived egocentric distance (i.e., Distance > Same), revealing that RSC represents images of scenes viewed from either a proximal or distal perspective as two different images.

Similarly, for OPA, a 3 level (condition: Different, Distance, Same) repeated measures ANOVA on the peak response (i.e., 5 s) revealed a significant main effect of condition ($F_{(2,48)} = 5.14$, $p < 0.01$, $\eta_P^2 = 0.18$), with a significantly greater response to the Different condition than to the Same condition (main effect contrast, $p < 0.002$, $d = 0.50$). But, like RSC, there was not a significant difference in the response to the Distance condition compared to either the Different or the Same condition (main effect contrasts, both p values = 0.11, both d's < 0.35), revealing an ambiguous result. However, a 3 level (condition: Different, Distance, Same) repeated measures ANOVA on the average response (i.e., from 5 s to 9 s) revealed a significant main effect of condition ($F_{(2,48)} = 8.95$, $p < 0.001$, $\eta_P^2 = 0.27$), with a significantly greater response to the Different condition than to both the Same and Distance conditions (main effect contrasts, both p values < 0.05, d's > 0.50), and a marginally significant difference between the Distance and Same conditions (main effect contrast, p =

0.07) (Fig. 2C). These results demonstrate the expected fMRI adaptation effect (i.e., Different > Same) in OPA, but no adaptation across perceived egocentric distance (i.e., Distance > Same), revealing that OPA, like RSC, represents images of scenes viewed from either a proximal or distal perspective as two different images.

The above analyses suggest that the three scene-selective regions encode egocentric distance information in images of scenes differently, so we directly tested this suggestion by comparing the differences in response across the three ROIs. Specifically, for each ROI the difference between the peak responses for two different images of scenes and the same images (i.e., expected adaptation) was compared to the difference between the peak responses for proximal vs. distal images and the same images (Fig. 3). Crucially, a 3 (ROI: OPA, RSC, PPA) x 2 (difference score: Different-Same, Distance-Same) repeated-measures ANOVA revealed a significant interaction ($F_{(2,48)} = 3.66$, $p < 0.04$, $\eta_P^2 = 0.13$), with a significantly greater difference between the Different and Same conditions than between the Distance and Same conditions for PPA, relative to both RSC and OPA (interaction contrasts, both p values < 0.04, both $\eta_P^2 > 0.16$). There was not a significant difference in the responses between the RSC and OPA (interaction contrast, $p = 0.56$, $\eta_P^2 = 0.01$). Likewise, we found a significant interaction when we compared the differences in response across the three ROIs on the average response ($F_{(2,48)} = 3.76$, $p < 0.03$, $\eta_P^2 = 0.14$), with a significantly greater difference between the Different and Same conditions than between the Distance and Same conditions for PPA, relative to RSC (interaction contrast, $p < 0.01$, $\eta_P^2 = 0.28$), a marginally significant difference between PPA and OPA (interaction contrast, p value = 0.06,

$\eta_P^2 = 0.14$), and no significant difference in the responses between RSC and OPA (interaction contrast, $p = 0.70$, $\eta_P^2 < 0.01$). These results show that the scene selective regions encode egocentric distance information differently: PPA is not sensitive to egocentric distance information in scenes, while RSC and OPA are sensitive to such information.

Given that our stimuli included objects as well as scenes, we were also able to investigate how OPA, PPA, and RSC might respond to changes in egocentric distance information in images of objects (the nonpreferred category). We found that none of the responses within scene-selective regions exhibit the expected adaptation effect (i.e., Different > Same) to object stimuli (all p values > 0.05; RSC showed a greater response to the Same condition than the Different condition). Thus, the question of sensitivity to egocentric distance information in objects for scene-selective regions is moot.

Might it be the case that the sensitivity to egocentric distance information in images of scenes in OPA or RSC is simply due to a feed-forward effect from earlier visual areas, rather than characteristic of egocentric distance sensitivity to scenes in particular? While we do not think this could be the case (because participants were asked to fixate, and thus the stimuli were moving across the fovea), we directly addressed this question by comparing both the peak response and the average response to the three conditions in an independently defined region of cortex representing the fovea, and found that 'foveal cortex' did not even show fMRI adaptation for Different versus Same scenes (all p values > 0.18), thus confirming that neither OPA nor RSC's sensitivity to egocentric distance information in scenes is due to adaptation in early visual cortex.

Might it be the case that the invariance to perceived egocentric distance in images of scenes in PPA is due to ‘boundary extension’, instead of actual insensitivity to egocentric distance information? Boundary extension is a process in which people, when asked to remember a photograph of a scene, remember a more expansive view than was shown in the original photograph. Thus, the representation of the scene extends beyond the pictures boundaries, particularly when the view is close up (Intraub and Richardson 1989). Consistent with these behavioral data, Park and colleagues (2007), using an fMRI adaptation paradigm, found adaptation in PPA when a wide-view of a scene followed a close-up view of a scene, but not when the wide view preceded the close-up view, and concluded that PPA is involved in boundary extension. If the effect in our study can be explained by boundary extension, then we should see the same pattern of results as reported by Park and colleagues. To test this hypothesis, we divided the Distance trials in half: one half was made up of trials in which a wide-view of a scene followed a close-up view of a scene, and the other half was the reverse condition. If we are observing a boundary extension effect in PPA, then adaptation should be greater when a wide-view of a scene is followed by a close-up view than on the reverse condition. A 4 level (condition: Different, Wide-Close, Close-Wide, Same) repeated measures ANOVA on either the peak response or the average response revealed a significant main effect of condition (both p values < 0.01), with a significantly greater response to the Different versus Same condition ($p < 0.01$), demonstrating the expected fMRI adaptation. Crucially, however, there were no significant differences between the

Wide-Close, Close-Wide, or Same conditions (p values > 0.30). Thus, these results confirm that the effects found in our study are due to insensitivity to egocentric distance information in images of scenes in PPA, rather than to boundary extension. The reason for this conflicting result is not entirely clear, but could be due to differences between the two studies with respect to i) the level of processing (i.e., perception versus memory: boundary extension does not occur while sensory information is present, as is the case in this study, but rather involves distortion of the scene representation over time), ii) task demands (i.e., in our study participants were performing an orthogonal task, while in Park et al.'s study participants were asked to memorize the layout and overall details of the scene), or iii) the definition of the PPA (we defined the PPA using the contrast scenes versus objects, while Park et al. defined the PPA using the contrast scenes versus faces).

DISCUSSION

The current study asked how scene-selective regions encode egocentric distance information. As predicted, the results demonstrate that egocentric distance information is represented differentially across scene-selective cortex. Specifically, using an fMRI adaptation paradigm we found that two scene-selective regions (i.e., RSC and OPA) were sensitive to egocentric distance information, while the PPA, another scene-selective region, was not sensitive to such information. These results are specific to images of scenes, not to images of objects, and cannot be explained by a feed-forward effect from earlier visual areas, or by 'boundary extension.'

Our finding of insensitivity to egocentric distance information in scene-selective PPA provides further evidence against the hypothesis that the PPA is directly involved in navigation (Epstein and Kanwisher 1998; Ghaem et al. 1997; Janzen and van Turennout 2004; Cheng and Newcombe 2005; Rosenbaum et al. 2004; Rauchs et al. 2008; Spelke et al. 2010). Recall that Dilks and colleagues (2011) also found that PPA was not sensitive to sense information, another type of information crucial for navigation. While navigation is no doubt crucial to our successful functioning (e.g., walking to the market or even getting around one's own house), it is reasonable to argue that the ability to recognize a scene as belonging to a specific category (e.g., kitchen, beach, or city) also plays a necessary role in one's everyday life. After all, our ability to categorize a scene makes it possible to know what to expect from, and how to behave in, different environments (Bar 2004). Taken together, these arguments support the necessity of both navigation and scene categorization systems, and the current data suggest that visual scene processing may not serve a single purpose (i.e., for navigation), but rather has multiple purposes guiding us not only through our environments, but also guiding our behaviors within them. If our two-streams-for-scene-processing hypothesis is correct, then the PPA may contribute to the 'categorization stream', while the RSC, OPA, or both may contribute to the 'navigation stream'.

While our finding that PPA is not sensitive to egocentric distance information coupled with its insensitivity to sense information (Dilks et al., 2011) strongly suggest that PPA is not involved in navigation, ongoing studies are directly testing this hypothesis by correlating behavioral measures of navigation and categorization tasks to the fMRI signal in each scene-selective ROI. Further, the current study does not distinguish the precise roles of OPA and RSC in navigation. It is possible that OPA and RSC may both be involved in navigation more generally, but support different functions within navigation. Ongoing studies are investigating this possibility.

The hypothesis that the scene processing system may be divided into two streams might sound familiar. For example, Epstein (2008) proposed that human scene processing is divided into two streams – both serving the primary function of navigation – with PPA representing the local scene, and RSC supporting orientation within the broader environment. This hypothesis is quite different from what we propose here. While we agree that the primary role of the RSC is navigation, we disagree that the PPA shares this role. Rather, we hypothesize that the PPA is a part of a functionally distinct pathway devoted to scene recognition and categorization. Thus, our two-streams-for-scene-processing hypothesis is instead more like the two functionally distinct streams of visual object processing proposed by Goodale and Milner (1992), with one stream responsible for recognition, and another for visually-guided action. Note that our hypothesis of two distinct systems for human scene processing – a categorization stream including PPA, and a navigation stream including RSC and OPA – does not mean that the two systems cannot and do not interact. Indeed, two recent studies found functional correlations between the RSC and anterior portions of the PPA, and between the OPA and posterior PPA (Baldassano et al. 2013; Nasr et al. 2013), suggesting these two regions are functionally (and most likely anatomically) connected, thereby facilitating crosstalk between the two systems.

Findings from several other studies support our hypothesis that a scene-processing pathway devoted to the recognition of scene category (including PPA) exists separate from a pathway primarily involved in representing information necessary for navigation (including RSC and OPA), although most of the findings were never discussed in this way. For example, when the PPA is damaged due to stroke, patients have difficulty identifying places and landmarks, and report that their sense of a scene as a coherent whole has been lost (Aguirre and D'Esposito 1999; Barrash et al. 2000; Mendez and Cherrier 2003). A patient studied by Takahashi et al. (1989) was unable to recognize photographs of his home and neighborhood, but he could describe visual features such as patterns on a fence and the shape of a mailbox, as well as draw an accurate map of his home and neighborhood. Another patient with PPA damage was reported to have lost semantic knowledge of places and buildings (McCarthy et al. 1996). By contrast, patients with RSC damage are able to recognize landmarks, but cannot use these same landmarks to orient themselves in the environment, and often report that they are 'lost in space' (Bottini et al. 1990; Takahashi et al. 1997; Greene et al. 2006; Ino et al. 2007; Osawa et al. 2008). Further, MVPA studies have shown that while activity patterns in both PPA and RSC contain information about scene category (e.g., beaches, forests, highways), only the activation patterns in PPA, not RSC, are related to behavioral performance (Walther et al. 2011; Walther et al. 2009). Finally, a number of recent studies have shown that PPA, not RSC, is sensitive to object information, especially object information that might facilitate the categorization of a scene. For example, several studies found that PPA responds to i) objects that are good exemplars of specific scenes (e.g., a bed or a refrigerator) (MacEvoy and Epstein 2009; Harel et al. 2013), ii) objects that are strongly associated with a given context (e.g., a toaster) versus low ‘contextual’ objects (e.g., an apple) (Bar 2004; Bar and Aminoff 2003; Bar et al. 2008), and iii) objects that evoke a strong sense of the surrounding space (e.g., a bed or a couch) versus objects that do not (e.g., a small fan or a box) (Mullally and Maguire 2011). Taken together, the above findings are consistent with our hypothesis that scene processing may consist of two streams, one stream devoted to the recognition of scene category (including PPA), and another involved in navigation (including RSC and OPA).

It is well established that PPA responds to ‘spatial layout’, or the geometry of local space, initially based on evidence that this region responds significantly more strongly to images of sparse, empty rooms than to these same images when the walls, floors and ceilings have been fractured and rearranged (Epstein and Kanwisher 1998). At first glance, the idea that PPA encodes geometric information may seem contradictory to its involvement in the recognition of scene category. In fact, such spatial layout representation in PPA has even led to hypotheses that the PPA might be the neural locus for a ‘geometry module’ (Hermer and Spelke 1994), necessary for reorientation and navigation (Epstein and Kanwisher 1998). But spatial layout information need not be used for navigation only, and could also easily facilitate the recognition of scene category. Indeed, several behavioral and computer vision studies have found that scenes can be categorized based on their overall spatial layout (Walther et al. 2011; Oliva and Schyns 1997; Oliva and Torralba 2001; Greene and Oliva 2009). Thus, we propose that spatial layout information may be encoded by both a scene-processing pathway devoted to the recognition of scene category (including PPA) and a pathway primarily involved in representing information necessary for navigation (including RSC and OPA), but each stream either encodes different types of spatial layout information, or uses the same kind of spatial layout information for different purposes.

Finally, while we found that PPA is not sensitive to egocentric distance information in scenes – one type of viewpoint change (i.e., near or far to the viewer) – two other studies have shown that it is sensitive to another type of viewpoint change (i.e., left/right changes less than 180° relative to the viewer) (Epstein et al. 2003; Park and Chun 2009). Why the possible discrepancy between our finding and the Epstein et al. and Park and Chun findings? First, these different kinds of viewpoint changes might truly reflect different computations and hence different responses by PPA. Second, the same researchers found that PPA became insensitive to such viewpoint changes as participants became familiar with the scenes over the course of an experimental session (Epstein et al. 2005; Epstein and Higgins 2007), or strangely when the viewpoint change across panoramic scenes was not continuous (Park and Chun, 2009). Furthermore, another study found that PPA was insensitive to 180° viewpoint changes (mirror reversals) (Dilks et al., 2011), revealing insensitivity to yet another viewpoint change in PPA.

In conclusion, we have shown that PPA is not sensitive to egocentric distance information in images of scenes, while RSC and OPA are sensitive to such information. The findings from this study, along with those from Dilks et al. (2011), demonstrate that while the RSC and OPA are sensitive to both egocentric distance (proximal-distal) and sense (left-right) information, both critical for visually-guided navigation, the PPA is insensitive to both kinds of information, suggesting that the computations directly involved in visually-guided navigation do not occur in the PPA. These results support the hypothesis that there exist two distinct pathways for processing scenes: one for navigation that consists of RSC and OPA, and another for the recognition of scene category that includes PPA.

REFERENCES

Aguirre GK, D'Esposito M. 1999. Topographical disorientation: a synthesis and taxonomy. Brain. 122 (Pt 9):1613–1628.

Aguirre GK, Zarahn E, D'esposito M. 1998. The variability of human, BOLD hemodynamic responses. NeuroImage. 8:360–369.

Baldassano C, Beck DM, Fei-Fei L. 2013. Differential connectivity within the Parahippocampal Place Area. NeuroImage. 75:228–237.

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, Schacter DL. 2008. Scenes Unseen: The Parahippocampal Cortex Intrinsically Subserves Contextual Associations, Not Scenes or Places Per Se. J Neurosci. 28:8539–8544.

Barrash J, Damasio H, Adolphs R, Tranel D. 2000. The neuroanatomical correlates of route learning impairment. Neuropsychologia. 38:820–836.

Bottini G, Cappa S, Geminiani G, Sterzi R. 1990. Topographic disorientation--a case report. Neuropsychologia. 28:309–312.

Burock MA, Buckner RL, Woldorff MG, Rosen BR, Dale AM. 1998. Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. Neuroreport. 9:3737–3739.

Cheng K. 1986. A purely geometric module in the rat's spatial representation. Cognition. 23:149–178.

Cheng K, Newcombe NS. 2005. Is there a geometric module for spatial orientation? Squaring theory and evidence. Psychon Bull Rev. 12:1–23.

Dale AM, Greve DN, Burock MA. 1999. Stimulus sequences for event-related fMRI. Hum Brain Mapp. 8:109–114.

Dilks DD, Julian JB, Kubilius J, Spelke ES, Kanwisher N. 2011. Mirror-image sensitivity and invariance in object and scene processing pathways. J Neurosci Off J Soc Neurosci. 31:11305–11312.

Dilks DD, Julian JB, Paunov AM, Kanwisher N. 2013. The occipital place area is causally and selectively involved in scene perception. J Neurosci Off J Soc Neurosci. 33:1331–1336a.

Dougherty RF, Koch VM, Brewer AA, Fischer B, Modersitzki J, Wandell BA. 2003. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. J Vis. 3:586–598.

Epstein R, Graham KS, Downing PE. 2003. Viewpoint-specific scene representations in human parahippocampal cortex. Neuron. 37:865–876.

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

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, Higgins JS, Thompson-Schill SL. 2005. Learning places from views: variation in scene processing as a function of experience and navigational ability. J Cogn Neurosci. 17:73–83.

Fajen BR, Warren WH. 2003. Behavioral dynamics of steering, obstable avoidance, and route selection. J Exp Psychol Hum Percept Perform. 29:343–362.

Ghaem O, Mellet E, Crivello F, Tzourio N, Mazoyer B, Berthoz A, Denis M. 1997. Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. Neuroreport. 8:739–744.

Goodale MA, Milner AD. 1992. Separate visual pathways for perception and action. Trends Neurosci. 15:20–25.

Gouteux S, Thinus-Blanc C, Vauclair J. 2001. Rhesus monkeys use geometric and nongeometric information during a reorientation task. J Exp Psychol Gen. 130:505–519.

Gray ER, Spetch ML, Kelly DM, Nguyen A. 2004. Searching in the Center: Pigeons (Columba livid) Encode Relative Distance From Walls of an Enclosure. J Comp Psychol. 118:113–117.

Greene KK, Donders J, Thoits T. 2006. Topographical heading disorientation: a case study. Appl Neuropsychol. 13:269–274.

Greene M, Oliva A. 2009. Recognition of natural scenes from global properties: Seeing the forest without representing the trees. Cognit Psychol. 58:137–176.

Grill-Spector K. 2003. The neural basis of object perception. Curr Opin Neurobiol. 13:159–166.

Grill-Spector K, Malach R. 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol Amst. 107:293–321.

Handwerker DA, Ollinger JM, D'Esposito M. 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. NeuroImage. 21:1639–1651.

Harel A, Kravitz DJ, Baker CI. 2013. Deconstructing visual scenes in cortex: gradients of object and spatial layout information. Cereb Cortex N Y N 1991. 23:947–957.

Hermer L, Spelke ES. 1994. A geometric process for spatial reorientation in young children. Nature. 370:57–59.

Intraub H, Richardson M. 1989. Wide-angle memories of close-up scenes. J Exp Psychol Learn Mem Cogn. 15:179–187.

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

Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. NeuroImage. 17:825–841.

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

Linsley D, MacEvoy SP. 2014. Evidence for participation by object-selective visual cortex in scene category judgments. J Vis. 14:19–19.

MacEvoy SP, Epstein RA. 2009. Decoding the Representation of Multiple Simultaneous Objects in Human Occipitotemporal Cortex. Curr Biol. 19:943–947.

MacEvoy SP, Yang Z. 2012. Joint neuronal tuning for object form and position in the human lateral occipital complex. NeuroImage. 63:1901–1908.

Maguire EA. 2001. The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. Scand J Psychol. 42:225–238.

McCarthy RA, Evans JJ, Hodges JR. 1996. Topographic amnesia: spatial memory disorder, perceptual dysfunction, or category specific semantic memory impairment? J Neurol Neurosurg Psychiatry. 60:318–325.

Mendez MF, Cherrier MM. 2003. Agnosia for scenes in topographagnosia. Neuropsychologia. 41:1387–1395.

Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. 2000. Characterizing the Hemodynamic Response: Effects of Presentation Rate, Sampling Procedure, and the Possibility of Ordering Brain Activity Based on Relative Timing. NeuroImage. 11:735–759.

Mullally SL, Maguire EA. 2011. A new role for the parahippocampal cortex in representing space. J Neurosci Off J Soc Neurosci. 31:7441–7449.

Nasr S, Devaney KJ, Tootell RBH. 2013. Spatial encoding and underlying circuitry in scene-selective cortex. NeuroImage. 83:892–900.

Oliva A, Schyns PG. 1997. Coarse Blobs or Fine Edges? Evidence That Information Diagnosticity Changes the Perception of Complex Visual Stimuli. Cognit Psychol. 34:72–107.

Oliva A, Torralba A. 2001. Modeling the shape of a scene: a holistic representation of the spatial envelope. Int J Comput Vis. 42:145–175.

Osawa A, Maeshima S, Kunishio K. 2008. Topographic disorientation and amnesia due to cerebral hemorrhage in the left retrosplenial region. Eur Neurol. 59:79–82.

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

Park S, Intraub H, Yi D-J, Widders D, Chun MM. 2007. Beyond the Edges of a View: Boundary Extension in Human Scene-Selective Visual Cortex. Neuron. 54:335–342.

Rauchs G, Orban P, Balteau E, Schmidt C, Degueldre C, Luxen A, Maquet P, Peigneux P. 2008. Partially segregated neural networks for spatial and contextual memory in virtual navigation. Hippocampus. 18:503–518.

Rosenbaum RS, Ziegler M, Winocur G, Grady CL, Moscovitch M. 2004. "I have often walked down this street before": fMRI studies on the hippocampus and other structures during mental navigation of an old environment. Hippocampus. 14:826–835.

Schöner G, Dose M, Engels C. 1995. Dynamics of behavior: Theory and applications for autonomous robot architectures. Robot Auton Syst. 16:213–245.

Smith SM. 2002. Fast robust automated brain extraction. Hum Brain Mapp. 17:143–155.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage. 23 Suppl 1:S208–219.

Sovrano VA, Bisazza A, Vallortigara G. 2002. Modularity and spatial reorientation in a simple mind: encoding of geometric and nongeometric properties of a spatial environment by fish. Cognition. 85:B51–B59.

Spelke E, Lee SA, Izard V. 2010. Beyond core knowledge: Natural geometry. Cogn Sci. 34:863–884.

Takahashi N, Kawamura M, Shiota J, Kasahata N, Hirayama K. 1997. Pure topographic disorientation due to right retrosplenial lesion. Neurology. 49:464–469.

Walther DB, Caddigan E, Fei-Fei L, Beck DM. 2009. Natural Scene Categories Revealed in Distributed Patterns of Activity in the Human Brain. J Neurosci. 29:10573–10581.

Walther DB, Chai B, Caddigan E, Beck DM, Fei-Fei L. 2011. Simple line drawings suffice for functional MRI decoding of natural scene categories. Proc Natl Acad Sci U S A. 108:9661–9666.

Wang RF, Spelke ES. 2002. Human spatial representation: insights from animals. Trends Cogn Sci. 6:376–382.

Wehner, Michel, Antonsen. 1996. Visual navigation in insects: coupling of egocentric and geocentric information. J Exp Biol. 199:129–140.

FIGURE LEGENDS

Figure 1. Example stimuli from each condition (i.e., Same, Distance, Different) (**A**), and scene-selective cortical regions from an example participant (**B**). Using independent data, PPA, RSC, and OPA (shown in white) were localized as regions that responded more strongly to scenes than objects ($p < 10^{-4}$).

Figure 2. Hemodynamic time courses (percentage signal change) of three scene selective regions of cortex, PPA (**A**), RSC (**B**), and OPA (**C**) to (1) two completely different images of scenes (red line labeled "Different"), (2) the same image of a scene presented twice (blue line labeled "Same"), and (3) an image of a scene viewed from either a proximal or distal perspective followed by the opposite version of the same stimulus (green line labeled "Distance"). Note invariance to egocentric distance information in PPA, but sensitivity to such information in both RSC and OPA.

Figure 3. For each scene-selective ROI, the difference between the peak responses for two different images of scenes and the same images (labeled "Different-Same") was compared to the difference between the peak responses for two images of the same scene viewed from either a proximal or distal perspective and the opposite version of the same image and the same images (labeled "Distance-Same"). A 3 (ROI: OPA, RSC, PPA) x 2 (difference score: Different-Same, Distance-Same) repeated-measures ANOVA revealed a significant interaction ($F_{(2,48)} = 3.66$, $p < 0.04$, $\eta_P^2 = 0.13$), with a significantly greater difference between the Different and Same

conditions than between the Distance and Same conditions for PPA, relative to RSC or OPA. This result suggests that egocentric distance information is represented differently by the scene-selective regions: PPA is not sensitive to egocentric distance information in scenes, while RSC and OPA are sensitive to such information.

Figure 1.

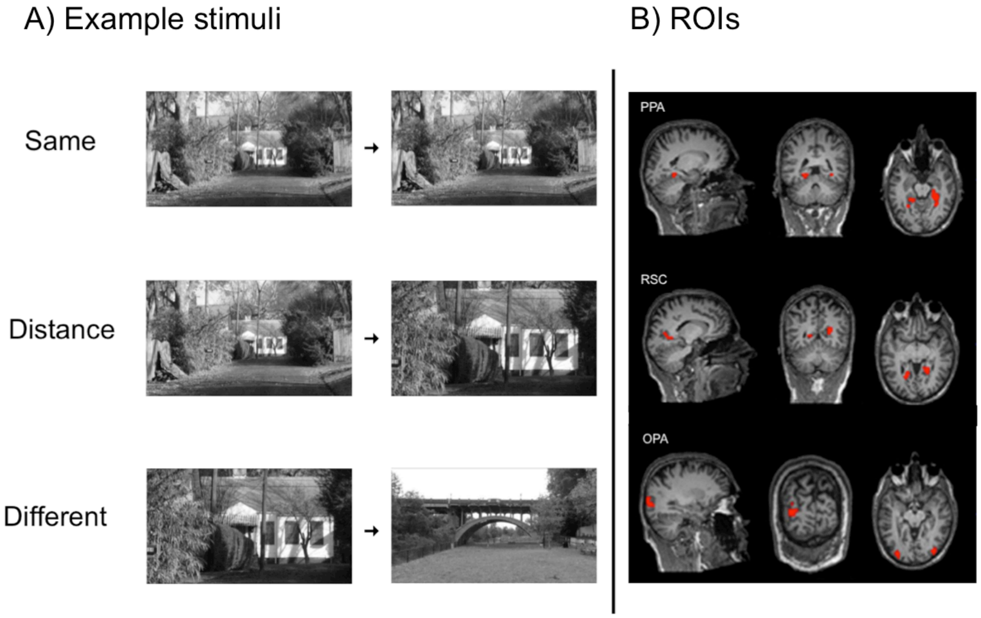
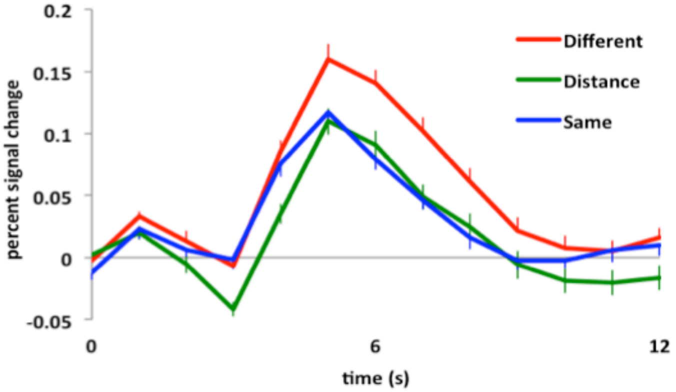
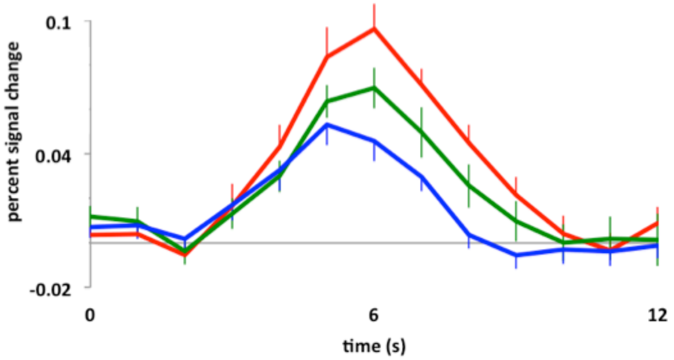

A) Example stimuli
B) ROIs
Same
Distance
Different
PPA
RSC
OPA

Figure 2.

A) PPA



B) RSC



C) OPA

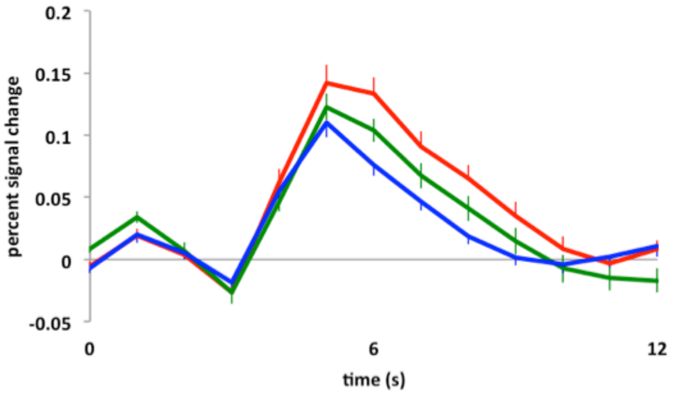
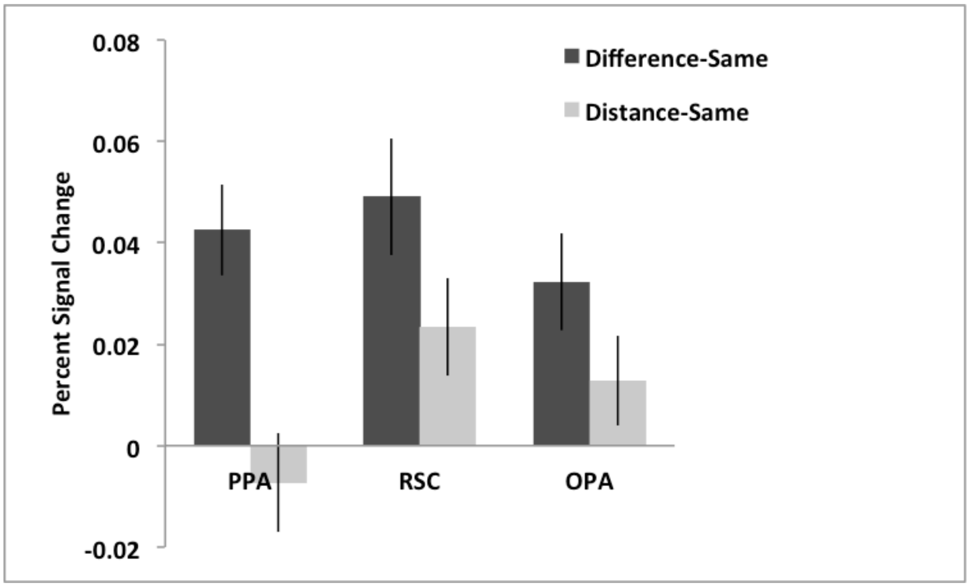


Figure 3.


0.08
0.06
0.04
0.02
0
-0.02
Percent Signal Change
■ Difference-Same
■ Distance-Same
PPA
RSC
OPA