Refreshing and Integrating Visual Scenes in Scene-selective Cortex

Soojin Park[1], Marvin M. Chun[2], and Marcia K. Johnson[2]

Abstract

■ Constructing a rich and coherent visual experience involves maintaining visual information that is not perceptually available in the current view. Recent studies suggest that briefly thinking about a stimulus (*refreshing*) can modulate activity in category-specific visual areas. Here, we tested the nature of such perceptually refreshed representations in the parahippocampal place area (PPA) and retrosplenial cortex (RSC) using fMRI. We asked whether a refreshed representation is specific to a restricted view of a scene, or more view-invariant. Participants saw a panoramic scene and were asked to think back to (refresh) a part of the scene after it disappeared. In some trials, the refresh cue appeared twice on the same side (e.g., refresh left–refresh left), and other trials, the refresh cue appeared on different sides (e.g., refresh left–refresh right). A control condition presented halves of the scene twice on same sides (e.g., perceive left–perceive left) or different sides (e.g., perceive left–perceive right). When scenes were physically repeated, both the PPA and RSC showed greater activation for the different-side repetition than the same-side repetition, suggesting view-specific representations. When participants refreshed scenes, the PPA showed view-specific activity just as in the physical repeat conditions, whereas RSC showed an equal amount of activation for different- and same-side conditions. This finding suggests that in RSC, refreshed representations were not restricted to a specific view of a scene, but extended beyond the target half into the entire scene. Thus, RSC activity associated with refreshing may provide a mechanism for integrating multiple views in the mind. ■

INTRODUCTION

Constructing a rich and coherent visual experience involves maintaining visual information that is not perceptually available in the current view. The visual information a person can see at one time is extremely limited, so we constantly move our eyes to perceive the whole environment. At the same time, we need to somehow maintain information that was just present on our retina a moment earlier but is currently not part of the perceptual stimulus. Once visual input disappears, keeping a perceptual representation intact and updated depends on higher-level operations rather than a simple bottom–up input (Hochberg, 1978). There is neuroimaging evidence that activity in visual associative areas, such as those that respond selectively to scenes (Epstein & Kanwisher, 1998; Aguirre, Detre, Alsop, & D'Esposito, 1996), is modulated by higher-level cognition such as working memory, mental imagery, or attention (Johnson, Mitchell, Raye, D'Esposito, & Johnson, 2007; Gazzaley, Cooney, McEvoy, Knight, & D'Esposito, 2005; Ranganath & D'Esposito, 2005; Yi & Chun, 2005; Yi, Woodman, Widders, Marois, & Chun, 2004; Curtis & D'Esposito, 2003; O'Craven & Kanwisher, 2000; Wojciulik, Kanwisher, & Driver, 1998).

A cognitive framework that highlights the interface between bottom–up and top–down processes is the multiple entry, modular (MEM) memory model (Johnson et al., 2005; Johnson & Hirst, 1991; Johnson, 1983). The MEM model assumes that there are two cognitive systems: perceptual and reflective. Thinking back to a representation that was activated just a moment ago is termed *refreshing*, and it is a component process in the reflective system. Refreshing, like rehearsing, is a component of working memory, but rehearsing usually involves maintaining multiple items over longer intervals (e.g., several seconds), whereas refreshing involves reflecting on a specific item for a very brief period within a second of its offset. Such momentary refreshing allows a representation to be foregrounded relative to other active items, giving it an advantage during ongoing tasks. Although refreshing may involve representing the physical appearance of an item, hence, imagery, it can be distinguished from mental imagery tasks that are more complex. For example, typical imagery tasks require the construction of a mental image from verbal descriptions (O'Craven & Kanwisher, 2000; Mellet et al., 1996) or the retrieval of information from long-term memory (Kosslyn, Thompson, Sukel, & Alpert, 2005; Ishai, Haxby, & Ungerleider, 2002; Mellet, Petit, Mazoyer, Denis, & Tzourio, 1998). That is, both working memory and mental imagery may involve refreshing, but working memory and mental imagery tasks frequently involve additional component processes as well.

Previous studies have found that refreshing activates dorsolateral prefrontal cortex (Raye, Johnson, Mitchell, Reeder, & Greene, 2002), and refreshing visual stimuli can modulate posterior visual areas (Yi, Turk-Browne,

[1]Massachusetts Institute of Technology, [2]Yale University

Chun, & Johnson, 2008; Johnson et al., 2007). For example, in one study, participants saw a scene and a face, and were cued to selectively think back to one of the stimuli after the stimuli had disappeared. Although both the face and the scene were presented for the same amount of time, the fMRI BOLD activity depended on which item participants refreshed: When participants refreshed a scene, the parahippocampal place area (PPA) was more active, and when participants refreshed a face, the fusiform face area (FFA) was more active (Johnson et al., 2007). Moreover, as in physical repetition priming, refreshing a scene facilitated subsequent perceptual processing of a scene (Yi, Turk-Browne, Chun, et al., 2008). Using an fMRI adaptation paradigm, which typically shows lower activity for the second of two perceptually repeated stimuli, Yi et al. tested whether having briefly refreshed a scene also reduced the BOLD response in the PPA when the same scene was viewed again. There was as much fMRI repetition attenuation for scenes that had been previously refreshed as for scenes that had been physically repeated. This finding suggests that the same population of neurons in the PPA may be recruited for bottom–up and top–down representations of scenes. The studies above demonstrate that briefly thinking back to a stimulus can have as powerful an effect on perceptual memory as seeing the physical stimulus again. However, the question remains as to what the nature of these refreshed representations are.

Due to acuity constraints, the amount of visual information around fixation is extremely limited, thus observers need to integrate information from multiple views. The ability to link the current view to prior views of a scene allows observers to form a continuous world representation of space. This ability requires previously perceived information to be available. That is, a perceptual representation of a scene likely depends on high-level operations working in a complementary manner with bottom–up input. Refreshing is a component process that would seem to be well-suited for this purpose. Thus, we investigated the potential role of refreshing in scene representation. More specifically, we asked whether discrete views of scenes are integrated during refreshing of these views. Are these refreshed scene representations specific to the view that people are actively thinking back to, or do they extend beyond this into related views?

The specificity of visual representations has been long debated in the object and scene recognition literature (Epstein, Graham, & Downing, 2003; Tarr, Williams, Hayward, & Gauthier, 1998; Biederman, 1987), and recent neuroimaging studies of scene perception suggest that there are multiple levels of scene representation coexisting in the brain (Park & Chun, 2009; Epstein et al., 2003). However, no study has tested the nature of representations that are refreshed after the visual stimulus is no longer in view. The current study examined the nature of such *post-perceptual, reflective* representations.

To address this question, we tested two scene-specific brain regions. Two brain regions in a scene network have been shown to be involved in scene recognition and spa-

tial navigation: the PPA near the medial temporal lobe and retrosplenial cortex (RSC) near posterior cingulate cortex (Henderson, Larson, & Zhu, 2008; Epstein & Higgins, 2007; Epstein, Higgins, Jablonksi, & Feiler, 2007; Park, Intraub, Widders, Yi, & Chun, 2007; Epstein et al., 2003; Maguire, 2001; Epstein & Kanwisher, 1998; Maguire et al., 1998). For example, Epstein, Higgins, et al. (2007) and Epstein, Parker, et al. (2007) found that during perception, the PPA represents spatial layouts within an individual scene, whereas RSC represents the scene in relation to a broader map of the environment, and also to familiar scenes in memory. A recent study from our lab using panoramic scene presentation showed that the PPA has a viewpoint-specific representation, whereas RSC has a viewpoint-invariant representation that encompasses multiple-related views of a scene (Park & Chun, 2009). Neurological studies and rodent navigation studies also support the idea of dissociated roles of the PPA and RSC (Mendez & Cherrier, 2003; Best, White, & Minai, 2001; Maguire, 2001; Aguirre & D'Esposito, 1999; O'Keefe & Nadel, 1978). Based on these studies of perception, we examined whether reflectively refreshed scenes are also differently represented in the PPA and RSC. More specifically, we examined whether the PPA codes a view-specific representation of a previously refreshed scene, whereas RSC represents a more view-invariant representation of a previously refreshed scene.

To test this idea, we used an fMRI adaptation paradigm, which can reveal overlap in visual representations. If a particular neuronal population treats two stimuli as the same, then the fMRI activation within the region is lower for the same repeated items compared to novel items (Grill-Spector, Henson, & Martin, 2005). For example, if two specific views of a same scene are treated as the same, then the fMRI response in the PPA will show less activity for the repeated view; if two views are treated as different, then the fMRI response will show less or no attenuation for the repeated view (Epstein et al., 2003). In the current experiment, we presented panoramic scenes that were horizontally extended so that they could be divided into a left side and a right side (see Figure 1).

The terms "scene" and "view" are not used consistently in the field (Henderson & Hollingworth, 1999). Here we use these terms as follows: A panoramic "scene" includes a broad space encompassing multiple "views," whereas a view refers to the subset of a scene that is bounded by an aperture. Integration of multiple discrete views into a scene requires cues to link how the views were sampled from a continuous environment. In our study, we presented an entire panoramic scene first, and decomposed it into left and right views that participants clearly understood to be subsets of the scene. In the refresh condition, participants first saw an entire panoramic scene, and were later asked to refresh specific sides only (e.g., refresh the left side). If a refreshed representation is view-specific, refreshing different sides of a scene (e.g., refresh left side and then refresh right side) will be treated differently from each other, and there will be no decrease in the

Figure 1. Experimental conditions. A panoramic scene was presented for 1.5 sec, followed by a 500-msec blank interval. In the refresh same-side condition, participants refreshed the same half side of the scene twice when the white box cue appeared (e.g., refresh left–refresh left; refresh right–refresh right); in the refresh different-side condition, participants refreshed different half sides of the scene (e.g., refresh left–refresh right; refresh right–refresh left). In the perceive same-side condition, participants

viewed the repetition of half scenes on the same sides (e.g., perceive left–perceive left; perceive right–perceive right); in the perceive different-side condition, participants viewed the repetition of half scenes on different sides (e.g., perceive left–perceive right; perceive right–perceive left).
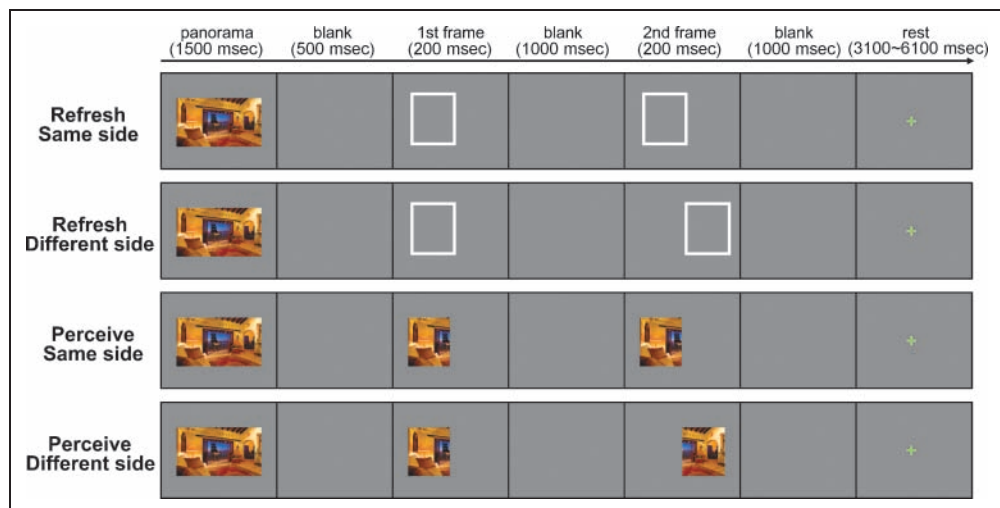
fMRI response for refreshing different sides each once. However, if a refreshed representation is view-invariant, then there should be a decrease in the fMRI response for refreshing different sides (e.g., refresh left side–refresh right side), as expected in refreshing the same side twice (e.g., refresh left side–refresh left side).

In short, the primary aims of this study were to investigate (1) whether refreshing, a postperceptual act of reflection, represents view-specific scene information or view-invariant scene information; and (2) whether the PPA and RSC represent similar or distinct levels of scene information when scenes are refreshed.

METHODS

Experimental Design and Procedure

Eighteen participants (13 women, 1 left-handed, 18–29 years old) from the Yale University community participated for financial compensation. All had normal or corrected-to-normal vision. Informed consent was obtained, and the study protocol was approved by the Yale University Human Investigation Committee at the School of Medicine. An additional four participants were tested, but excluded from the analyses: Two participants were excluded because of excessive head movement, and two other participants were excluded because there were no hemodynamic responses for refresh conditions in either of the PPA and RSC, indicating that they failed to follow instructions.

Participants completed two runs in the scanner with 144 trials per run. Refresh and perceive conditions were intermixed within a run. For all conditions, a panoramic scene (34° × 18°) was presented for 1.5 sec, followed by a 500-msec blank interval (Figure 1). In refresh conditions, participants were then asked to refresh part of

the scene after it disappeared. In half of the refresh trials, a white box cue appeared twice on the same side (e.g., refresh left–refresh left; refresh right–refresh right); in the other half of the refresh trials, the refresh cue appeared on different sides (e.g., refresh left–refresh right; refresh right–refresh left). Each box cue was presented for 200 msec, with 1 sec blank interval between the repetition. Participants were instructed to think back to the side of the scene cued by the box. For example, when they saw a box appear at the right side, they were asked to think back to the right half of the panoramic scene. In the perceive conditions, the left or right half image of the previously shown panoramic scene was presented instead of box cues. As in the refresh conditions, the repetition of half scenes could be on the same sides (e.g., perceive left–perceive left; perceive right–perceive right) or different sides (e.g., perceive left–perceive right; perceive right–perceive left). Participants were instructed to simply view these scenes. Each of these presentations was followed by a green fixation cross that remained for an intertrial interval of 3100, 4600, or 6100 msec. Participants were asked to simply rest when they saw the green fixation. Participants did not make any overt responses during the main experiment, and a surprise memory test was conducted after the scan.

For both the refresh and perceive conditions, there were 12 filler trials that only presented the refresh box cue once or the half image once. For example, in a refresh filler trial, participants saw a panoramic image, a white box cue once either on the left or right side, and a fixation cross for the rest of the remaining trial time. These filler trials were included to reduce the predictability of what would be perceived/refreshed next.

There were a total of 144 panoramic scenes assigned to six sets of 24 used in the refresh same-side condition, the

refresh different-side condition, the perceive same-side condition, the perceive different-side condition, the refresh filler condition, and the perceive filler condition. These six sets were counterbalanced across six participants so that each set appeared equally in each condition. This counterbalancing scheme was used three times. Half of each set were indoor and the other half were outdoor scenes.

After the main experiment, a separate scene localizer was used to localize the PPA and RSC. This run presented six scene blocks and six face blocks that alternated with each other. Each block had 12 images, and participants performed a repetition detection of consecutively repeated images (Park et al., 2007; Yi et al., 2004). Scene stimuli were novel panoramic scenes similar to those presented in the main experiment. For face stimuli, three identical faces were included side-by-side in each display in order to roughly equate the amount of peripheral visual stimulation to panoramic scene stimuli. The PPA and RSC were defined by contrasting brain activity for scene versus face blocks.

After the scan, there was a surprise behavioral scene recognition test. As each of 180 panoramic scenes (144 old from the main experiment and 36 new scenes) was presented, participants responded whether it had appeared in the main experiment or not by responding old, new, or unsure. Old scenes and new scenes were intermixed, and these scenes were counterbalanced so that each scene appeared equally in each condition.

fMRI Data Acquisition

Participants were scanned in a Siemens Trio 3-T scanner with a standard birdcage head coil. Anatomical images were acquired using conventional parameters. Functional images were acquired with a gradient-echo, echo-planar T2* sequence using BOLD contrast. Each functional volume for the main experiment comprised 26 axial slices (1.5 sec repetition time; 25 msec echo time; 90° flip angle; 5 mm thickness with no gap) acquired parallel to the anterior commissure–posterior commissure line. The main experiment was conducted in two functional scan runs, each acquiring 429 image volumes. A separate scene localizer followed to localize the PPA and RSC (130 image volumes, 2 sec repetition time). The first five image volumes of each functional scan were discarded to allow for T1 equilibration effects. Stimuli were presented through an LCD projector on a rear-projection screen.

fMRI Data Analysis

Image preprocessing and statistical analyses were conducted using SPM2 (Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK). The first five volumes of each run were discarded due to magnetization equilibration. The remaining volumes were then corrected for slice timing and realigned. The average realigned image was normalized to an EPI template in Montreal Neurological Institute space (resampling voxel size = 3 × 3 × 3 mm), then motion-corrected and smoothed (Gaussian kernel = 8 × 8 × 8 mm). A high-pass frequency filter (cutoff = 128 sec period) and autocorrelation correction were applied to the time series. The same slice properties and preprocessing procedures were used for the main and the localizer runs.

A linear contrast ($p < .0001$, uncorrected; cluster threshold = 5 voxels) was used to identify clusters of voxels in bilateral ventral visual areas that responded significantly more to scenes than to faces using the scene block versus face block contrast in the individual localizer. Median number of voxels (3 × 3 × 3 mm) for the PPA clusters were 132 voxels (LPPA) and 119 voxels (RPPA), and median number of voxels for RSC clusters were 55 voxels (LRSC) and 66 voxels (RRSC). A maximally scene-selective voxel was localized for each hemisphere: the PPA (average Talairach coordinates: −25, −46, 0; 26, −43, −1) and RSC (−13, −56, 18; 18, −54, 17). To restrict ambiguous boundaries of the functionally defined PPA and RSC, we used SPM Marsbar to construct boxes that were within 2 standard deviations of the mean peak voxel of all our past participants ($n = 46$). Peak voxels were selected for each individual participant in the present study within these defined regions. The peak voxels were typically around the center of mass for the participant's activation: the PPA center of mass (average Talairach coordinates: −24, −45, −3; 26, −46, −4) and RSC center of mass (−14, −56, 17; 18, −55, 17), and there was a high degree of consistency in the anatomical locations across participants. All six conditions (refresh same side, refresh different side, perceive same side, perceive different side, refresh filler, perceive filler) were modeled using 13 finite impulse response (FIR) functions with six motion parameters as covariates of no interest. For each ROI of each participant, the mean time courses for the six main events were extracted across voxels. To determine the time point to include in the ANOVA, the time courses were averaged across conditions and hemispheres, and the numerical peak was compared to each of the other time points. After the comparison of time points, the peak of the PPA response included both time points 6 and 7.5 sec after the trial onset. The peak of RSC response included only the time point at 6 sec, which was significantly different from the others, including the time point at 7.5 sec (t test, $p < .05$ one-tailed) (Marois, Yi, & Chun, 2004; Epstein et al., 2003). Exploratory whole-brain analyses were conducted to examine refresh-related activity. The fMRI data were modeled with an HRF including time derivatives, and used as regressors in a multiple regression analysis, along with six movement parameter regressors. A refresh > perceive contrast was specified a priori. Both the refresh same-side and the refresh different-side conditions were included to form the overall "refresh" condition. Likewise, both the perceive same-side and the perceive different-side conditions were included to form the overall "perceive" condition. This within-subject contrast was

combined into a group random effects analyses using SPM2 with a threshold of $p < .001$ (uncorrected; cluster threshold = 5 voxels). The threshold adopted follows prior convention for exploratory whole-brain analyses in related studies (Yi, Turk-Browne, Flombaum, et al., 2008; Epstein, Higgins, et al., 2007; Johnson et al., 2007).

Due to constraints in timing for refreshing (refreshing should immediately follow the previously presented scene) and constraints in deconvolving each event in the fMRI BOLD signal, we compared the overall activation of the different-side condition (e.g., refresh left side–refresh right side) to the overall activation of the same-side condition (e.g., refresh left side–refresh left side). If a refreshed representation is view-specific, then the overall activation for the different-side condition will be greater than the same-side condition. If a refreshed representation is view-invariant, then the overall activation for the different-side condition will be equal to the same-side condition because there will be attenuation for the two views that constitute a single scene.

RESULTS

ROI-based Analyses

A repeated measures ANOVA and paired t tests between conditions were performed on the averages of the peak amplitude responses in the PPA and RSC ROIs. The main effect of hemisphere was marginally significant in the PPA [$F(1, 17) = 4, p < .07$], suggesting that the right PPA had larger activation than the left PPA. However, hemisphere did not interact with any combination of other main factors of interest (all $Fs < 2, p > .2$), thus, both hemispheres were collapsed for analysis. A repeated measures ANOVA with region (PPA, RSC), condition (refresh, perceive), and repetition type (same side, different side) revealed a significant main effect of region [$F(1, 17) = 83.9, p < .001$] and repetition type [$F(1, 17) = 13.3, p < .005$]. There was a significant interaction between condition and repetition type, suggesting different patterns of scene representation in the refresh and perceive conditions [$F(1, 17) = 19, p < .001$]. Furthermore, there was a significant three-way interaction of region, condition, and repetition type [$F(1, 17) = 13, p < .005$], suggesting refresh and perceive activation patterns differed across regions.

To examine these effects more specifically, a repeated measures ANOVA with condition (refresh, perceive) and repetition type (same side, different side) was conducted separately for the PPA and RSC. Figure 2 illustrates mean peak hemodynamic responses for refresh same-side, refresh different-side, perceive same-side, and perceive different-side conditions in each of the ROIs. In the PPA (Figure 2A), when same-side scenes were perceived, the activation was significantly attenuated compared to when different-side scenes were perceived [$t(17) = 3.7, p < .005$]. When participants refreshed scenes, the activation in the PPA was very similar to when participants actually perceived the

physical repetition. The refresh same-side condition was significantly attenuated compared to the refresh different-side condition [$t(17) = 2.5, p < .05$]. There was no difference in the amount of attenuation between the refresh and perceive conditions [$F(1, 17) < 1, p = .76$]. This lack of interaction suggests that participants were not only able to specifically refresh the instructed half side of a scene, but furthermore, the refreshing of a specific side a second time activated the same specific neuronal population that was activated in refreshing it the first time. Yi, Turk-Browne, Chun, et al. (2008) previously showed that the same neuronal population was involved in refreshing a stimulus and in perceptually seeing the stimulus. The current study extends those findings by showing that multiple refreshing of the same stimulus will target the same neuronal properties in the PPA.

In contrast to the PPA, RSC showed a significant interaction between condition and repetition type [$F(1, 17) = 24, p < .001$], indicating different patterns in the refresh and perceive conditions in RSC. When scene images were perceived on the same side, the activation was significantly attenuated compared to when scene images were perceived on different sides [$t(17) = 5.3, p < .001$], as in the PPA perceive condition. However, in contrast to the perceive condition, there was no difference in the amount of activity between same-side and different-side conditions when participants refreshed scenes [$t(17) < 1, p = .71$]. Refreshing different sides looked similar to refreshing the same side (Figure 2B). Most importantly, compared to the perceive condition, there was a relatively reduced activation in the refresh different-side condition [$t(17) = 2, p = .06$], not an increase in the refresh same-side condition compared to the perceive same-side condition [$t(17) = 1.4, p = .17$]. Thus, when people refreshed different sides, there was attenuation for the repetition similar to when people refreshed (or perceived) the same side.

There was an interaction across region and repetition type within the refresh condition [$F(1, 17) = 10, p < .05$]. This interaction, along with the overall three-way interaction [Region × Condition × Repetition type: $F(1, 17) = 13, p < .005$], suggests that there are different levels of representation for refreshed scenes in the PPA and RSC: The PPA can refresh information restricted to a specific part of a scene, whereas RSC refreshes information that evidently is less restricted to a specific part of a scene.

Refresh-related Activity

To examine refresh-related activity in the whole brain, we contrasted the overall refresh condition (including both the refresh same-side and refresh different-side conditions) to the overall perceive condition (including both the perceive same-side and perceive different-side conditions). As shown in Table 1, when people refreshed scenes, there was greater activation in frontal areas, including superior frontal, middle frontal gyrus, and inferior frontal gyrus. These results are consistent with prior studies of refreshing, which
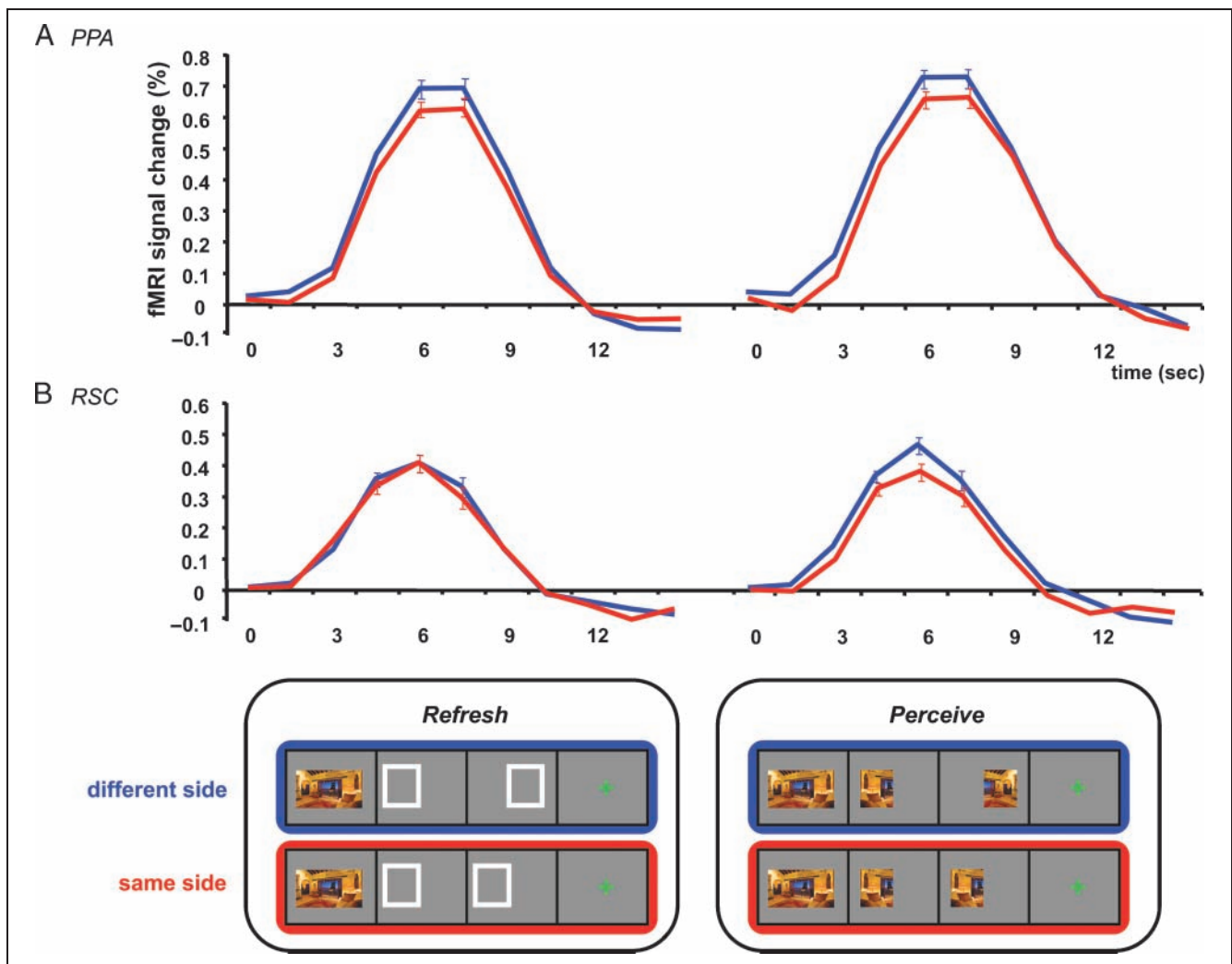
Figure 2. Mean peak hemodynamic responses for refresh same-side, refresh different-side, perceive same-side, and perceive different-side conditions are shown for each ROI. Error bars indicate ±*SEM*. (A) Parahippocampal place area (PPA). Paired comparisons between refresh same-side and refresh different-side conditions and between perceive same-side and perceive different-side conditions both showed a significant difference indicating greater attenuation in the same versus different conditions. (B) Retrosplenial cortex (RSC). Paired comparisons between refresh same-side and refresh different-side conditions showed no difference between conditions. Paired comparisons between perceive same-side and perceive different-side conditions showed significant attenuation, resulting in a significant interaction between condition and repetition type in RSC.

also found similar regions of prefrontal cortex (Johnson et al., 2005, 2008; Yi, Turk-Browne, Chun, et al., 2008; Raye et al., 2002).

Subsequent Scene Recognition

When participants came out of the scanner, they performed a behavioral scene recognition test. A panoramic scene was presented and participants indicated whether the scene was "old," "new," or if they were "unsure." As shown in Figure 3, participants correctly responded to most old scenes as "old," and most novel scenes as "new" or "unsure." Only 15% of the novel scenes were falsely recognized as old (false alarm). There was no difference in recognition rate of old scenes depending on whether the scene was refreshed or perceived during the main experiment, or whether the

scene was presented in the same-side condition or in the different-side condition (all $ts < 1$, $p > .5$).

DISCUSSION

The current study asked whether different views of scenes are integrated during refreshing in the PPA and RSC. The results suggest that the PPA refreshed a view-specific representation of a previously seen image, whereas RSC refreshed a relatively view-invariant representation. That is, the pattern of results indicates that the RSC representation during refreshing was not restricted to a specific view, but extended across the scene, presumably permitting an integrated representation of different views from a single panoramic scene. These results support previous suggestions that the PPA and RSC may represent different levels

Table 1. Talairach Coordinates of Areas That Showed Refresh-related Activity in the Refresh > Perceive Contrast (Uncorrected, $p < .001$, $k = 5$)

Brain Region	$k\oplus$	BA	x	y	z	T
Left Hemisphere						
Medial frontal gyrus/anterior cingulate gyrus	115	6/32	−2	17	44	5.75
Superior frontal sulcus/precentral gyrus	27	6/8	−27	−3	49	4.51
Inferior frontal gyrus/insula	87	13/45	−35	15	5	6.08
Middle frontal gyrus/precentral gyrus	34	6/46	−41	−2	37	4.59
Anterior cingulate gyrus	37	32	−8	26	28	5.80
Precuneus	11	7	−9	−62	46	4.45
Putamen/caudate nucleus	19		−15	9	2	4.38
Right Hemisphere						
Inferior frontal gyrus/insula	35	13/45	45	15	4	5.46
Anterior cingulate gyrus	9	24/32	5	26	25	4.04
Cerebellum	45		30	−61	−27	6.78

$k\oplus$ = number of contiguous voxels; BA = Brodmann's area.

of scene information (Epstein & Higgins, 2007; Epstein, Parker, et al., 2007; Park et al., 2007; Bar, 2004; Bar & Aminoff, 2003). The novel finding here is to extend the prior results for perception of scenes to postperceptual representations that are reflectively generated after the stimulus disappeared. Considering the limits of our visual system, maintaining a representation of what was present just a moment earlier beyond what is currently visually available may be an important part of constructing our visual world. When a view shifts from one frame to another, the viewer may construct a continuous perception by refreshing the previous view that slipped away from sight. Such a refreshing process may assist our visual system to smoothly glue the previous view and the current view together over time. Thus, refreshing may be one mechanism

by which perception and reflection interact (e.g., Johnson & Hirst, 1993).

In contrast to the more "panoramic" representation of refreshed scenes in RSC, the specific representation observed in the PPA for the refresh condition suggests that participants were able to specifically refresh the instructed half side of a scene. Less activation found for refreshing the same compared to the different side demonstrates that the same population of neurons was recruited when refreshing a specific side a second time.

Results for the perceive condition are especially interesting in light of a previous experiment (Park & Chun, 2009). Park and Chun (2009) presented three separate views of a single panoramic image, mimicking viewpoint changes of natural environmental scanning. When these panoramic

Figure 3. Scene recognition. Error bars indicate within-subject standard errors of the mean proportions.
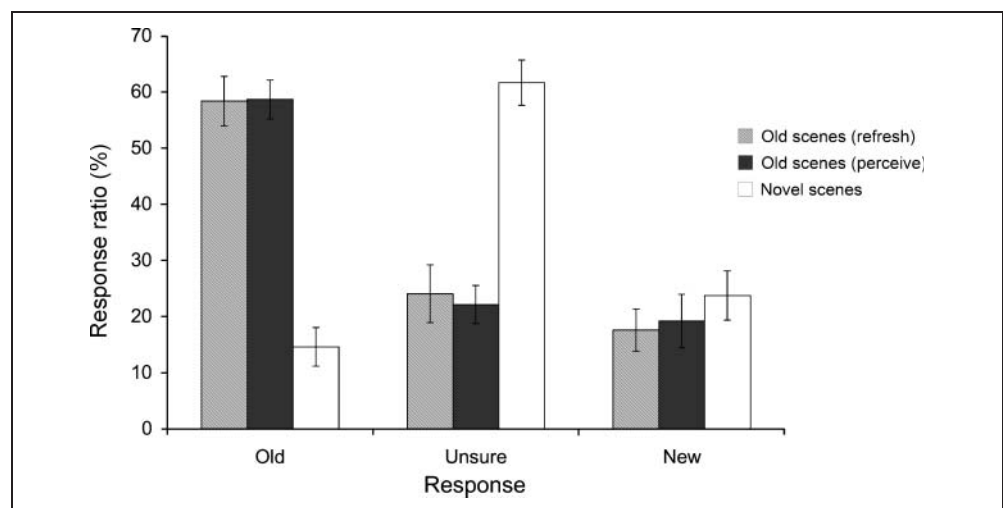
viewpoint changes were presented, corresponding to the present perceive condition, the PPA showed viewpoint specificity, consistent with the current PPA results, whereas RSC showed viewpoint invariance, consistent with the present refresh condition. However, in the current study, RSC showed viewpoint specificity in the perceive condition, partially conflicting with Park and Chun. Differences between the previous and current experiments may explain why we did not find scene integration in RSC for the current perceive condition. In Park and Chun, the different perceive condition consisted of views that overlapped 33%, and moreover, the three overlapping views were presented in panoramic sequence, which made the overlap and the continuity cues across these views very strong. In the current study, when different sides were presented, they did not overlap in physical detail at all. Thus, in the present study, there evidently was not enough overlap across the discrete views for RSC to integrate perceptual scene information in a view-invariant manner. In contrast, RSC demonstrated scene integration in the refresh condition. This highlights a possibly important functional difference between the RSC response during perceiving and refreshing. It may be that when physical stimuli are present, scene construction and integration tends to be largely driven by specific bottom–up, perceptual details, such as the amount of visual overlap in successive glances. In contrast, when views are refreshed, the representation may be less restricted to cued (or intended) boundaries, and thus, may extend to provide overlap cues across refreshed views. Ordinarily, bottom–up and top–down processes may contribute to scene integration, each providing a potential "correction" on limitations of the other. This intriguing possibility deserves further investigation in a systematic study that manipulates the amount of overlap in perceptual and refreshed views. The key finding that one can draw from both studies is that the PPA is generally viewpoint specific, and that RSC is viewpoint invariant when strong perceptual cues for continuity exist for different views, or when distinct views from the same scene are refreshed.

Although there is increasing evidence about the individual functions of human PPA and RSC (Park & Chun, 2009; Epstein, Higgins, et al., 2007; Epstein, Parker, et al., 2007; Epstein & Higgins, 2007), how these two neural regions work in concert to support coherent scene perception is not yet known. Anatomical connectivity data from monkeys suggest that RSC is located such that it might communicate between viewpoint-invariant spatial information in the parietal regions and viewpoint-specific information in the medial temporal regions (which may roughly correspond to human PPA). It is possible that the PPA plays a role in analyzing the geometric structure of the current view, whereas RSC is a hub region that combines such view-specific information from the PPA with top–down information from other areas contributing to the presence of a continuous context across views.

The lack of differences in RSC between refreshing same and different sides suggests that the two types of views are represented similarly, but it is important to rule out the possibility that this result is due to a floor effect— an overall reduced level of response in RSC for refreshing. Arguing against this, and replicating prior studies (Johnson & Johnson, 2009; Johnson et al., 2007), refresh and perceive conditions showed comparable levels of activation, sufficient to support robust differences for different- and same-side manipulations in the perceive condition. The significant increase to different perceptual views in the PPA reflects greater neural involvement in response to novelty. Generally speaking, floor effects have been rejected as a confounding factor in prior studies that manipulated activation strength directly with experimental manipulations (Xu, Turk-Browne, & Chun, 2007; Yi, Kelley, Marois, & Chun, 2006). Thus, it seems unlikely that the null effect for refreshing in RSC can be attributed to an overall low level of activation. Likewise, we do not think that our results can be explained in terms of differences in effort required to refresh a view that was presented a longer time ago in the different- than same-side condition. Although we cannot completely exclude this possibility with our current design, the interaction between the PPA and RSC in the refresh condition suggests that differences in general effort cannot fully explain our overall pattern of results (cf., Xu et al., 2007).

Finally, we should note that to operationalize refreshing and bring it under experimental control, we explicitly cued participants to refresh. However, we assume that in everyday cognition, refreshing takes place spontaneously, with varying degrees of conscious awareness, for example, as people process scenes, sentences, or their own thoughts. In such cases, refreshing is likely guided by an active agenda, or initiated by a gap in stimuli, or by novel information that cues refreshing. Although all of such cases should involve top–down modulation of representational areas, we would expect changes in the relative contributions of different regions in the refresh network, depending on the conditions under which it takes place (e.g., Johnson et al., 2005, Experiment 5; Raye, Mitchell, Reeder, Greene, & Johnson, 2008).

In summary, the current experiment showed that a refreshed scene is represented at multiple coexisting levels in a scene network. Consistent with recent evidence for different roles of the PPA and RSC in scene perception, a refreshed representation in the PPA was specific to a view that a participant was selectively thinking back to, whereas a refreshed representation in RSC was not restricted to a specific view of a scene, but extended into the scene. Such reflective acts of refreshing a view that is no longer present may play an important role in constructing a visual world beyond the view that is currently physically available.

REFERENCES

Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain, 122,* 1613–1628.

Aguirre, G. K., Detre, J. A., Alsop, D. C., & D'Esposito, M. (1996). The parahippocampus subserves topographical learning in man. *Cerebral Cortex, 6,* 823–829.

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.

Best, P. J., White, A. M., & Minai, A. (2001). Spatial processing in the brain: The activity of hippocampal place cells. *Annual Review of Neuroscience, 24,* 459–486.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7,* 415–423.

Epstein, R., Graham, K. S., & Downing, P. E. (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, R. A., & Higgins, J. S. (2007). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cerebral Cortex, 17,* 1680–1693.

Epstein, R. A., Higgins, J. S., Jablonksi, K., & Feiler, A. M. (2007). Visual scene processing in familiar and unfamiliar environments. *Journal of Neurophysiology, 97,* 3670–3683.

Epstein, R. A., Parker, W. E., & Feiler, A. M. (2007). Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *Journal of Neuroscience, 27,* 6141–6149.

Gazzaley, A., Cooney, J. W., McEvoy, K., Knight, R. T., & D'Esposito, M. (2005). Top–down enhancement and suppression of the magnitude and speed of neural activity. *Journal of Cognitive Neuroscience, 17,* 507–517.

Grill-Spector, K., Henson, R., & Martin, A. (2005). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10,* 14–23.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50,* 243–271.

Henderson, J. M., Larson, C. L., & Zhu, D. C. (2008). Full scenes produce more activation than close-up scenes and scene-diagnostic objects in parahippocampal and retrosplenial cortex: An fMRI study. *Brain and Cognition, 66,* 40–49.

Hochberg, J. E. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: Effects of memory and attention revealed by fMRI. *Neuroimage, 17,* 1729–1741.

Johnson, M. K. (1983). A multiple-entry, modular memory system. *The Psychology of Learning and Motivation: Advances in Research and Theory, 17,* 81–123.

Johnson, M. K., & Hirst, W. (1991). Processing subsystems of memory. In R. G. Lister & H. J. Weingartner (Eds.), *Perspectives on cognitive neuroscience* (pp. 197–217). New York: Oxford University Press.

Johnson, M. K., & Hirst, W. (1993). MEM: Memory subsystems as processes. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 241–286). East Sussex, England: Erlbaum.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 339–361.

Johnson, M. R., & Johnson, M. K. (2009). Top–down enhancement and suppression of activity in category-selective extrastriate cortex from an act of reflective attention. *Journal of Cognitive Neuroscience, 21,* 2320–2327.

Johnson, M. R., Mitchell, K. J., Raye, C. L., D'Esposito, M., & Johnson, M. K. (2007). A brief thought can modulate activity in extrastriate visual areas: Top–down effects of refreshing just-seen visual stimuli. *Neuroimage, 37,* 290–299.

Kosslyn, S. M., Thompson, W. L., Sukel, K. E., & Alpert, N. M. (2005). Two types of image generation: Evidence from PET. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 41–53.

Maguire, E. A. (2001). The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology, 42,* 225–238.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: A human navigation network. *Science, 280,* 921–924.

Marois, R., Yi, D.-J., & Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron, 41,* 465–472.

Mellet, E., Petit, L., Mazoyer, B., Denis, M., & Tzourio, N. (1998). Reopening the mental imagery debate: Lessons from functional anatomy. *Neuroimage, 8,* 129–139.

Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., & Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *Journal of Neuroscience, 16,* 6504–6512.

Mendez, M. F., & Cherrier, M. M. (2003). Agnosia for scenes in topographagnosia. *Neuropsychologia, 41,* 1387–1395.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12,* 1013–1023.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.

Park, S., & Chun, M. M. (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., Widders, D., Yi, D. J., & Chun, M. M. (2007). Beyond the edges of a view: Boundary extension in human scene-selective visual cortex. *Neuron, 54,* 335–342.

Ranganath, C., & D'Esposito, M. (2005). Directing the mind's eye: Prefrontal, inferior and medial temporal mechanisms for visual working memory. *Current Opinion in Neurobiology, 15,* 175–182.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a single thought: Dorsolateral PFC activity associated with refreshing just-activated information. *Neuroimage, 15,* 447–453.

Raye, C. L., Mitchell, K. J., Reeder, J. A., Greene, E. J., & Johnson, M. K. (2008). Refreshing one of several active representations: Behavioral and functional magnetic resonance imaging differences between young and older adults. *Journal of Cognitive Neuroscience, 20,* 852–862.

Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience, 1,* 275–277.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology, 79,* 1574–1578.

Xu, Y., Turk-Browne, N. B., & Chun, M. M. (2007). Dissociating task performance from fMRI repetition attenuation in ventral visual cortex. *Journal of Neuroscience, 27,* 5981–5985.

Yi, D.-J., & Chun, M. M. (2005). Attentional modulation of learning-related repetition attenuation effects in human parahippocampal cortex. *Journal of Neuroscience, 25,* 3593–3600.

Yi, D.-J., Kelley, T. A., Marois, R., & Chun, M. M. (2006). Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Research, 1080,* 53–62.

Yi, D.-J., Turk-Browne, N. B., Chun, M. M., & Johnson, M. K. (2008). When a thought equals a look: Refreshing enhances perceptual memory. *Journal of Cognitive Neuroscience, 20,* 1371–1380.

Yi, D.-J., Turk-Browne, N. B., Flombaum, J. I., Scholl, B. J., Kim, M.-S., & Chun, M. M. (2008). Spatiotemporal object continuity in human ventral visual cortex. *Proceedings of the National Academy of Sciences, U.S.A., 105,* 8840–8845.

Yi, D.-J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience, 7,* 992–996.