Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Perceived egocentric distance sensitivity and invariance across scene-selective cortex

Andrew S. Persichetti and Daniel D. Dilks *

Department of Psychology, Emory University, Atlanta, GA, USA

ARTICLE INFO

Article history:
Received 12 August 2015
Reviewed 21 September 2015
Revised 2 November 2015
Accepted 8 February 2016
Action editor Gereon Fink
Published online 17 February 2016

Keywords:
Scene recognition
Navigation
Occipital place area (OPA)
Parahippocampal place area (PPA)
Retrosplenial complex (RSC)

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 examined 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 and a regions-of-interest analysis approach in human adults, we found sensitivity to egocentric distance information in RSC and OPA, while PPA was not sensitive to such information. These findings further support that RSC and OPA are directly involved in navigation, while PPA is not, consistent with the hypothesis that scenes may be processed by distinct systems guiding navigation and recognition.

1. 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 has demonstrated that both sense (left-right) and egocentric distance (proximal-distal) information are used in navigation by insects (Wehner, Michel, & Antonsen, 1996), fish (Sovrano, Bisazza, & Vallortigara, 2002), pigeons (Gray, Spetch, Kelly, & Nguyen, 2004), rats (Cheng, 1986), rhesus monkeys (Gouteux, Thinus-Blanc, & Vauclair, 2001), and humans (Fajen & Warren, 2003; Hermer & Spelke, 1994). Similarly, studies in robotics highlight the necessity of sense and egocentric distance information for successful visually-guided navigation (Schöner, Dose, & Engels, 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 environment (Gallistel, 1990; Wang & Spelke, 2002). Here we use this standard definition of navigation.

* *Corresponding author*. Department of Psychology, Emory University, Atlanta, GA 30309, USA.
E-mail address: dilks@emory.edu (D.D. Dilks).

http://dx.doi.org/10.1016/j.cortex.2016.02.006

A recent fMRI adaptation study (Dilks, Julian, Kubilius, Spelke, & Kanwisher, 2011) found sensitivity to 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, Julian, Paunov, & Kanwisher, 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 (Dilks et al., 2011; Epstein, 2008; Maguire, 2001). By contrast, another scene-selective region, the parahippocampal place area (PPA) (Epstein & Kanwisher, 1998), was not sensitive to sense information, challenging hypotheses that this region is directly involved in navigation (Cheng & Newcombe, 2005; Epstein & Kanwisher, 1998; Ghaem et al. 1997; Janzen & van Turennout, 2004; Rauchs et al. 2008; Rosenbaum, Ziegler, Winocur, Grady, & Moscovitch, 2004; Spelke, Lee, & Izard, 2010). Here we investigate 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. Given that RSC and OPA are sensitive to sense information – one type of information that is crucial for navigation – we predict that these regions will also be sensitive to egocentric distance information. By contrast, since PPA is not sensitive to sense information, we predict that this region will also not be sensitive to egocentric distance information.

To test our predictions, we used an event-related fMRI adaptation paradigm (Grill-Spector & 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 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. On the other hand, if scene representations 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. We examined the representation of egocentric distance information in the three known scene-selective regions (PPA, RSC, and OPA) in human cortex.

2. Methods

2.1. *Participants*

Thirty healthy individuals (ages 18–54; 17 females; 26 right handed) 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. Thus, we report the results from 28 participants.

2.2. *Design*

We localized scene-selective regions of interest (ROIs) and then used an independent set of data to investigate the responses 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 & 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 sec 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 sec blocks. Each image was presented for 300 msec, followed by a 500 msec interstimulus interval (ISI). We also included five 16 sec 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 (48 ‘scene’ trials and 48 ‘object’ trials, intermixed), and an average of 47 fixation trials, used as a baseline condition. Each run was 397 sec long. On each fixation trial, a white fixation cross (subtending .5° 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 msec, followed by an ISI of 400 msec and then by another image of the same stimulus category presented for 300 msec – 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 sec (ranging from 1 to 6 sec) 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 total, each subject viewed 128 trials of each condition (Same, Different, Distance). Note, in the Distance condition, we were careful to manipulate only perceived egocentric distance information, while not changing the angle from which the scenes were viewed. To ensure that viewing angle did not change between the Distance conditions in our stimuli, we first identified the same point in both the proximal and distal perspectives of each image (e.g., a window) and measured its distance (in pixels) away from two other points (to the right and left) in each image (e.g., fence posts). Next, we calculated the ratio of the distance from the central point and the point on the left to the distance between the central point and the point on the right, and finally, compared the ratios between the two perspectives. We found no difference in viewing angle between the near and far images of scenes [mean ratio: near = 2.35, far = 2.33; $t_{(9)} = .25$, $p = .81$]. Further, 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

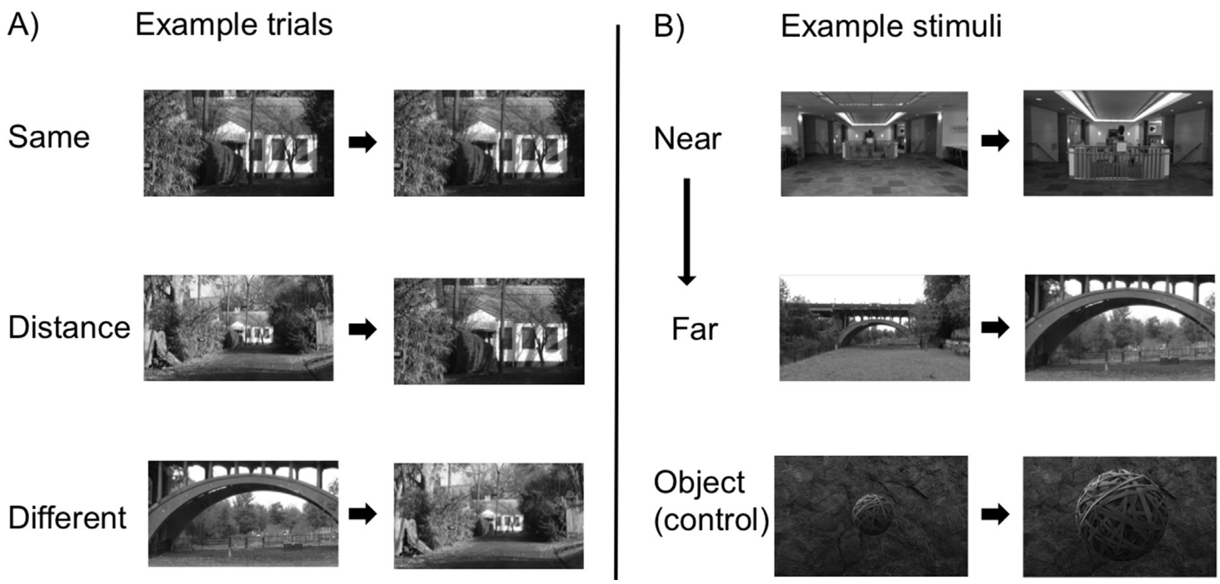

Fig. 1 – (A) Example trials from each condition (i.e., Same, Distance, Different). (B) Example stimuli from the Distance condition, ranging from Near to Far. Two independent sets of 25 participants each rated the stimuli as either near or far to ensure that our stimuli spanned a wide range of distances. The top row of panel B shows the scene pair that participants rated as the nearest distance change, and the middle row as the farthest distance change. The bottom row is an example of an object trial that depicts the object from both proximal and distal perspectives against a background texture that provided depth cues, ensuring that participants perceived the objects as proximal and distal, and not simply as small or big.

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 i) angle changes, ii) the potential perception of navigating through the scene, which might be perceived if every trial consisted of a proximal image preceding a distal image, or iii) 'boundary extension' (Intraub & Richardson, 1989), which is discussed in more detail in the Results sections. Trial sequence was generated using the FreeSurfer optseq2 function, optimized for the most accurate estimations of hemodynamic response (Burock, Buckner, Woldorff, Rosen, & Dale, 1998; Dale, Greve, & Burock, 1999). The images used as stimuli were photographs of 10 different scenes (5 indoor, 5 outdoor) from both a proximal and distal perspective. Thus, there were 20 different images of scenes (Supplementary Fig. 1). Each set of images was created by first taking a photo from a distal perspective, and then walking in a straight line ~20 feet–~100 feet and taking a photo from this proximal perspective. The camera zoom function was never utilized when generating the stimulus set to ensure that the stimuli did not induce a percept of zooming in/out between the 'near' and 'far' conditions. Two independent groups of participants rated the stimuli as either 'near' or 'far' to ensure that changes in our stimuli were indeed perceived as changes in egocentric distance during the fMRI experiment, and that the stimuli spanned a wide range of distances (Fig. 1B). One group was sitting upright at a computer when making these ratings, while the other group was supine in a mock scanner when making the ratings. The second group was included to make certain that lying down in the scanner did not affect judgments of egocentric distance. Indeed the near/far judgments were highly correlated across groups ($r^2 = .95$, $p < .0001$). Thus, we can conclude that changes in egocentric distance in the stimuli are perceived as nearer or farther away both while sitting upright at a computer, and while supine in a scanner. Similarly, we included 10 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. Importantly, the background textures provided depth cues ensuring that participants perceived the objects as proximal and distal, and not simply as small or big. All stimuli were grayscale and $9° \times 7°$ in size. Subjects were instructed to remain fixated on a white cross that was presented on the screen in between each pair of stimuli. Each image was presented at the central fixation and then moved 1° of visual angle either 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 particularly chosen to eliminate any early retinotopic confounds, and to further disrupt the potential perception of navigating through the scene.

2.3. *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 sec), and the experimental scans (repetition time = 1 sec). For all scans: echo time = 30 msec; voxel size = 3.1 × 3.1 × 4.0 mm with a .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.

2.4. 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, Bannister, Brady, & Smith, 2002). Additionally, localizer data, but not experimental data, were spatially smoothed (6 mm kernel), as described previously (e.g., Dilks et al., 2011), detrended, and fit using a double-gamma function. However, we also analyzed the experimental data after spatially smoothing with a 6 mm kernel, and the overall results did not change. 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) – following the method of Epstein and Kanwisher (1998) (Fig. 2). PPA was identified bilaterally in all 28 participants, RSC was identified in the right hemisphere in all 28 participants, and in the left hemisphere in 26 participants, and OPA was identified bilaterally in 26 participants. As a control region, we also functionally defined a bilateral foveal confluence (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 28 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 (Linsley & MacEvoy, 2014; MacEvoy & Yang, 2012; Persichetti, Aguirre, & Thompson-Schill, 2015). For each ROI

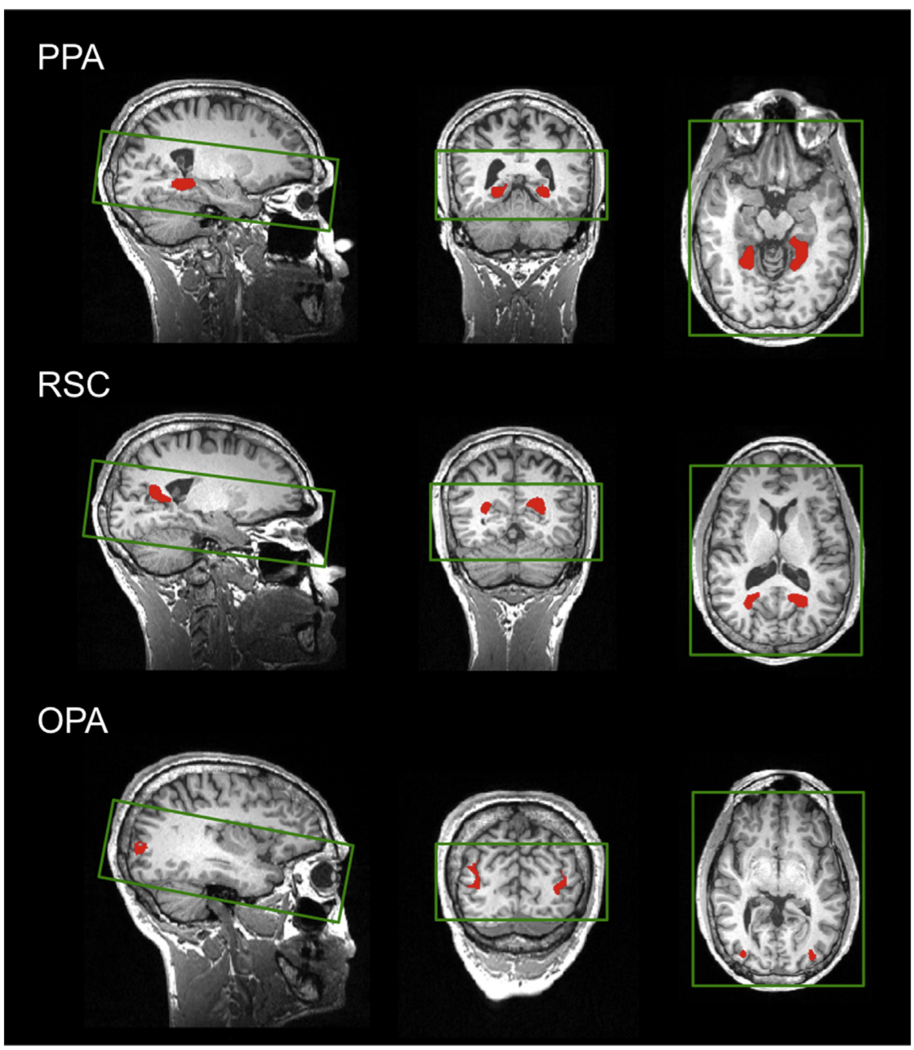

Fig. 2 – Scene-selective ROIs from an example participant. Using independent data, PPA, RSC, and OPA (shown in red) were localized as regions that responded more strongly to scenes than objects ($p < 10^{-4}$). The green rectangles overlaid on the brain images show the slice prescription used in the experiment (16 slices). Note that the slice prescription covered the entirety of the occipital and temporal lobes in all subjects, and thus we were able to capture all of the ROIs.

of each participant, the mean time courses (percentage signal change relative to a baseline fixation) for the experimental conditions (i.e., Same, Different, Distance) were extracted across voxels.

Next, the Same and Different condition time courses were separately averaged across the scene-selective ROIs and participants to identify an average response across 'peak' time points. 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 sec to 9 sec after trial onset (all *p* values < .05). Conversely, none of the time points before 5 sec or after 9 sec showed this adaptation effect (all *p* values > .50). Finally, for each participant, these average responses for each scene-selective ROI were then extracted for each condition (Different, Distance, Same), and repeated-measures ANOVAs were performed on each.

A 3 (ROI: PPA, OPA, RSC) × 3 (condition: Different, Distance, Same) × 2 (hemisphere: Left, Right) repeated-measures ANOVA was conducted. We found no significant ROI × condition × hemisphere interaction at the average response [$F_{(4,96)} = .97$, $p = .43$, $\eta_p^2 = .04$]. Thus, both hemispheres were collapsed for further analyses.

3. Results

As predicted, 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 on the average response from 5 sec to 9 sec (see Methods for details) revealed a significant main effect of condition [$F_{(2,52)} = 6.04$, $p < .005$, $\eta_p^2 = .19$], with a significantly greater response to the Different condition compared to the Same condition (main effect contrast, $p < .001$, $d = .91$), and a marginally significant difference between the Distance and Same conditions (main effect contrast, $p = .05$, $d = .47$). There was no significant difference between the Distance and Different conditions (main effect contrast, $p = .17$, $d = .40$) (Fig. 3). 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 is sensitive to changes in egocentric distance information in images of scenes.

Similarly, for OPA, a 3 level (condition: Different, Distance, Same) repeated measures ANOVA on the average response revealed a significant main effect of condition [$F_{(2,50)} = 7.93$, $p < .001$, $\eta_p^2 = .24$], with a significantly greater response to the Different condition compared to the Same condition (main effect contrast, $p < .001$, $d = 1.05$), and a significant difference between the Distance and Same conditions (main effect contrast, $p < .05$, $d = .62$). There was no significant difference between the Distance and Different conditions (main effect contrast, $p = .12$, $d = .43$) (Fig. 3). 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 is sensitive to changes in egocentric distance information in images of scenes.

By contrast, PPA was not sensitive to egocentric distance information in images of scenes. A 3 level (condition: Different, Distance, Same) repeated-measures ANOVA on the average response revealed a significant main effect of condition [$F_{(2,54)} = 11.36$, $p < .001$, $\eta_p^2 = .30$], with a significantly

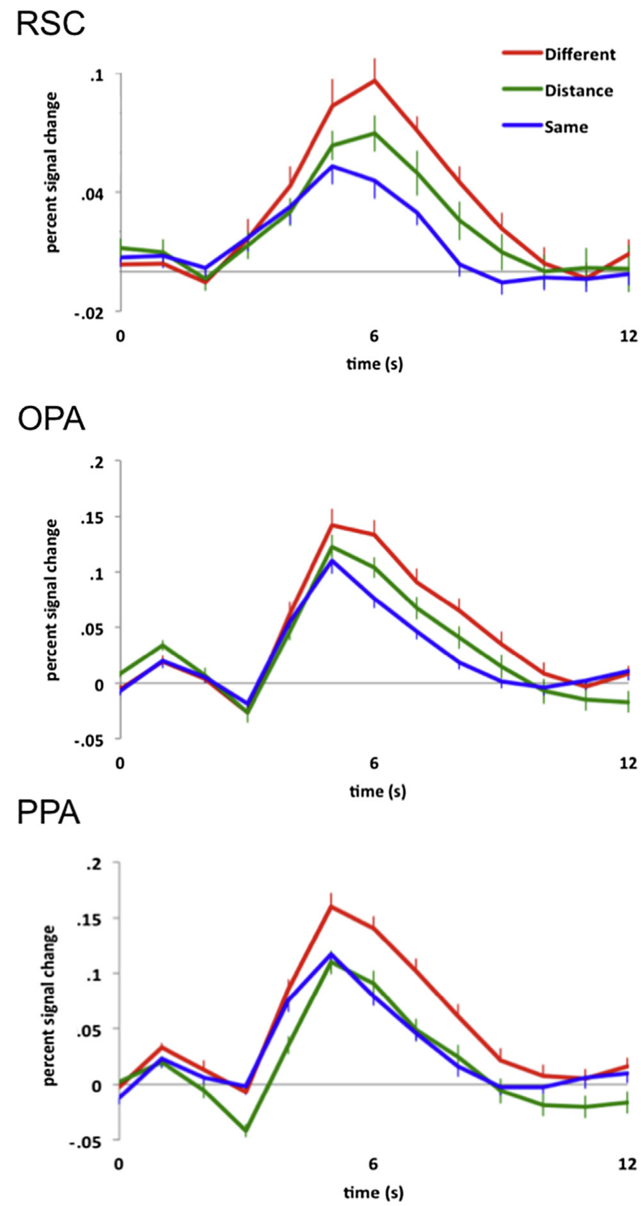

Fig. 3 – Hemodynamic time courses (percentage signal change) of three scene selective regions of cortex, RSC, OPA, and PPA to (1) two completely different images of scenes (red line labeled "Different"), (2) the same image of a scene presented twice (blue line labeled "Same"), and (3) an image of a scene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

greater response to the Different condition compared to either the Same or Distance conditions (main effect contrasts, both p values $< .005$, both d's $> .90$), and no significant difference between the Distance and Same conditions (main effects contrast, $p = .73$, d $= .08$) (Fig. 3).

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 average responses for two different images of scenes and the same images (i.e., expected adaptation) was compared to the difference between the average responses for proximal versus distal images and the same images (Fig. 4). Crucially, a 3 (ROI: OPA, RSC, PPA) × 2 (difference score: Different-Same, Distance-Same) repeated-measures ANOVA revealed a significant interaction [$F_{(2,50)} = 4.35$, $p < .02$, $\eta^2_P = .15$], 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 $< .05$, both $\eta^2_P > .15$). There was not a significant difference in the responses between the RSC and OPA (interaction contrast, $p = .85$, $\eta^2_P = .001$). These results show that the scene selective regions encode egocentric distance information differently: RSC and OPA are sensitive to egocentric distance information in scenes, while PPA is not. To further probe this difference across the three ROIs, we conducted three additional analyses. First, results from paired t-tests comparing the Different-Same and Distance-Same conditions for each ROI independently revealed no significant difference in RSC or OPA [$t_{(26)} = 1.43$, $p = .17$; $t_{(25)} = 1.60$, $p = .12$, respectively], but a significant difference in PPA [$t_{(27)} = 3.49$, $p < .01$]. Second, results from one-sample t-tests comparing the Distance-Same condition to 0 for each ROI independently revealed a significant difference in RSC and OPA [$t_{(26)} = 2.10$; $t_{(25)} = 2.19$, both p-values$<.05$], but no significant difference in PPA [$t_{(27)} = .35$, $p = .73$].

Third, we ran a one-way repeated-measures ANOVA asking whether the signal change for the Distance-Same condition was different across the three regions. Indeed, we found a significant effect [$F_{(2,50)} = 3.43$, $p < .05$], with PPA responding significantly less to Distance-Same compared to RSC and OPA (both p-values $< .02$), and no difference between RSC and OPA ($p = .67$). Taken together, these results demonstrate significant adaptation to egocentric distance information in PPA only, not RSC or OPA.

But might it be the case that the sensitivity to egocentric distance information in images of scenes in RSC or OPA is due to a feed-forward effect from earlier visual areas, rather than indicative 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 the average response to the three conditions in an independently defined region of cortex representing the fovea, and found that 'FC' did not even show fMRI adaptation for Different versus Same scenes (main effect contrast, $p = .30$, d $= .22$), thus confirming that neither OPA nor RSC's sensitivity to egocentric distance information in scenes is due to adaptation in early visual cortex.

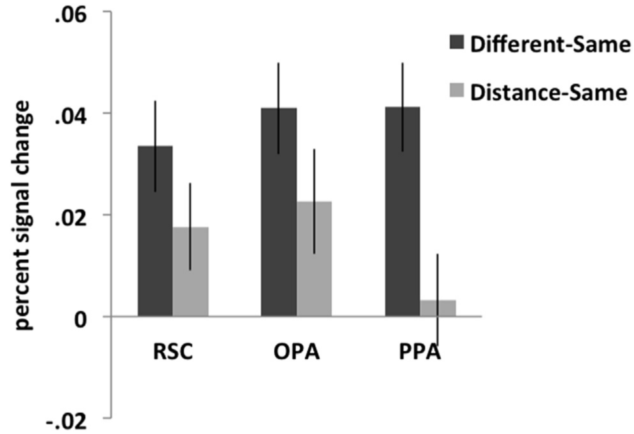

Fig. 4 – For each scene-selective ROI, the difference between the peak responses for two different images of scenes and the same images (labeled "Different-Same") was compared to the difference between the peak responses for two images of the same scene viewed from either a proximal or distal perspective and the opposite version of the same image and the same images (labeled "Distance-Same"). A 3 (ROI: RSC, OPA, PPA) × 2 (difference score: Different-Same, Distance-Same) repeated-measures ANOVA revealed a significant interaction [$F_{(2,50)} = 4.35$, $p < .02$, $\eta^2_P = .15$], 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 the scene-selective regions represent egocentric distance information differently: RSC and OPA are sensitive to egocentric distance information in scenes, while PPA is not sensitive to such information.

Finally, 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 boundaries of the pictures, particularly when the view is close up (Intraub & Richardson, 1989). Consistent with these behavioral data, Park, Intraub, Yi, Widders, and Chun (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 the average response revealed a significant main effect of condition

[$F_{(3,81)} = 5.22$, $p < .005$, $\eta_p^2 = .16$], with a significantly greater response to the Different versus Same condition ($p < .001$, d = .92), demonstrating the expected fMRI adaptation. Crucially, however, there were no significant differences between the Wide-Close, Close-Wide, or Same conditions (p values > .40, both d's < .20). 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) (Intraub & Richardson, 1989), 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).

Given that our stimuli included objects as well as scenes, we were also able to investigate how RSC, OPA, and PPA might respond to changes in egocentric distance information in images of objects (the non-preferred 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 > .15). Thus, the question of sensitivity to egocentric distance information in objects for scene-selective regions is moot.

4. Discussion

The current study asked how scene-selective regions represent egocentric distance information. As predicted, the results demonstrate that the regions of scene-selective cortex are differentially sensitive to perceived egocentric distance information. 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 viewing angle changes across scene images, by a feed-forward effect from earlier visual areas, or by 'boundary extension'.

But, might it be the case that the sensitivity to egocentric distance information in images of scenes in RSC or OPA is simply due to size changes of features in the scenes (i.e., proximal features in the scene subtend larger visual angles than distal features), rather than characteristic of egocentric distance sensitivity to scenes in particular? We do not think this could be the case because as a scene image switches from a distal to proximal perspective (and vice versa) some features of the scene (e.g., a tree or a bridge) increase (or decrease) in size, while other features (e.g., the ground plane or sky) decrease (or increase) in size. Given the varying size changes within each pair of scene images, it seems highly unlikely then that a scene region that simply tracks size (i.e., responding to either 'big' or 'small' features in the scene) would respond, and thus size alone cannot explain the sensitivity to changes in egocentric distance in scenes found in RSC and OPA.

Our finding that PPA is not sensitive to egocentric distance information, one kind of information critical for navigation, provides further evidence challenging the hypothesis that PPA is directly involved in navigation (Cheng & Newcombe, 2005; Epstein & Kanwisher, 1998; Ghaem et al. 1997; Janzen & van Turennout, 2004; Rauchs et al. 2008; Rosenbaum et al. 2004; Spelke et al. 2010). Recall that Dilks et al. (2011) also found that PPA was not sensitive to sense information, another type of information crucial for navigation. Rather, we hypothesize then that human scene processing may be composed of two systems: one responsible for navigation, including RSC and OPA, and another responsible for the recognition of scene category, including PPA. 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 kinds of 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-systems-for-scene-processing hypothesis is correct, then the PPA may contribute to the 'categorization system', while the RSC, OPA, or both may contribute to the 'navigation system'. Indeed, support for this hypothesis comes from two multi-voxel pattern analysis (MVPA) studies demonstrating 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 on a scene categorization task (Walther, Chai, Caddigan, Beck, & Fei-Fei, 2011, Walther, Caddigan, Fei-Fei, & Beck, 2009).

Our hypothesis that the scene processing system may be divided into two systems might sound familiar. For example, Epstein (2008) proposed that human scene processing is divided into two systems – 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-systems-for-scene-processing hypothesis is instead more like the two functionally distinct systems of visual object processing proposed by Goodale and Milner (1992), with one system responsible for recognition, and another for visually-guided action. Note that our hypothesis of two distinct systems for human scene processing – a categorization system including PPA, and a navigation system 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, Beck, & Fei-Fei, 2013; Nasr, Devaney, & Tootell, 2013), suggesting these two regions are functionally (and most likely anatomically) connected, thereby facilitating crosstalk between the two systems.

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 & 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 & Spelke, 1994), necessary for reorientation and navigation (Epstein & 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 (Greene & Oliva, 2009; Oliva & Schyns, 1997; Oliva & Torralba, 2001; Walther et al. 2011). However, spatial layout representation in PPA is only half of the story. A number of recent studies have found that PPA is also 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) (Harel, Kravitz, & Baker, 2013; MacEvoy & Epstein, 2009), 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 & Aminoff, 2003; Bar, Aminoff, & Schacter, 2008), and iii) objects that are large and not portable, thus defining the space around them (e.g., a bed or a couch versus a small fan or a box) (Mullally & Maguire, 2011). Taken together, the above findings are consistent with our idea that PPA may be involved in the recognition of scene category.

In conclusion, we have shown that RSC and OPA are sensitive to egocentric distance information in images of scenes, while PPA is not. This finding coupled with the finding that RSC and OPA are also sensitive to sense information, while PPA is not, suggest that the computations directly involved in navigation do not occur in the PPA. These results are consistent with the hypothesis that there exist two distinct systems for processing scenes: one for navigation, including RSC and OPA, and another for the recognition of scene category, including PPA. 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. Furthermore, the current study does not distinguish the precise roles of RSC and OPA in navigation. It is possible that RSC and OPA may both be involved in navigation more generally, but support different functions within navigation. Specifically, OPA may be involved in navigating the local visual environment (Kamps, Julian, Kubilius, Kanwisher, & Dilks, in press), while RSC is more involved in more complex forms of navigation (i.e., orienting the individual to the broad environment) (Marchette, Vass, Ryan, & Epstein, 2014; Vass & Epstein, 2013). Ongoing studies are investigating this possibility.

Acknowledgments

We would like to thank the Facility for Education and Research in Neuroscience (FERN) Imaging Center in the Department of Psychology, Emory University, Atlanta, GA. We also thank Frederik S. Kamps and Samuel Weiller for insightful conversation and technical support, as well as Russell Epstein and the members of his lab for helpful feedback. This work was supported by Emory College, Emory University (DD) and National Eye Institute grant T32 EY007092-28 (AP). The authors declare no competing financial interests.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cortex.2016.02.006.

REFERENCES

Baldassano, C., Beck, D. M., & Fei-Fei, L. (2013). Differential connectivity within the parahippocampal place area. *NeuroImage, 75*, 228–237.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617–629.

Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron, 38*, 347–358.

Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *The Journal of Neuroscience, 28*, 8539–8544.

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (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, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review, 12*, 1–23.

Dale, A. M., Greve, D. N., & Burock, M. A. (1999). Stimulus sequences for event-related fMRI. *Human Brain Mapping, 8*, 109–114.

Dilks, D. D., Julian, J. B., Kubilius, J., Spelke, E. S., & Kanwisher, N. (2011). Mirror-image sensitivity and invariance in object and scene processing pathways. *The Journal of Neuroscience, 31*, 11305–11312.

Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *The Journal of Neuroscience, 33*, 1331–1336a.

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision, 3*, 586–598.

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

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences, 12*, 388–396.

Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstable avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 343–362.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., et al. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *NeuroReport, 8*, 739–744.
Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*, 20–25.
Gouteux, S., Thinus-Blanc, C., & Vauclair, J. (2001). Rhesus monkeys use geometric and nongeometric information during a reorientation task. *Journal of Experimental Psychology: General, 130*, 505–519.
Gray, E. R., Spetch, M. L., Kelly, D. M., & Nguyen, A. (2004). Searching in the Center: pigeons (Columba livid) encode relative distance from walls of an enclosure. *Journal of Comparative Psychology, 118*, 113–117.
Greene, M., & Oliva, A. (2009). Recognition of natural scenes from global properties: seeing the forest without representing the trees. *Cognitive Psychology, 58*, 137–176.
Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology, 13*, 159–166.
Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica, 107*, 293–321.
Harel, A., Kravitz, D. J., & Baker, C. I. (2013). Deconstructing visual scenes in cortex: gradients of object and spatial layout information. *Cerebral Cortex (New York, N.Y.: 1991), 23*, 947–957.
Hermer, L., & Spelke, E. S. (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. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 179–187.
Janzen, G., & van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience, 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.
Kamps F.S., Julian J.B., Kubilius J., Kanwisher N. and Dilks D.D., The occipital place area represents the local elements of scenes, NeuroImage, (in press).
Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293*, 1506–1509.
Linsley, D., & MacEvoy, S. P. (2014). Evidence for participation by object-selective visual cortex in scene category judgments. *Journal of Vision, 14*, 19–19.
MacEvoy, S. P., & Epstein, R. A. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Current Biology, 19*, 943–947.
MacEvoy, S. P., & Yang, Z. (2012). Joint neuronal tuning for object form and position in the human lateral occipital complex. *NeuroImage, 63*, 1901–1908.
Maguire, E. A. (2001). The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology, 42*, 225–238.
Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience, 17*, 1598–1606.
Mullally, S. L., & Maguire, E. A. (2011). A new role for the parahippocampal cortex in representing space. *Journal of Neuroscience, 31*, 7441–7449.
Nasr, S., Devaney, K. J., & Tootell, R. B. H. (2013). Spatial encoding and underlying circuitry in scene-selective cortex. *NeuroImage, 83*, 892–900.
Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34*, 72–107.
Oliva, A., & Torralba, A. (2001). Modeling the shape of a scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision, 42*, 145–175.
Park, S., Intraub, H., Yi, D.-J., Widders, D., & Chun, M. M. (2007). Beyond the edges of a view: boundary extension in human scene-selective visual cortex. *Neuron, 54*, 335–342.
Persichetti, A. S., Aguirre, G. K., & Thompson-Schill, S. L. (2015). Value is in the eye of the beholder: early visual cortex codes monetary value of objects during a diverted attention task. *Journal of Cognitive Neuroscience, 27*, 893–901.
Rauchs, G., Orban, P., Balteau, E., Schmidt, C., Degueldre, C., Luxen, A., et al. (2008). Partially segregated neural networks for spatial and contextual memory in virtual navigation. *Hippocampus, 18*, 503–518.
Rosenbaum, R. S., Ziegler, M., Winocur, G., Grady, C. L., & 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. *Robotics and Autonomous Systems, 16*, 213–245.
Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*, 143–155.
Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage, 23*(Suppl. 1), S208–S219.
Sovrano, V. A., 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, S. A., & Izard, V. (2010). Beyond core knowledge: natural geometry. *Cognitive Sciences, 34*, 863–884.
Vass, L. K., & Epstein, R. A. (2013). Abstract representations of location and facing direction in the human brain. *Journal of Neuroscience, 33*, 6133–6142.
Walther, D. B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience, 29*, 10573–10581.
Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2011). Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 9661–9666.
Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences, 6*, 376–382.
Wehner, Michel, & Antonsen. (1996). Visual navigation in insects: coupling of egocentric and geocentric information. *Journal of Experimental Biology, 199*, 129–140.